SENSOR ARRAY SIGNAL PROCESSING

Prabhakar S. Naidu

SENSOR ARRAY SIGNAL PROCESSING

SENSOR ARRAY SIGNAL PROCESSING

Prabhakar S. Naidu

Prologue

An array of sensors is often used in many diverse fields of science and engineering, particularly where the goal is to study propagating wavefields. Some examples are astronomy (radio astronomy), medical diagnosis, radar, communication, sonar, nonrestrictive testing, seismology, and seismic exploration (see [1] for different applications of the array signal processing). The main goal of array signal processing is to deduce the following information through an analysis of wavefields:

- (a) Source localization as in radar, sonar, astronomy, and seismology, etc.
- (b) Source waveform estimation as in communication, etc.
- (c) Source characterization as in seismology
- (d) Imaging of the scattering medium as in medical diagnosis, seismic exploration, etc.

The tools of array signal processing remain the same, cutting across the boundaries of different disciplines. For example, the basic tool of beamformation is used in many areas mentioned above. The present book aims at unraveling the underlying basic principles of array signal processing without a reference to any particular application. However, an attempt is made to include as many tools as possible from different disciplines in an order which reflects the underlying principle.

In the real world, different types of wavefields are used in different applications, for example, acoustic waves in sonar, mechanical waves in seismic exploration, electromagnetic waves in radar and radio astronomy. Fortunately, all wavefields can be characterized under identical mathematical framework. This common mathematical framework is briefly summarized in chapter 1. Here we have described the basic equations underlying different wavefields and the structure of array signals and the background noise when the noise sources follow some simple geometrical distribution. The topics covered are wavefield in open space, bounded space including multipath propagation and layered medium. Also covered is the weak scattering phenomenon which is the basis for tomographic imaging. In chapter 2 we study different types of sensor configurations. The emphasis is however on commonly used uniform linear array (ULA), uniform circular array (UCA). Many practical sensor array systems can be studied in terms of the basic ULA and UCA systems (cylindrical array in radar and sonar, cross array in astronomy and seismology). Like sensors, the sources can also be configured in the form of an array. The

source array is useful in synthesizing a desired wavefront and/or waveform. In chapter 3 we examine the issues connected with the design of 2D digital filters for wavefield analysis. Since the propagating wavefields possess some interesting spectral characteristics in frequency wavenumber domain, for example, the spectrum of a propagating wavefront is always on a radial line, it is natural to take into account these features in the design of digital filters for separation of interfering wavefields. Specifically, we cover in detail the design of a fan filter and quadrant filter. Also, the classical Wiener filter as an optimum least squares filter is covered in this chapter.

The theme in chapters 4 and 5 is localization of a source. In chapter 4 we describe the classical methods based on the frequency wavenumber spectrum of the observed array output. We start with the Blackman Tukey type frequency wavenumber spectrum and then go on to modern nonlinear high resolution spectrum analysis methods such as Capon's maximum likelihood spectrum which is also known as minimum variance distortionless response (MVDR) beamformer and maximum entropy spectrum. Localization essentially involves estimation of parameters pertaining to the source position, for example, azimuth and elevation angles, range, speed if the source is moving, etc. In the last two decades a host of new methods of source localization have been invented. We elaborate these new approaches in chapter 5. These include subspace based methods, use of man-made signals such as in communication and finally multipath environment. Quite often localization must be done in the real time and it may be necessary to track a moving source. Adaptive techniques are best suited for such tasks. A brief discussion on adaptive approach is included. In chapter 6 we look into methods for source waveform separation and estimation. The direction of arrival (DOA) is assumed to be known or has been estimated. We shall describe a Wiener filter which minimizes the mean square error in the estimation of the desired signal coming from a known direction and a Capon filter which, while minimizing the power, ensures that the desired signal is not distorted. We also talk about the estimation of direction of arrival in a multipath environment encountered in wireless communication.

The next two chapters are devoted to array processing for imaging purposes. Firstly, in chapter 7 we look at different types of tomographic imaging systems: nondiffracting, diffracting and reflection tomography. The received wavefield is inverted under the assumption of weak scattering to map any one or more physical properties of the medium, for example, sound speed variations in a medium. For objects of regular shape, scattering points play an important role in geometrical diffraction theory. Estimation of these scattering points for the determination of shape is also discussed. In chapter 8 we study the method of wavefield extrapolation for imaging, extensively used in seismic

exploration. The raw seismic traces are stacked in order to produce an output trace from a hypothetical sensor kept close to the source (with zero- offset). A suite of such stacked traces may be modeled as a wavefield recorded in an imaginary experiment wherein small charges are placed on the reflector and exploded at the same time. The zero-offset wavefield is used for imaging of reflectors. The imaging process may be looked upon as a downward continuation of the wavefield or inverse source problem or propagation backward in time, i.e., depropagation to the reflector. All three view points are very briefly described.

      The book is based on a course entitled "Digital Array Processing" offered to the graduate students who had already taken a course on digital signal processing (DSP) and a course on modern spectrum analysis (MSA). It has been my conviction that a student should be exposed to all basic concepts cutting across the different disciplines without being burdened with the questions of practical applications which are usually dealt with in specialty courses. The most satisfying experience is that there is a common thread that connects seemingly different tools used in different disciplines. An example is beamformation, a commonly used tool in radar/sonar, which has a close similarity with stacking used in seismic exploration. I have tried to bring out in this exposition the common thread that exists in the analysis of wavefields used in a wide variety of application areas. The proposed book has a significantly different flavor, both in coverage and depth in comparison with the ones on the market [1-5]. The first book, edited by Haykin, is a collection of chapters, each devoted to an application. It rapidly surveys the state of art in respective application areas but does not go deep enough and describe the basic mathematical theory required for the understanding of array processing. The second book by Ziomek is entirely devoted to array signal processing in underwater acoustics. It covers in great depth the topic of beamformation by linear and planar arrays but confines to linear methods. Modern array processing tools do not find a place in this book. The third book by Pillai [3] has a very narrow scope as it deals with in great detail only the subspace based methods. The fourth book by Bouvet and Bienvenu (Eds) is again a collection of papers largely devoted to modern subspace techniques. It is not suitable as a text. Finally, the present book has some similarities with a book by Johnson and Dudgeon [3] but differs in one important respect, namely, it does not cover the application of arrays to imaging though a brief mention of tomography is made. Also, the present book covers newer material which was not available at the time of the publication of the book by Johnson and Dudgeon. During the last two decades there has been intense research activity in the area of array signal processing. There have been at least two review papers summarizing the

new results obtained during this period. The present book is not a research monograph but it is an advanced level text which focuses on the important developments which, the author believes, should be taught to give a broad "picture" of array signal processing.

I have adopted the following plan of teaching. As the entire book cannot be covered in one semester (about 35 hours) I preferred to cover it in two parts in alternate semesters. In the first part, I covered chapter 1 (exclude §1.6), chapter 2, chapters 4, 5 and 6. In the second part, I covered chapter 1, chapter 2 (exclude §2.3), chapter 3 (exclude §3.5), chapters 7 and 8. Exercises are given at the end of each chapter. (The solution guide may be obtained from the publisher).

1. S.Haykin (Ed), *Array Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1985.
2. L. J. Ziomek, *Underwater Acoustics, A Linear Systems Theory*, Academic Press, Orlando, 1985.
3. S. U. Pillai, *Array Signal Processing*, Springer-Verlag, New York, 1989.
4. M. Bouvet and G. Bienvenu, *High Resolution Methods in Underwater Acoustics*, Springer-Verlag, Berlin, 1991.
5. D. H. Johnson and D. E. Dudgeon, *Array Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1993.
6. H. Krim and M. Viberg, Two decades of array signl processing, IEEE Signal Proc. Mag., July 1996, pp. 67-94.
7. T. Chen (Ed) Highlights of statistical signal and array processing, IEEE Signal Proc. Mag., pp. 21-64, Sept. 1998.

Prabhakar S. Naidu   February, 2000
Prof, Dept of ECE,
Indian Institute of Science,
Bangalore 560012, India.

Sensor Array Signal Processing

Contents

Acknowledgment

Dedication

This work is dedicated to the memory of the great visionary , J. R. D Tata who shaped the Indian Institute of Science for many decades.

Chapter One

An Overview of Wavefields

A sensor array is used to measure wavefields and extract information about the sources and the medium through which the wavefield propagates. It is therefore imperative that some background in different types of wavefields and the basic equations governing the wavefield must be acquired for complete understanding of the principles of Array Signal Processing (ASP). In an idealistic environment of open space, homogeneous medium and high frequency (where ray approach is valid), a thorough understanding of the wave phenomenon may not be necessary (those who are lucky enough to work in such an idealistic environment may skip this chapter). But in a bounded inhomogeneous medium and at low frequencies where diffraction phenomenon is dominating, the physics of the waves plays a significant role in ASP algorithms. In this chapter our aim is essentially to provide the basics of the physics of the waves which will enable us to understand the complexities of the ASP problems in a more realistic situation. The subject of wave physics is vast and naturally no attempt is made to cover all its complexities.

§1.1 **Types of Wavefields and Governing Equations:**
The most commonly encountered wavefields are: (i) Acoustic waves including sound waves, (ii) Mechanical waves in solids including vibrations and (iii) Electromagnetic waves including light. The wavefields may be classified into two types, namely, scalar and vector waves. In the scalar wavefield we have a scalar physical quantity that propagates through the space, for example, hydrostatic pressure is the physical quantity in acoustic scalar wavefields. In a vector wavefield, the physical quantity involved is a vector, for example, the displacement vector in mechanical waves, electric and magnetic vectors in electromagnetic waves. A vector has three components all of which travel independently in a homogeneous medium without any exchange of energy. But at an interface separating two different media the components do interact. For example, at an interface separating two solids a pressure wave will produce a shear wave and vice versa. In a homogeneous medium without any reflecting boundaries there is no energy transfer among components. Each component of a vector field then behaves as if it is a scalar field, like an acoustic pressure field.
1.1.1 <u>Acoustic Field</u>: Acoustic field is a pressure (hydrostatic) field. The energy is transmitted by means of propagation of compression and rarefaction waves. The governing equation in a homogeneous medium is given by

$$\nabla^2 \phi = \frac{\rho}{\gamma \phi_0} \frac{d^2 \phi}{dt^2} \qquad (1.1a)$$

where ϕ_0 is ambient pressure, γ is ratio of specific heats at constant pressure and volume and ρ is density. The wave propagation speed is given by

$$c = \sqrt{\frac{\gamma\phi_0}{\rho}} = \sqrt{\frac{\kappa}{\rho}} \tag{1.1b}$$

where κ is compressibility modulus and the wave propagates radially away from the source. In an inhomogeneous medium the wave equation is given by

$$\frac{1}{c^2(\mathbf{r})}\frac{d^2\phi}{dt^2} = \rho(\mathbf{r})\nabla \cdot (\frac{1}{\rho(\mathbf{r})}\nabla\phi)$$
$$= \rho(\mathbf{r})\nabla(\frac{1}{\rho(\mathbf{r})}) \cdot \nabla\phi + \nabla^2\phi \tag{1.2a}$$

After rearranging the terms in (1.2a) we obtain

$$\nabla^2\phi - \frac{1}{c^2(\mathbf{r})}\frac{d^2\phi}{dt^2} = \frac{\nabla\rho(\mathbf{r})}{\rho(\mathbf{r})} \cdot \nabla\phi \tag{1.2b}$$

where $\mathbf{r}$ stands for position vector. The acoustic impedance is equal to the product of density and propagation speed ρc and the admittance is given by the inverse of the impedance or it is also defined in terms of the fluid speed and the pressure,

$$\text{Acoustic admittance} = \frac{fluid\ speed}{pressure} = \frac{\nabla\phi}{j\omega\phi}$$

Note that the acoustic impedance in air is 42 but in water it is 1.53×10^5.

1.1.2 <u>Mechanical Waves in Solids</u>: The physical quantity which propagates is the displacement vector, that is, particle displacement with respect to its stationary position. Let $\mathbf{d}$ stand for the displacement vector. The wave equation in a homogeneous medium is given by [1, p142]

$$\rho\frac{\partial^2\mathbf{d}}{\partial t^2} = (2\mu + \lambda)grad\,div\mathbf{d} - \mu\,curlcurl\mathbf{d}$$

where μ is shear constant and λ is Young's modulus. In terms of these two basic lame constants we define other more familiar parameters:

$$\text{Pressure wave speed:} \quad \alpha = \sqrt{\frac{(2\mu + \lambda)}{\rho}}$$

$$\text{Shear wave speed:} \quad \beta = \sqrt{\frac{\mu}{\rho}}$$

$$\text{Poisson ratio:} \quad \sigma = \frac{\lambda}{2(\mu + \lambda)}$$

$$\text{Bulk modules} \quad \kappa = (\frac{2}{3}\mu + \lambda)$$

The above parameters are observable from experimental data.

A displacement vector can be expressed as a sum of gradient of a scalar function ϕ and curl of a vector function ψ (Helmholtz theorem)

$$\mathbf{d} = \nabla\phi + \nabla \times \psi \tag{1.3a}$$

ϕ and ψ satisfy two different wave equations:

$$\nabla^2\phi = \frac{1}{\alpha^2}\frac{\partial^2\phi}{\partial t^2}$$

$$\nabla \times \nabla \times \psi = -\frac{1}{\beta^2}\frac{\partial^2\psi}{\partial t^2} \qquad \nabla.\psi = 0 \tag{1.3b}$$

where $\nabla \times$ is a curl operator on a vector. The operator is defined as follows:

$$\nabla \times \mathbf{F} = \det\begin{bmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f_x & f_y & f_z \end{bmatrix}$$

$$= (\frac{\partial f_z}{\partial y} - \frac{\partial f_y}{\partial z})\mathbf{e}_x + (\frac{\partial f_x}{\partial z} - \frac{\partial f_z}{\partial x})\mathbf{e}_y + (\frac{\partial f_y}{\partial x} - \frac{\partial f_x}{\partial y})\mathbf{e}_z$$

where $\mathbf{e}_x$, $\mathbf{e}_y$ and $\mathbf{e}_z$ are unit vectors in the direction of x, y, and z, respectively and f_x, f_y and f_z components of vector $\mathbf{F}$. The scalar potential gives rise to longitudinal waves or pressure waves (p-waves) and the vector

potential gives rise to transverse waves or shear waves (s-waves). The p-waves travel with speed α and s-waves travel with speed β. The components of displacement vector can be expressed in terms of ϕ and ψ. From (1.3a) we obtain

$$\mathbf{d} = (d_x, d_y, d_z)$$

$$d_x = \frac{\partial \phi}{\partial x} + \frac{\partial \psi_z}{\partial y} - \frac{\partial \psi_y}{\partial z}$$

$$d_y = \frac{\partial \phi}{\partial y} + \frac{\partial \psi_x}{\partial z} - \frac{\partial \psi_z}{\partial x} \qquad (1.4)$$

$$d_z = \frac{\partial \phi}{\partial z} + \frac{\partial \psi_y}{\partial x} - \frac{\partial \psi_x}{\partial y}$$

where $\psi = (\psi_x, \psi_y, \psi_z)$. In solids we must speak of stress and strain tensors. An element of solid is not only compressed but also twisted while an element of fluid is only capable of being compressed but not twisted. We have to use tensors for characterizing the phenomenon of twisting. We shall define the stress and strain tensors and relate them through Hooke's law. A stress tensor is a matrix of nine components

$$\underline{\underline{S}} = \begin{bmatrix} s_{xx} & s_{yx} & s_{zx} \\ s_{xy} & s_{yy} & s_{zy} \\ s_{xz} & s_{yz} & s_{zz} \end{bmatrix} \qquad (1.5a)$$

The components of the stress tensor represent stresses on different faces of a cube (see fig. 1.1). A strain tensor is given by

$$\underline{\underline{\varepsilon}} = \begin{bmatrix} e_{xx} & e_{yx} & e_{zx} \\ e_{xy} & e_{yy} & e_{zy} \\ e_{xz} & e_{yz} & e_{zz} \end{bmatrix} \qquad (1.5b)$$

The first subscript refers to the plane perpendicular to the axis denoted by the subscript and the second subscript denotes the direction in which the vector is pointing. For example, s_{xx} is a stress in a plane perpendicular to the x-axis (i.e., y-z plane) and pointing along the x-axis.
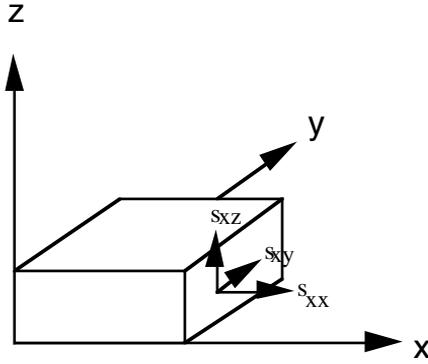
Figure 1.1: An element of volume (cuboid) and stresses are shown on a face perpendicular to the x-axis.

The stress components on different faces of a cuboid are shown in fig. 1.1. The torque on the cuboid should not cause any rotation. For this, we must have $s_{xy} = s_{yx}$ and similarly all other nondiagonal elements in the stress matrix. Thus, $\underline{\underline{s}}$ must be a symmetric matrix. The components of a strain matrix are related to the displacement vector

$$\varepsilon_{xx} = \frac{\partial d_x}{\partial x}, \varepsilon_{yy} = \frac{\partial d_y}{\partial y}, \varepsilon_{zz} = \frac{\partial d_z}{\partial z}$$

$$\varepsilon_{xy} = \varepsilon_{yx} = \frac{\partial d_x}{\partial y} + \frac{\partial d_y}{\partial x};$$

$$\varepsilon_{yz} = \varepsilon_{zy} = \frac{\partial d_y}{\partial z} + \frac{\partial d_z}{\partial y}; \tag{1.6}$$

$$\varepsilon_{zx} = \varepsilon_{xz} = \frac{\partial d_z}{\partial x} + \frac{\partial d_x}{\partial z}.$$

Finally, the stress and strain components are related through Hooke's Law:

$$s_{xx} = \rho\alpha^2\varepsilon_{xx} + \rho(\alpha^2 - 2\beta^2)(\varepsilon_{yy} + \varepsilon_{zz})$$
$$s_{yy} = \rho\alpha^2\varepsilon_{yy} + \rho(\alpha^2 - 2\beta^2)(\varepsilon_{xx} + \varepsilon_{zz})$$
$$s_{zz} = \rho\alpha^2\varepsilon_{zz} + \rho(\alpha^2 - 2\beta^2)(\varepsilon_{xx} + \varepsilon_{yy}) \tag{1.7}$$

$$s_{xy} = s_{yx} = \rho\beta^2\varepsilon_{xy}$$

$$s_{yz} = s_{zy} = \rho\beta^2\varepsilon_{yz}$$

$$s_{zx} = s_{xz} = \rho\beta^2\varepsilon_{zx}$$

Using equations (1.4), (1.6) and (1.7) we can express all nine stress components in terms of the scalar and vector potential functions, ϕ and ψ. For example, it is possible to show that s_{xx} is given by

$$s_{xx} = \rho\alpha^2\frac{\partial^2\phi}{\partial x^2} + \rho(\alpha^2 - 2\beta^2)(\frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2}) + 2\rho\beta^2(\frac{\partial^2\psi_z}{\partial x\partial y} - \frac{\partial^2\psi_y}{\partial x\partial z})$$

A general solution of (1.3b) may be given by

$$\phi(x,y,z,\omega) = \frac{1}{4\pi^2}\int\int_{-\infty}^{\infty}\Phi(u,v,\omega)e^{-\sqrt{u^2+v^2-k_\alpha^2}z}e^{j(ux+vy)}dudv \qquad \text{(1.8a)}$$

$$\psi(x,y,z,\omega) = \frac{1}{4\pi^2}\int\int_{-\infty}^{\infty}\Psi(u,v,\omega)e^{-\sqrt{u^2+v^2-k_\beta^2}z}e^{j(ux+vy)}dudv$$

$$\text{(1.8b)}$$

where $k_\alpha = \frac{\omega}{\alpha}$, $k_\beta = \frac{\omega}{\beta}$. $\Phi(u,v,\omega)$ and $\Psi(u,v,\omega)$ are respectively the Fourier transforms of the displacement potentials ϕ and ψ evaluated on the surface z=0. Furthermore, ψ must satisfy zero divergence condition (1.3b). This will place additional constraints on $\Psi(u,v,\omega)$, namely,

$$ju\Psi_x(u,v,\omega) + jv\Psi_y(u,v,\omega) - \sqrt{u^2+v^2-k_\beta^2}\Psi_z(u,v,\omega) = 0 \quad \text{(1.8c)}$$

Recall that the pressure waves (p-waves) travel at speed α and the shear waves travel at speed β, where α is generally greater than β. The displacement vector is in the direction of the gradient of the scalar potential but it is in the direction of curl of the vector potential (1.4), that is, normal to the vector potential. Thus there is a fundamental difference in the nature of propagation of the shear and the pressure waves. The shear waves are polarized; the displacement vector is always perpendicular to the direction of wave propagation. The displacement vector executes a motion depending upon the phase difference between the components of the displacement vector; a line

when the phase difference is zero, a circular path when the phase difference is 90^0, or a random path when the phase difference is randomly varying. These factors play an important role in the design of sensor array systems and processing of vector potential signals.

1.1.3 <u>Electromagnetic Fields</u>: In electromagnetic fields there are two vectors, namely, electric vector $\mathbf{E}$ and magnetic vector $\mathbf{H}$, each with three components, thus a six component vector field. The basic equations governing the electromagnetic (EM) fields are the Maxwell's equations (in mks units),

$$
\left.\begin{aligned}
\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\
\nabla \cdot \mathbf{D} &= \rho
\end{aligned}\right\} \qquad \text{Faraday's law} \qquad (1.9a)
$$

$$
\left.\begin{aligned}
\nabla \times \mathbf{H} &= \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \\
\nabla \cdot \mathbf{B} &= 0
\end{aligned}\right\} \qquad \text{Ampere's law} \qquad (1.9b)
$$

where

$$
\begin{aligned}
\mathbf{D} &= \varepsilon \mathbf{E} \qquad &\varepsilon&: \text{ dielectric constant} \\
\mathbf{B} &= \mu \mathbf{H} \qquad &\mu&: \text{ magnetic susceptibility} \qquad (1.9c) \\
\mathbf{J} &= \sigma \mathbf{E} \qquad &\sigma&: \text{ conductivity}
\end{aligned}
$$

We shall introduce two potential functions ϕ, a scalar potential, and ψ, a vector potential, in terms of which the components of the EM field are expressed. The electric and magnetic fields are defined in terms of ϕ and ψ as follows:

$$
\begin{aligned}
\mathbf{H} &= \nabla \times \psi \\
\mathbf{E} &= -\nabla \phi - \frac{\partial \psi}{\partial t}
\end{aligned} \qquad (1.10a)
$$

Th vector potential is further required to satisfy (Lorentz gauge)

$$
\nabla \cdot \psi + \varepsilon \mu \frac{\partial \phi}{\partial t} = 0 \qquad (1.10b)
$$

Using (1.10) in (1.9) we obtain two decoupled equations,

$$
\nabla^2 \phi - \varepsilon \mu \frac{\partial^2 \phi}{\partial t^2} = -\frac{1}{\varepsilon} \rho \qquad (1.11a)
$$

$$\nabla \times \nabla \times \psi - \varepsilon\mu \frac{\partial^2 \psi}{\partial t^2} = -\mu \mathbf{J} \qquad (1.11b)$$

When there is no free charge, that is, $\rho = 0$, the scalar potential may be canceled by a suitable choice, using the principle of gauge transformation [1, p207]. Then, equations in (1.10a) reduce to

$$\mathbf{H} = \nabla \times \psi$$
$$E = -\frac{\partial \psi}{\partial t} \qquad (1.12)$$

and the vector potential satisfies

$$\nabla \times \nabla \times \psi - \varepsilon\mu \frac{\partial^2 \psi}{\partial t^2} = -\mu \mathbf{J}$$
$$\nabla \cdot \psi = 0 \qquad (1.13)$$

Both electric and magnetic fields travel with the same speed unlike p and s waves in solids. The electric and magnetic vectors lie in a plane $\perp$ to the direction of propagation. The tip of a field vector will execute a smooth curve known as a polarization ellipse. For example, in a vertically polarized electric field the tip of the electric vector lies on a vertical line while the magnetic field lies on a horizontal line. The minor and major axes of the polarization ellipse are a and b respectively. Define an angle $\chi = \tan^{-1}(\frac{a}{b})$. The major axis of the polarization ellipse is inclined at an angle ζ (see fig. 1.2). Consider a plane wave with its wave vector in y-z plane making an angle θ with z axis. The electric vector will lie in a plane $\perp$ to direction of propagation. This plane is known as the x-θ plane. The electric vector can be split into two components,

Ex and Eθ, lying in x-θ plane.

$$E_x = E \cos \gamma \qquad \text{horizontal component}$$
$$E_\theta = E \sin \gamma \, e^{j\kappa} \qquad \text{vertical component}$$

where γ and κ are related to χ and ζ [2],

$$\cos(2\gamma) = \cos(2\chi)\cos(2\zeta) \qquad 0 \le \gamma \le \frac{\pi}{2}$$
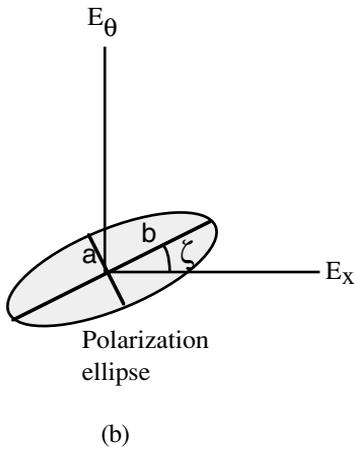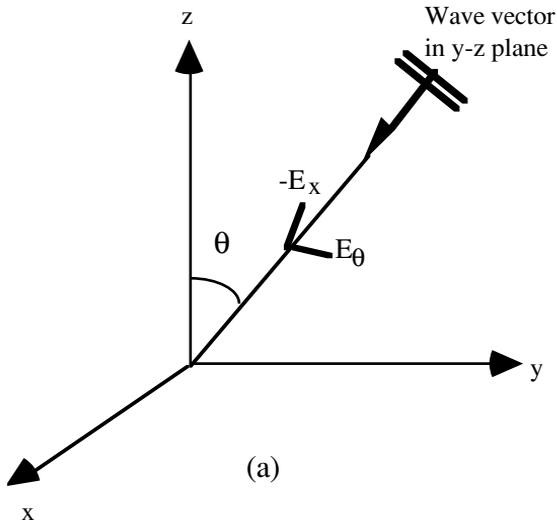
Figure 1.2: (a) EM wave propagation vector is in y-z plane. Electric and magnetic vectors lie in a plane ⊥ to the propagation vector and it contains the x-axis. (b) The tip of the electric vector executes a smooth curve, for example, an ellipse. The electric vector can be split into two components, E_X and E_θ.

$$\tan(\kappa) = \tan(2\chi)\csc(2\zeta) \qquad -\pi \le \kappa \le \pi$$

A plane EM wave is characterized by four parameters, namely, E (amplitude), θ (direction), χ and ζ (polarization parameters).

When there is no free charge current (free moving charge, i.e., $\rho = 0$) the vector potential for a plane wave EM field is given by

$$\psi = \mathbf{a}_p A \, \exp(j(\mathbf{k} \cdot \mathbf{r} - \omega t))$$

where $\mathbf{k} = \dfrac{\omega}{c} \mathbf{a}_k$ and A is a complex constant. $\mathbf{a}_p$ and $\mathbf{a}_k$ are mutually $\perp$ unit vectors. The electric and magnetic field vectors are given by

$$\mathbf{E} = j\omega \mathbf{a}_p A \exp(j(\mathbf{k} \cdot \mathbf{r} - \omega t))$$
$$\mathbf{H} = j\omega (\mathbf{a}_k \times \mathbf{a}_p) A \exp(j(\mathbf{k} \cdot \mathbf{r} - \omega t)) = \mathbf{a}_k \times \mathbf{E}$$

$$(1.14)$$

1.1.4 <u>Properties of Wavefields</u>: We shall briefly state some properties of the wavefields: (a) Radiation condition: A solution to a wave equation must satisfy the radiation condition or Sommerfeld condition given by [26, p. 499]

$$\frac{\partial \phi}{\partial r} \xrightarrow[r \to \infty]{} O(\frac{1}{r})$$

$$\frac{\partial \phi}{\partial r} + jk\phi \xrightarrow[r \to \infty]{} o(\frac{1}{r})$$

(b) Time Reversal: All wave equations without time varying coefficients or time derivatives of odd order share an interesting property of time reversal. If $\phi(x, y, z, t)$ is a solution of the wave equation it is easy to verify that $\phi(x, y, z, t_0 - t)$ is also a solution of the same wave equation for any constant t_0 [3]. This fact has been used for reverse time propagation of seismic wavefields. The seismic wavefield observed on a surface is time reversed and fed into a loudspeaker broadcasting the recorded wavefield.

1.1.5 <u>Sensing of Wavefields</u>: A device is required to convert a physical wavefield, say, pressure wave, into an electrical signal which is then sampled and digitized. The resulting numbers are stored in computer memory for further processing. Such a device is known as a sensor. While the mechanism of conversion of a physical wavefield into an electrical signal is not important from the point of array signal processing, the speed of conversion and the dynamic range are very relevant; accordingly the sampling rate and the number of bits per sample are fixed. The speed of conversion controls the bandwidth of a sensor, faster is the conversion larger is its bandwidth. As we have two types of wavefields, scalar wavefield and vector wavefield, there are also two types of

sensors. A scaler sensor is used to sense a scaler wavefield such as pressure or any one component of the EM field. The most common example of the scaler sensor is a microphone or a hydrophone. A vector sensor will measure all components of the vector wavefield, three components of mechanical waves in solid or six components of EM field. Three component seisometers are sometimes used in seismic exploration. Six component EM sensors are likely to be available very soon off the shelf [4]. Modern sensor arrays consist of several tens or hundreds of sensors. One major problem is the lack of uniformity in sensor response. Ideally, it is assumed that all sensors are omnidirectional with unit response. But in practice, the response of a sensor may depend upon the direction of incident wavefront and it may be complex. This behaviour in itself is not so disturbing as the variation of this behaviour from sensor to sensor. It is often required to carefully estimate the response of each sensor in a large array. This process is known as array calibration. Commonly used sensors measure just one component of the wavefield. This seems to be adequate as all components travel with the same speed (except p and s waves in solids) yielding the same delay information on which many array processing algorithms are based. Additional information such as the polarization and the particle motion has also been lately used in array processing. Some interesting developments in sensor technology, which will have considerable bearing on the future of sensor array processing, are taking place. For example, the vector sensors capable of measuring all six components of the EM field and four components of the mechanical field (pressure and three particle velocity components) have become commercially available [4].

        In addition to the above wavefield sensors, we have chemical sensors capable of detecting a very small quantity of chemicals in the vapour state. The distribution of vapour is governed by a diffusion equation in place of a wave equation. Extremely sensitive detectors of magnetic fields based on the principle of superconducting quantum interference have also appeared and have been used in magnetoencephalography.

§1.2 **Wavefield in Open Space**:
In a homegenous space without any reflecting boundaries the wave equation is easily solved. The propagation of wavefields may be described within the framework of filter theory. The propagating and nonpropagating (transient) waves are well demarcated in the frequency wavenumber domain. Continuation wavefield from one plane to another plane is easily achieved through a filtering operation. We shall introduce some of these basic concepts here in this section. Another important constraint imposed by practical considerations is that the wavefield is measured on a set of discrete points with discrete sensors. Thus, the spatial sampling is always implied; in contrast a temporal sampling of the analog sensor output is required only when digital processing is desired.
1.2.1 <u>Fourier Representation of Wave Field</u>: The generic form of a wave equation in a homogeneous medium is given by

$$\nabla^2 f = \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} \tag{1.15}$$

where $f(\mathbf{r},t)$ stands for any one of the wave types, for example, pressure or one of the components of a vector field. We introduce the Fourier integral representation of a wavefield,

$$f(\mathbf{r},t) = \frac{1}{8\pi^3} \int \int\limits_{-\infty}^{\infty} \int F(u,v,\omega) H(u,v,z) e^{j(\omega t - ux - vy)} du dv d\omega \tag{1.16}$$

in wave equation (1.15). We observe that $H(u,v,z)$ must satisfy an ordinary differential equation given by

$$\frac{d^2 H(u,v,z)}{dz^2} = (u^2 + v^2 - \frac{\omega^2}{c^2}) H(u,v,z) \tag{1.17}$$

whose solution is

$$H(u,v,z) = \exp(\pm\sqrt{(u^2 + v^2 - k^2} z) \tag{1.18}$$

where $k = \dfrac{\omega}{c}$ is known as a wavenumber. When $\sqrt{(u^2 + v^2)} > k$, we choose (-) sign for z>0 and (+) sign for z<0 so that the wavefield does not diverge. In both cases the field will rapidly decay as $|z| \to \infty$. These are known as evanescent waves. When $\sqrt{(u^2 + v^2)} < k$ we get propagating waves whose integral representation reduces to

$$f(x,y,t) =$$

$$\frac{1}{8\pi^3} \int \int\limits_{-\infty}^{\infty} \int F(u,v,\omega) e^{\pm j(\sqrt{k^2 - u^2 - v^2} z)} e^{j(\omega t - ux - vy)} du dv d\omega \tag{1.19}$$

where the sign in $e^{\pm j(\sqrt{k^2 - u^2 - v^2} z)}$ is selected depending on whether the waves are diverging or converging. The convention is (-) sign for diverging waves and (+) sign for converging waves (see fig. 1.3). Note that in a bounded space both diverging and converging waves can coexist and hence it would be necessary to use both signs in describing wavefields in a bounded space.
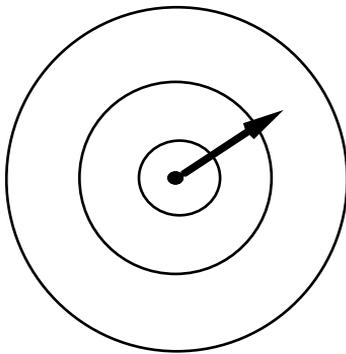
(-) sign   (+) sign

Diverging wavefronts   Converging wavefronts
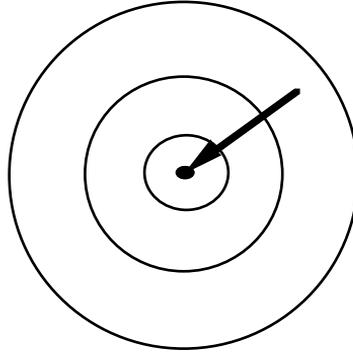
Figure 1.3: Sign convention in diverging and converging waves.

Equation (1.19) suggests an interesting possibility, that is, if a wavefield is observed on an x-y plane, it is possible to extrapolate it into the space above or below the plane of observation. Further, (1.19) may be looked upon as a sum of an infinitely large number of plane waves of the type $e^{j(\omega t - ux - vy - \sqrt{k^2 - u^2 - v^2}\,z)}$ with complex amplitude $F(u, v, \omega)$. The direction of propagation of individual wave is prescribed by the spatial frequencies u and v,

$$u = k \sin\theta \cos\varphi$$
$$v = k \sin\theta \sin\varphi$$

where φ and θ are respectively azimuth and elevation angles of a plane wave. The elevation angle is an angle between the z-axis and the wave vector and azimuth angle is an angle between the x-axis and the projection of the wave vector on the x-y plane. The representation given in (1.19) is also known as a plane wave decomposition of a wavefield. In (u, v, ω) space a single frequency plane wave is represented by a point and a wideband wave by a line passing through the center of the coordinate system (see fig. 1.4). The slope of the line is inversely proportional to the direction cosines (defined later on page 17) of the vector ⊥ to the plane wavefront and directly proportional to the speed of propagation.
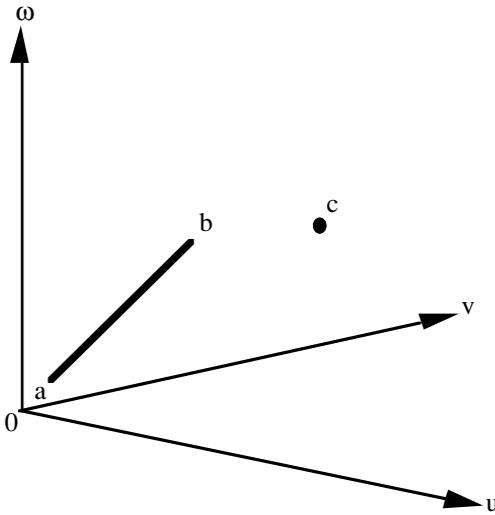
Figure 1.4: Representation of a plane wave in (u, v, ω) space. Point c represents a single frequency (narrowband) plane wave and line ab represents a wideband plane wave.

1.2.2 <u>Domain of Propagation</u>: We have noted earlier that propagation of waves is possible only when $\sqrt{(u^2 + v^2}} < k$. For a fixed ω, the spatial frequencies u and v must lie within a circle of radius k ($= \dfrac{\omega}{c}$), that is, a disc defined by

$$(u^2 + v^2) \leq (\frac{\omega}{c})^2 \tag{1.20}$$

Equation (1.20) represents, as a function of ω, a conical surface in (u, v, ω) space. It is a vertical cone with an apex at the center of the coordinates and the angle of the cone is inversely proportional to the speed of propagation. For a real signal, plane wave representation in (u, v, ω) space (see fig.1.4) extends below the (u,v) plane. The domain of wave propagation is obtained by reflecting the conical surface below (u,v) plane. This results into an hourglass figure shown in fig. 1.5.

1.2.3 <u>Apparent Propagation Speed</u>: Apparent propagation speed refers to the speed with which a wave appears to travel across an array of sensors placed on a line. For example, consider an array of sensors on the x-axis and a plane wavefront incident at angle θ and φ as shown in fig. 1.6. Travel time from p to

o is $dt = \dfrac{po\cos(\theta)}{c}$. Hence, the speed of propagation along the vertical axis is

$$\frac{po}{dt} = c_z = \frac{c}{\cos(\theta)} \qquad (1.21a)$$

To compute travel time from q to o, do the following construction:

    i) draw a wavefront $qs_2q_1 \perp$ to incident ray

    ii) draw a vertical plane through os_2. Since qq_1 is in the horizontal plane xoy and or is in the vertical plane, $\angle orq = \dfrac{\pi}{2}$.

    iii) Since rs_2 is in the wavefront plane, $\angle rs_2o = \dfrac{\pi}{2}$

From (i), (ii), and (iii) we obtain the following result:

$$or = oq \cos\varphi$$
$$os_2 = or \sin\theta$$
$$= oq \sin\theta \cos\varphi$$

The travel time from s_2 to o is $dt = \dfrac{os_2}{c} = \dfrac{oq}{c}\sin\theta\cos\varphi$. Hence the apparent speed along the x-axis is given by

$$c_x = \frac{oq}{dt} = \frac{c}{\sin\theta\cos\varphi} \qquad (1.21b)$$

Similarly we can show that

$$c_y = \frac{c}{\sin\theta\sin\varphi} \qquad (1.21c)$$

Equations (1.21b) and (1.21c) can be rewritten in terms of spatial frequencies

$$c_x = \frac{\omega}{k\sin\theta\cos\varphi} = \frac{\omega}{u}$$

$$c_y = \frac{\omega}{k\sin\theta\sin\varphi} = \frac{\omega}{v}$$
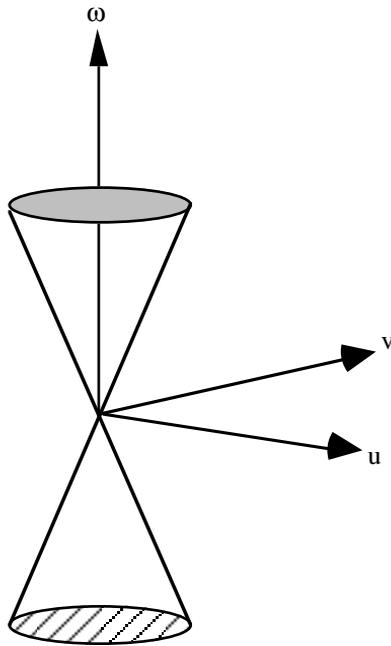
Figure 1.5: Domain of propagation in (u, v, ω) space (an hourglass figure).
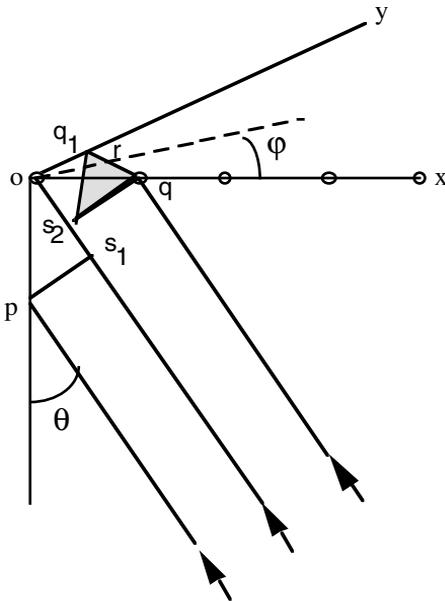
Figure 1.6: Geometrical derivation of apparent speeds.

$$c_z = \frac{\omega}{k\cos\theta} = \frac{\omega}{\sqrt{k^2 - u^2 - v^2}} \tag{1.22}$$

From (1.22) it follows that

$$\frac{1}{c_x^2} + \frac{1}{c_y^2} + \frac{1}{c_z^2} = \frac{1}{c^2} \tag{1.23}$$

and that the apparent speeds are always greater than or equal to the wave speed. Closely related to the apparent speed are the so-called ray parameters,

$$p_x = \frac{\sin\theta\cos\varphi}{c}, \; p_y = \frac{\sin\theta\sin\varphi}{c}, \; p_z = \frac{\cos\theta}{c} \tag{1.24}$$

The significance of ray parameters is that as a ray propagates through media of different wave speeds, the angles of incidence and emergence will change such that the ray parameters remain fixed, equal to that at the start of the ray. From (1.22) it is easy to show that the ray parameters are related to the apparent speeds,

$$p_x = \frac{1}{c_x}, \; p_y = \frac{1}{c_y}, \; p_z = \frac{1}{c_z} \tag{1.25a}$$

The direction cosines of a ray are defined as

$$\alpha = \sin\theta\cos\varphi, \; \beta = \sin\theta\sin\varphi, \; \gamma = \cos\theta \tag{1.25b}$$

From (1.24) and (1.25) the ray parameters can be expressed in terms of direction cosines

$$\alpha = p_x c, \; \beta = p_y c, \; \gamma = p_z c$$

1.2.4 <u>Continuation of Wavefield</u>: Consider a thin layer of sources on the z=0 plane. Let $f_0(x,y,t)$ be the wavefield observed close to the z=0 plane. The wavefield on a horizontal plane z=z_1 can be expressed in terms of that on z=0 plane. For this we make use of (1.19),

$$f(x,y,z_1,t) =$$

$$\frac{1}{8\pi^3}\int\int\int_{-\infty}^{\infty}F_0(u,v,\omega)e^{\pm j(\sqrt{k^2-u^2-v^2}z_1)}e^{j(\omega t-ux-vy)}dudvd\omega \quad (1.26)$$

where $F_0(u,v,\omega)$ is the Fourier transform of the wavefield observed on the z=0 surface. Thus, it is possible to extrapolate a wavefield observed on one plane to another plane provided the intervening space is source free. Note that the Fourier transforms of the wavefield observed on two parallel surfaces differ by a phase factor only. In the case of horizontal surfaces, the phase factor is simply given by $e^{\pm j(\sqrt{k^2-u^2+v^2}\Delta z)}$ where Δz is vertical separation. Each component of the plane wave decomposition is subjected to a phase shift whose magnitude depends upon the spatial frequencies, u and v, or on the direction of the wave vector. For example, for $\Delta z = \lambda$ the phase shift applied to different plane waves is shown in fig. 1.7 as a function of u (keeping v=0). It may be recalled that the Fourier transform of the wavefield for $(u^2+v^2)>k^2$ is rapidly vanishing and hence the phase shift is set to zero in this range.

1.2.5 <u>Point Source</u>: A point source is often used as a source of illumination, but it will only generate spherical waves which produce a much more complex (mathematically speaking) response from a target than what a plane wave does. Fortunately, a point source wavefield may be written as a sum of infinitely many plane waves; a plane wave decomposition was described in subsection (1.2.1). Additionally, a response of a target may be obtained as a sum of plane wave responses. To obtain the Fourier integral representation of a point source wavefield we go back to (1.19) where we shall assume that the Fourier transform of the wavefield has radial symmetry,

$$f(x,y,z,t)$$

$$=\frac{1}{8\pi^3}\int\int\int_{-\infty}^{\infty}F_0(u,v,\omega)e^{\pm j(\sqrt{k^2-u^2-v^2}z)}e^{j(\omega t-ux-vy)}dudvd\omega \qquad (1.27a)$$

$$=\frac{1}{4\pi^2}\int F_0(s,\omega)e^{j\omega t}d\omega\int_0^{\infty}se^{\pm j(\sqrt{k^2-s^2}z)}ds\frac{1}{2\pi}\int_0^{2\pi}e^{-j(sr\cos(\theta-\varphi))}d\theta$$

where $F_0(u,v,\omega)$ is the Fourier transform of $f(x,y,z=0,t)$. We can further simplify (1.27a) as

$$f(r,z,t)\quad=\frac{1}{4\pi^2}\int_{-\infty}^{\infty}F_0(s,\omega)e^{j\omega t}d\omega\int_0^{\infty}sJ_0(sr)e^{\pm j(\sqrt{k^2-s^2}z)}ds \quad (1.27b)$$
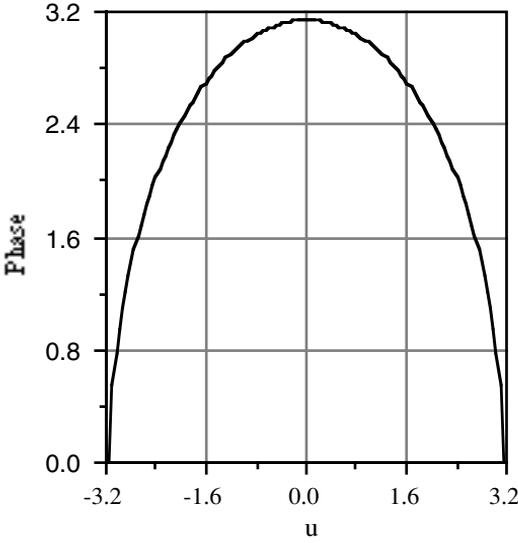
Figure 1.7: Phase shift as a function of u (keeping v=0) for vertical separation equal to wavelength (assumed to one).

where $s = \sqrt{u^2 + v^2} = k\sin\gamma$ where $0 \leq \gamma \leq \dfrac{\pi}{2}$. In (1.27b) we replace s by $k\sin\gamma$ and rewrite (1.27b) as

$$f(r,z,t) = \frac{1}{8\pi^2} \int\limits_{-\infty}^{\infty} d\omega \int\limits_{0}^{\frac{\pi}{2}} F_0(k\sin\gamma,\omega) e^{j\omega(t\pm\frac{z}{c}\cos\gamma)} k^2 \sin(2\gamma) J_0(kr\sin\gamma) d\gamma$$

$$(1.28)$$

Equation (1.28) is a plane wave decomposition (PWD) of a point source wavefield. In an r-z plane, for a fixed ω and γ the integrand in (1.28) represents a plane wave component, traveling with an angle of incidence γ as shown in fig. 1.8 and $\Delta t = \dfrac{z}{c}\cos\gamma$ is propagation time from surface to a depth z. The inverse of (1.28) is given by

point source

r

γ

z
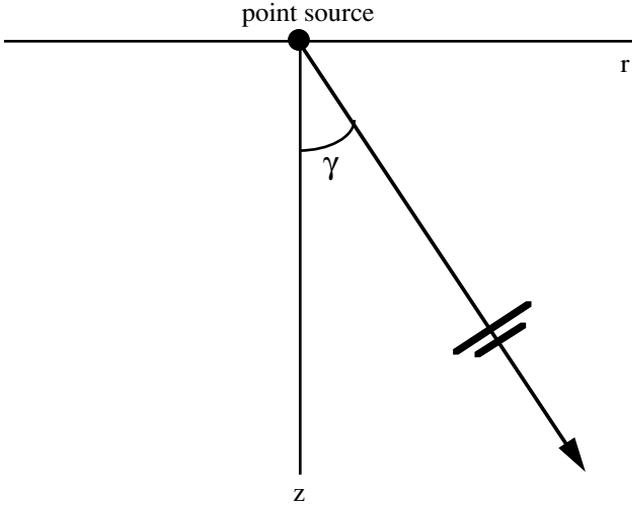
Figure 1.8: In a homogeneous medium the wavefield due to a point source may be expressed as a sum of plane waves traveling with an angle of incidence γ, where

$$0 \le \gamma \le \frac{\pi}{2} \ .$$

$$F_0(s,\omega)e^{\pm j(\sqrt{k^2 - s^2}\,z)} = \int_0^\infty F(r,z,\omega)rJ_0(sr)dr \qquad (1.29)$$

where $F(r,z,\omega)$ is a Fourier transform (temporal) of $f(r,z,t)$. On the z=0 surface (1.29) reduces to

$$F_0(k\sin\gamma,\omega) = \int_0^\infty F(r,z=0,\omega)rJ_0(kr\sin\gamma)dr \qquad (1.30)$$

where we have substituted $s = k\sin\gamma$. Equation (1.30) in the time domain may be expressed as

$$f(t,\sin\gamma) = \int_0^\infty \left\{ f(r,z=0,t) * \frac{2}{\sqrt{(\frac{r}{c}\sin\gamma)^2 - t^2}} \right\} r dr \quad (1.31)$$

Equation (1.31) enables us to convert the wavefield due to a point source into that due to a plane wave source. A linear array of sensors is used in a radial direction with respect to point source. The output of each sensor is first filtered with a filter having response

$$h(t) = \frac{2}{\sqrt{(\frac{r}{c}\sin\gamma)^2 - t^2}}$$

and then summed over all sensors. The impulse response function may be expressed as an integral involving a delta function [5],

$$h(t) = \frac{2}{\sqrt{(\frac{r}{c}\sin\gamma)^2 - t^2}} = \int_{-\pi}^{\pi} \delta(t - t_0\cos\beta)d\beta$$

where $t_0 = (\frac{r}{c}\sin\gamma)$. Using the above integral representation of the impulse response function in (1.31) we obtain

$$f(t,\sin\gamma) = \int_0^\infty \left\{ \int_{-t_0}^{t_0} f(r,z=0,t-t') \frac{2}{\sqrt{t_0^2 - (t')^2}} dt' \right\} r dr$$

$$= \int_0^\infty \left\{ \int_{-t_0}^{t_0} f(r,z=0,t-t')dt' \int_{-\pi}^{\pi} \delta(t'-t_0\cos\beta)d\beta \right\} r dr \quad (1.32)$$

$$= \int_0^\infty \left\{ \int_{-\pi}^{\pi} f(r,z=0,t-t_0\cos\beta)d\beta \right\} r dr$$

For a discrete array the integral is replaced by a sum,

$$f(n\Delta t_0,\sin\gamma) = \int_{-\pi}^{\pi}\sum_{n=0}^{\infty} n\Delta r^2 f(n\Delta r, z=0, t-n\Delta t_0\cos\beta)d\beta \quad (1.33)$$

where $\Delta t_0 = \frac{\Delta r}{c}\sin\gamma$. The inner sum in (1.33) is the sum-after-delay operation, commonly known as slant stacking in seismic exploration. Later in chapter 2 we shall show how this operation is related to the radon transform.

1.2.6 <u>Spatial Sampling and Aliasing</u>: A wavefield is measured on a set of discrete points with discrete sensors. Thus, spatial sampling is always implied; in contrast a temporal sampling of an analog sensor output is used only when digital processing is desired. Often the temporal sampling is preceded by some kind of low pass filtering, which determines the maximum frequency and the maximum sampling rate according to the well known sampling theorem. It is thus possible to avoid the temporal aliasing but it is a different story with spatial aliasing which is intimately related to propagation speed. For simplicity of analysis let us assume only spatial sampling. Consider a broadband plane wave incident on an infinitely long array of sensors,

$$f(x,t) = f(t - \frac{x}{c_x}) \sum_{i=-\infty}^{\infty} \delta(x - i\Delta x)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j(t - \frac{x}{c_x})\omega} d\omega \sum_{i=-\infty}^{\infty} \delta(x - i\Delta x)$$

Note that c_x stands for wave speed in the direction of x-axis or apparent speed. Taking the 2D Fourier transform of f(x,t)

$$F(u,\omega) = F(u,\omega) = \int_{-\infty}^{+\infty}\int f(x,t) e^{j(\omega t - ux)} dx dt$$

$$= \int_{-\infty}^{\infty} F(\omega') d\omega' \sum_{i=-\infty}^{\infty} \int_{-\infty}^{+\infty}\int e^{-j(\omega t - ux)} e^{j(t - \frac{x}{c})\omega'} \delta(x - i\Delta x) dx dt$$

$$= \int_{-\infty}^{\infty} F(\omega') \delta(\omega' - \omega) d\omega' \sum_{i=-\infty}^{\infty} \int_{-\infty}^{\infty} e^{j(u - \frac{\omega'}{c_x})x} \delta(x - i\Delta x) dx$$

$$= \int_{-\infty}^{\infty} F(\omega') \delta(\omega' - \omega) d\omega' \sum_{i=-\infty}^{\infty} e^{j(u - \frac{\omega'}{c_x})i\Delta x} \qquad (1.34)$$

$$= \int_{-\infty}^{\infty} F(\omega') \delta(\omega' - \omega) d\omega' \sum_{k=-\infty}^{\infty} \delta(u - \frac{\omega'}{c_x} - \frac{2\pi}{\Delta x} k)$$

The spatial Fourier transform of the array output is sketched in fig. 1.9. The signal spectrum is concentrated on sloping lines in the u-w plane as seen in fig. 1.9.

Depending upon the signal bandwidth and the apparent speed, an alias will show up within the principal band. In fig. (1.9) there is no aliasing when
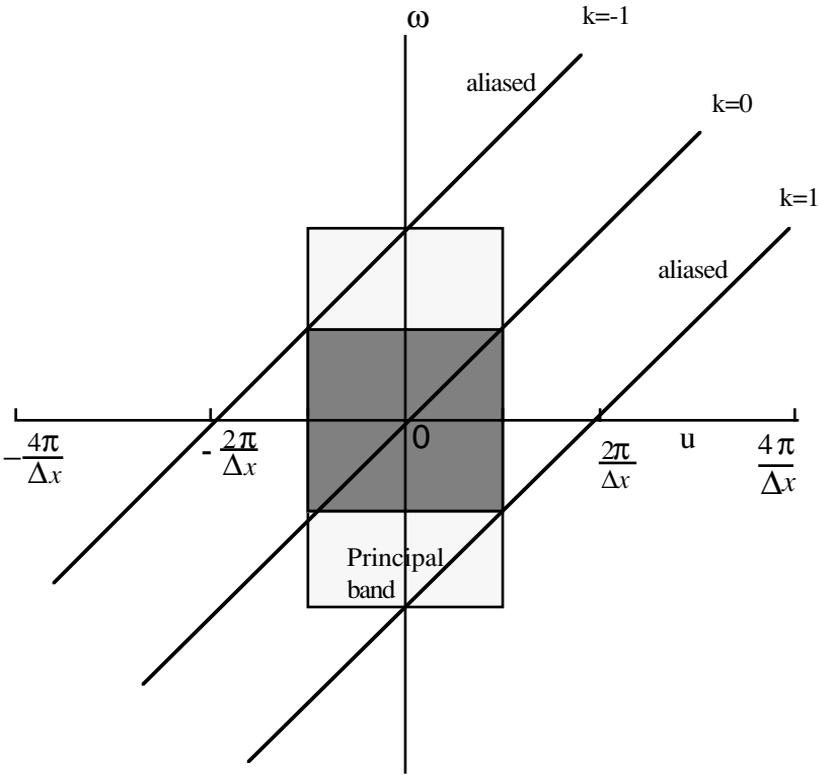
Figure 1.9: The Fourier transform of spatially sampled wavefield (plane wave) lies on a set of sloping lines (with slope= $1/C_x$) corresponding to different values of k in (1.34).

the signal bandwidth is as shown by the dark square, which in this case corresponds to a sensor spacing $\Delta x = \dfrac{\lambda_{\min}}{2}$ or $\Delta x = \dfrac{\pi}{\omega_{\max}} c_x$. Note that, since the sampling interval $\Delta t = \dfrac{\pi}{\omega_{\max}}$, $\Delta x = \Delta t \, c_x$. There is no aliasing whenever the sensor spacing and the time sampling interval are such that $\dfrac{\Delta x}{\Delta t} \leq c_x$. For the vertical angle of incidence, since $c_x = \infty$, there is no aliasing effect for any Δx.

1.2.7 <u>Dispersion</u>: A medium is said to be dispersive when a wavefront travels at a speed which is a function of frequency. Consider a wavefront at a fixed

temporal frequency, ω_0. Let us track a particular point, say, the crest of a sinusoidal waveform. The speed with which a point on the waveform travels is said to be phase speed, which is equal to $\dfrac{\lambda_0}{T_0}$ where λ_0 is the wavelength and T_0 is the period of a single frequency waveform. Thus, phase speed, by definition, is given by

$$c_{ph} = \frac{\lambda_0}{T_0} = \frac{\omega_0}{k_0} = c$$

and it is equal to the propagation speed.

Now consider a group of waves whose temporal spectrum is centered at frequency ω_0 and it is spread over a frequency interval $\Delta\omega$. The wavefield may be expressed as

$$f(\mathbf{r},t) = \frac{1}{2\pi} \int_{\omega_0 - \frac{\Delta\omega}{2}}^{\omega_0 + \frac{\Delta\omega}{2}} F(\omega)\exp(j(\omega t - \mathbf{k}\cdot\mathbf{r})d\omega \qquad (1.35a)$$

where $\mathbf{k} = (u, v, \sqrt{(\frac{\omega}{c})^2 - u^2 - v^2})$ is the wave vector. We assume that the propagation speed is a function of frequency, and hence the wave vector is also a function of frequency. Using a Taylor's series expansion of the wave vector

$$\mathbf{k}(\omega) = \mathbf{k}(\omega_0) + \frac{d\mathbf{k}(\omega)}{d\omega}\bigg|_{\omega=\omega_0} (\omega - \omega_0) + \ldots$$

in (1.35a) we obtain

$$f(\mathbf{r},t) = F(\omega_0)e^{j(\omega_0 t - \mathbf{k}(\omega_0)\cdot\mathbf{r})} \frac{1}{2\pi} \int_{-\frac{\Delta\omega}{2}}^{\frac{\Delta\omega}{2}} \exp[j(t - \frac{d\mathbf{k}(\omega)}{d\omega}\bigg|_{\omega=\omega_0} \cdot\mathbf{r})\tilde{\omega}]d\tilde{\omega}$$

$$= F(\omega_0)e^{j(\omega_0 t - \mathbf{k}(\omega_0)\cdot\mathbf{r})} \frac{\Delta\omega}{2\pi} \sin c[(t - \frac{d\mathbf{k}(\omega)}{d\omega}\bigg|_{\omega=\omega_0} \cdot\mathbf{r})\frac{\Delta\omega}{2}]$$

$$\qquad (1.35b)$$
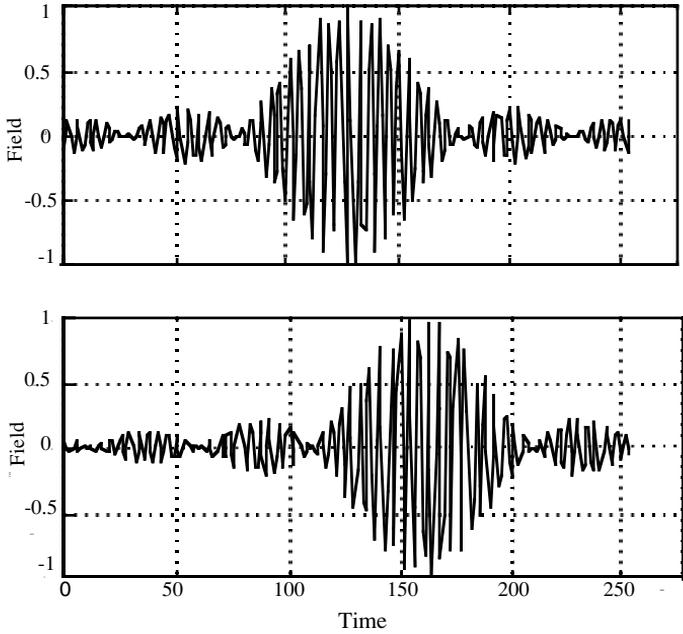
Figure 1.10: A narrowband signal with a center frequency at 0.23 Hz and a bandwidth equal to 0.02 Hz. The envelope travels at the group speed which in this is equal to the phase speed (assumed to be one). The signal in the lower panel arrives 32 time units later.

where $\tilde{\omega} = (\omega - \omega_0)$. Observe that the second term modulates the carrier wave (first term). The second term is a propagating waveform which travels at a speed of c_{gp} which satisfies an equation

$$\left|\frac{d\mathbf{k}(\omega)}{d\omega}\right|_{\omega=\omega_0} c_{gp} = 1 \tag{1.36a}$$

or

$$c_{gp} = \frac{1}{\left|\dfrac{d\mathbf{k}(\omega)}{d\omega}\right|_{\omega=\omega_0}} \tag{1.36b}$$

The group speed differs from the phase speed only when the medium is dispersive. In a nondispervive medium both speeds are equal. The modulating waveform (second term in (1.35b)) travels at the same speed as the carrier wave (first term). An example of propagation of a narrowband signal is shown in fig.

where we have assumed that $c_{ph} = c_{gp} = 1$. The arrival of the signal is denoted by the arrival of the crest of the modulating wave which travels with a speed equal to c_{gp}. Hence the energy is transmitted at a speed equal to the group speed.

§1.3 **Wavefield in Bounded Space:**

Wavefield in a space bounded by a plane reflecting boundaries is of interest in many practical problems. The examples of bounded space are (i) acoustic field in a room, (ii) acoustic field in a shallow water channel and (iii) mechanical waves in a layered medium, for example, in shallow earth, etc. The wavefield in a bounded space is normally studied under two different approaches, namely, rays and wavefronts, and the exact solution of the wave equation leading to modes. While the former is more versatile and easy to use, it is less accurate. The second approach is mathematically accurate but difficult to use except in simple geometries.

1.3.1 <u>Ray Propagation</u>: A point source emits a spherical wavefront, a surface of equal phase. Lines $\perp$ to the wavefront are rays along which the wave energy is transmitted. A spherical wavefront at a large distance from a point source may be approximated by a plane

$$f(\mathbf{r},t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j[\omega t - u_0(x-x_0) - v_0(y-y_0) - \sqrt{k^2-s_0^2}(z-z_0)]} d\omega$$

where (x_0, y_0, z_0) is the position of a point source and $\mathbf{r} = ((x-x_0),(y-y_0),(z-z_0))$ is the position vector of a point on a plane wavefront. Note that the direction cosines are related to the spatial frequencies, $u_0 = k\alpha_0$ and $v_0 = k\beta_0$ where $\alpha_0, \beta_0, \sqrt{1-\alpha_0^2-\beta_0^2}$ are direction cosines of a ray. If a sensor is placed at point (x_1, y_1, z_1), the wavefield received at (x_1, y_1, z_1) is given by

$$f(\mathbf{r}_1,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j[\omega t - \phi_1]} d\omega$$

where $\phi_1 = u_0(x_1-x_0) + v_0(y_1-y_0) + \sqrt{k^2-s_0^2}(z_1-z_0)$ is a constant phase. We can also write $\phi_1 = \omega t_1$ where
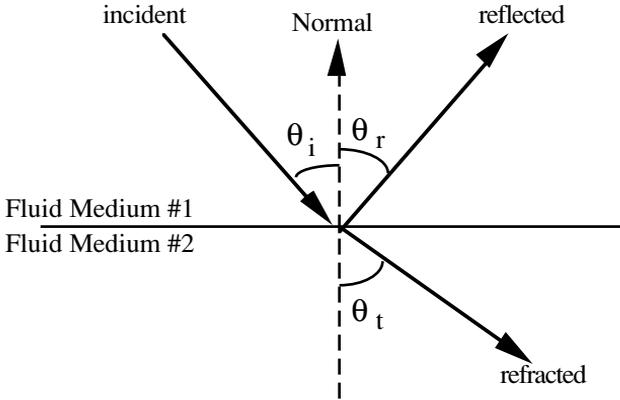
Figure 1.11: Laws of reflection and refraction: (1) Incident wave, reflected wave, refracted wave and the normal to the interface at the point of incidence lie in the same plane. (2). $\theta_i = \theta_r$ (3) $\dfrac{\sin\theta_i}{c_1} = \dfrac{\sin\theta_t}{c_2} = p$ (ray parameter).

$$
\begin{aligned}
t_1 &= \frac{\sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2}}{c} \\
&= \frac{(x_1 - x_0)}{c_x} + \frac{(y_1 - y_0)}{c_y} + \frac{(z_1 - z_0)}{c_z}
\end{aligned}
$$

where c_x, c_y, c_z are apparent speeds in x, y, z directions, respectively. When a ray encounters an interface separating two contrasting media it splits itself into two rays, a reflected and a refracted ray. The laws of reflection and refraction are summarized in fig. 1.11. The reflection and refraction coefficients are related to the impedance of the media on both sides of the interface. For example, the reflection and transmission coefficients at the interface separating two fluid media are given by

$$
\hat{r} = \frac{\dfrac{\rho_2}{\rho_1}\cos\theta_i - \sqrt{(\dfrac{c_2}{c_1})^2 - \sin^2\theta_i}}{\dfrac{\rho_2}{\rho_1}\cos\theta_i + \sqrt{(\dfrac{c_2}{c_1})^2 - \sin^2\theta_i}} \tag{1.37a}
$$
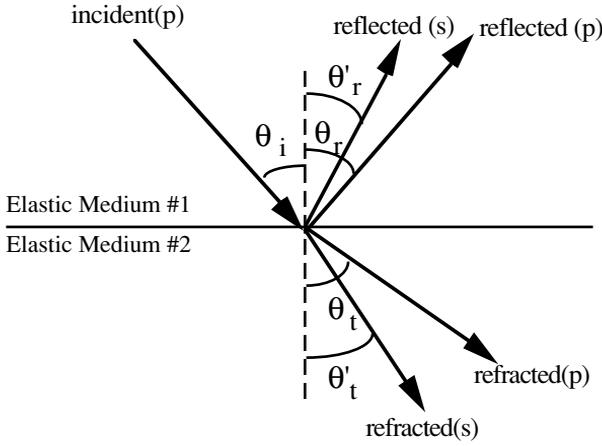
Figure 1.12: Reflection and refraction of a p-wave at an interface. Four rays as shown are generated. All rays including the incident ray and the normal to the interface lie in a plane. The angle of incidence and the angles of reflection and refraction are related. $\theta_i = \theta_r$ and $\dfrac{\sin \theta_i}{c_p^I} = \dfrac{\sin \theta_r'}{c_s^I} = \dfrac{\sin \theta_t}{c_p^{II}} = \dfrac{\sin \theta_t'}{c_s^{II}} = \text{const.}$

$$\hat{t} = \frac{2\cos\theta_i}{\dfrac{\rho_2}{\rho_1}\cos\theta_i + \sqrt{(\dfrac{c_2}{c_1})^2 - \sin^2\theta_i}} \qquad (1.37b)$$

Reflection and refraction at an interface separating two elastic media are more complex. An incident longitudinal wave will give rise to two reflected and two refracted rays. The first ray is a p-wave and the second ray is a s-wave. The laws of reflection and refraction for a p-wave incident at the interface are summarized in fig. 1.12. The coefficients of reflection and refraction, $(\hat{r}_p, \hat{r}_s, \hat{t}_p, \hat{t}_s)$, are obtained by solving the following system of four linear equations [6, p3-101]:

$$\begin{aligned}
\hat{r}_p \cos\theta_i - \hat{r}_s \sin\theta_r' + \hat{t}_p \cos\theta_t - \hat{t}_s \sin\theta_t' &= \cos\theta_i \\
-\hat{r}_p \sin\theta_i - \hat{r}_s \cos\theta_r' + \hat{t}_p \sin\theta_t + \hat{t}_s \cos\theta_t' &= \sin\theta_i
\end{aligned} \qquad (1.38a)$$

$$-\hat{r}_p \frac{(\lambda_1 + 2\mu_1 - 2\mu_1 \sin^2 \theta_r)}{c_p^I} + \hat{r}_s \frac{\mu_1}{c_s^I} \sin 2\theta_r'$$

$$+\hat{t}_p \frac{(\lambda_2 + 2\mu_2 - 2\mu_2 \sin^2 \theta_t)}{c_p^{II}} - \hat{t}_s \frac{\mu_2}{c_s^{II}} \sin 2\theta_t' = \frac{(\lambda_1 - 2\mu_1 - 2\mu_1 \sin^2 \theta_r)}{c_p^I}$$

$$\hat{r}_p \frac{\mu_1}{c_p^I} \sin 2\theta_r + \hat{r}_s \frac{\mu_1}{c_s^I} \cos 2\theta_r' + \hat{t}_p \frac{\mu_2}{c_p^{II}} \sin 2\theta_t - \hat{t}_s \frac{\mu_2}{c_s^{II}} \cos 2\theta_t' = \frac{\mu_1}{c_p^I} \sin 2\theta_r$$

$$(1.38b)$$

where (λ_1, μ_1) are the elastic constants and (c_p^I, c_s^I) are respectively the p-wave and the s-wave speed in the upper medium. Other parameters refer to the lower medium. For vertical incidence, that is, $\theta_i = 0$ from (1.38), it follows that $\hat{r}_s = \hat{t}_s = 0$. This is a consequence of the fact that there is no s-wave generation for vertical incidence.

Polarization in the plane of incidence: Finally, an EM wave at an interface between two different media undergoes reflection and refraction. However, there is only one reflected and one transmitted wave which travel at the same speed. The process of reflection and refraction is strongly influenced by the polarization angle. We shall consider one simple case of polarization in the plane of incidence (fig. 1.13). The coefficient of reflection and transmission is given by (Fresnel's equations)

$$\hat{r}_E = \frac{\chi - \zeta}{\chi + \zeta}, \quad \hat{t}_E = \frac{2}{\chi + \zeta} \qquad (1.39)$$

where $\chi = \frac{\cos \theta_t}{\cos \theta_i}$ and $\zeta = \frac{\mu_1 c_1}{\mu_2 c_2}$. Note, when $\chi = \zeta$, $\hat{r}_E = 0$ and $\hat{t}_E = 1$; that is, there is no reflected energy. This phenomenon takes place at a specific angle of incidence known as Brewster angle given by

$$\sin^2 \theta_B = \frac{1 - \zeta^2}{(\frac{c_2}{c_1})^2 - \zeta^2} \qquad (1.40)$$

1.3.2 Propagation in Channel - Ray Theory: Consider a source inside a fluid channel which is bounded from above by a free surface and from below by another fluid (this is known as the Pekeris model). The reflection coefficient at the
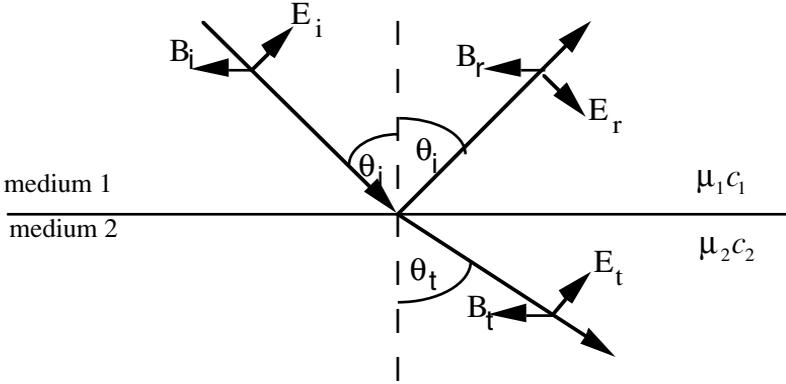
Figure 1.13: The electric vector is in the plane of incidence, that is, in the plane containing the incident ray, normal to interface and reflected and refracted rays.

free surface is (-1) and that at the bottom is $\hat{r}_b$ (see (1.37)). Because of multiple reflections at the free surface and at the bottom (see fig. 1.14) many waves will reach a sensor array at different times. It is convenient to model all these waves as emanating from a series of images whose position can be determined by following simple geometrical optics rules. An infinite set of images is formed between two parallel reflecting surfaces. We index these images with two integers (i,k); the first integer represents a group and the second integer represents an image within a group. There are four images in each group. For example, in the i=0 group the images are: s_{01} and s_{02} which are equal but opposite in strength, that is, a surface dipole, and s_{03} and s_{04}, which are caused by the reflection of the surface dipole onto the bottom. The surface dipole and its image are separated by a distance of 2H where H is the depth of the fluid channel. The next group of images, that is, for i=1, is obtained by sliding the surface dipole two depth units above and the image dipole two depth units below the bottom. The vertical distance with respect to the top sensor in the array to different images in the ith group is given by

$$
\begin{aligned}
H_{i1} &= 2iH + z_r - z_s \\
H_{i2} &= 2iH + z_r + z_s \\
H_{i3} &= 2(i+1)H - z_r - z_s \\
H_{i4} &= 2(i+1)H - z_r + z_s \qquad i = 0,1,2,...
\end{aligned}
\tag{1.41}
$$
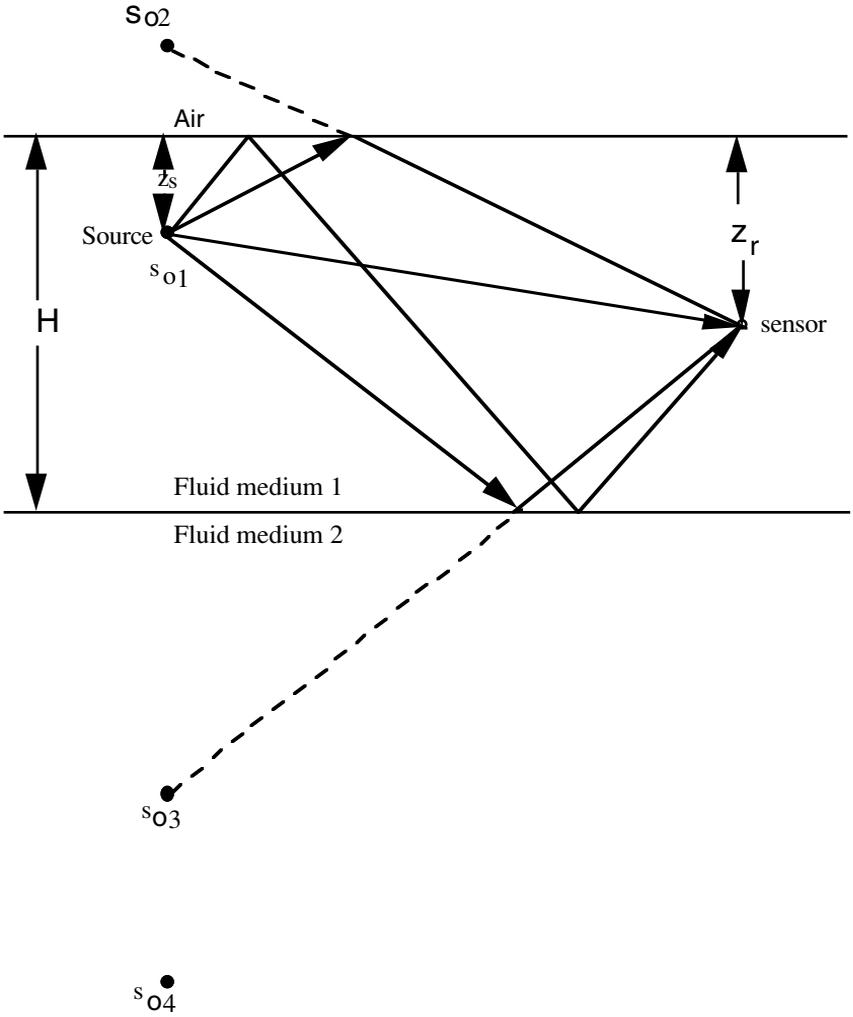
Figure 1.14: A shallow water channel with a sound source $(\bullet)$ at depth z_s and a sensor at depth z_r. Rays starting from the source are reflected at the surface and bottom and finally reach the sensor after one or more reflections. Some of the images $(\bullet)$ are also shown.

An image whose second index is 1 or 2 lies above the bottom and an image whose second image is 3 or 4 lies below the bottom. The strength of an image depends upon the number of bounces the ray has undergone before reaching the

sensor and also, on account of geometrical spreading and absorption, on the actual distance traveled which is given by

$$l_{i0} = \sqrt{d^2 + H_{i0}^2} \qquad l_{i1} = \sqrt{d^2 + H_{i1}^2}$$
$$l_{i2} = \sqrt{d^2 + H_{i2}^2} \qquad l_{i3} = \sqrt{d^2 + H_{i3}^2} \qquad (1.42)$$

where d is the horizontal distance to the sensor from the source. The strength of the images is given by

$$\alpha_{i0} = (-1)^i \hat{r}_b^i \frac{e^{-\beta l_{i0}}}{l_{i0}} \qquad \alpha_{i1} = (-1)^{i+1} \hat{r}_b^i \frac{e^{-\beta l_{i1}}}{l_{i1}}$$
$$\alpha_{i2} = (-1)^i \hat{r}_b^{i+1} \frac{e^{-\beta l_{i2}}}{l_{i2}} \qquad \alpha_{i3} = (-1)^{i+1} \hat{r}_b^{i+1} \frac{e^{-\beta l_{i3}}}{l_{i3}} \qquad (1.43)$$

where β is the attenuation coefficient in the top liquid layer. The signal reaching the sensor (fig 1.14) may be expressed as

$$p(t) = \frac{1}{2\pi} \sum_{i=0}^{\infty} \sum_{m=0}^{3} \int_{-\infty}^{\infty} \alpha_{im} P(\omega) e^{j(t-\tau_{im})\omega} d\omega \qquad (1.44)$$

where p(t) is the pressure field received by a sensor and $P(\omega)$ is the Fourier transform of the waveform emitted by the source. τ_{im} is propagation delay $\tau_{im} = \dfrac{l_{im}}{c}$. The pressure field given by (1.44) is a complex field. The magnitude and phase are found to vary rapidly with the depth of the sensor. For example, for a channel with the following parameters:

$$H = 100 \text{ m (meters)}$$
$$d = 2000 \text{ m}$$
$$z_s = 50 \text{ m}$$
$$\rho_1 = 1, \ c_1 = 1500$$
$$\rho_2 = 2, c_2 = 1600$$

the magnitude and phase variations as a function of the depth of a sensor are shown in fig. 1.15. Eighty images were taken into account in computing the pressure field.

As the range increases, while the position of image sources remain unchanged, the strength of the images may increase on account of the phenomenon of total internal reflection. Recalling equation (1.37) it may be

noticed that whenever $\sin(\theta_i) = \dfrac{c_1}{c_2}$ $(c_1 < c_2)$, $|r_b| = 1$. This angle of incidence is known as the *critical angle,* θ_c. For $\theta_i > \theta_c$, r_b becomes purely imaginary. The wave energy travels along the interface and as it travels some energy is re-radiated back into the channel at the critical angle. It is possible to account for long distance transmission of acoustic energy using the above ray model even at frequencies as low as 200Hz [7]. By introducing certain corrections (i.e., beam displacement) the ray model has been improved in the low frequency region (60 - 140 Hz)[8]. The ray model has been used for source localization in shallow water [9].

1.3.3 Propagation in a Channel - Normal Mode Theory: The pressure field due to a point source in a fluid channel, as depicted in fig. 1.14, is given by

$$p(r,z,t) = \frac{j}{2} \int\limits_{-\infty}^{\infty} \sum_{n=0}^{\infty} \phi_n(z)\phi_n(z_s) H_0^1(k_n r) P(\omega) e^{j\omega t} d\omega \qquad (1.45)$$

where $\phi_n(z)$ is an eigenfunction and k_n is its corresponding eigenvalue obtained by solving a differential equation with homogeneous boundary conditions

$$\frac{d^2\phi}{dz^2} + \gamma_n^2\phi = 0$$

$$\phi(0) = 0 \qquad\qquad (1.46)$$

$$\left.\frac{d\phi}{dz}\right|_{z=H} = 0 \quad (\textit{rigid bottom})$$

where $k_n^2 + \gamma_n^2 = \dfrac{\omega^2}{c^2(z)}$. k_n is a horizontal wavenumber and γ_n is a vertical wavenumber. The source is assumed to be at a depth z_s. Note that both wavenumbers are functions of frequency and also of depth when the sound speed is a function of depth. Consider a special case of isospeed channel and large r. Firstly, the Hankel function can be approximated as $H_0^1(k_n r) \approx \sqrt{\dfrac{2}{k_n \pi r}} e^{j(k_n r + \frac{\pi}{4})}$; secondly both wavenumbers are independent of depth and therefore the solution of (1.46) is given by
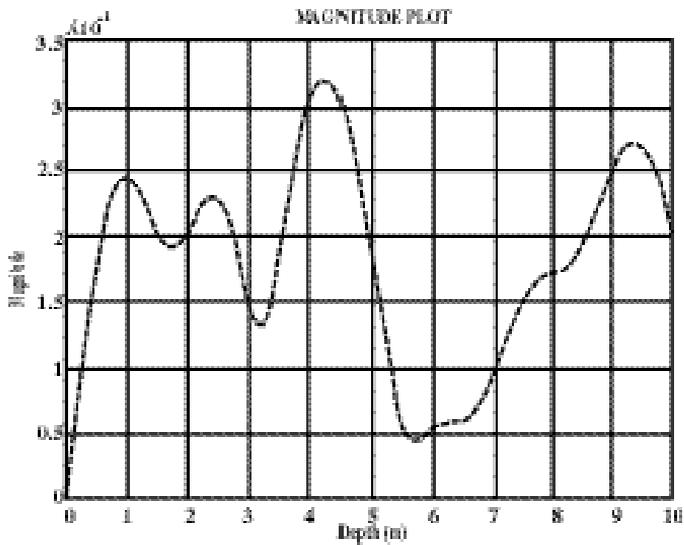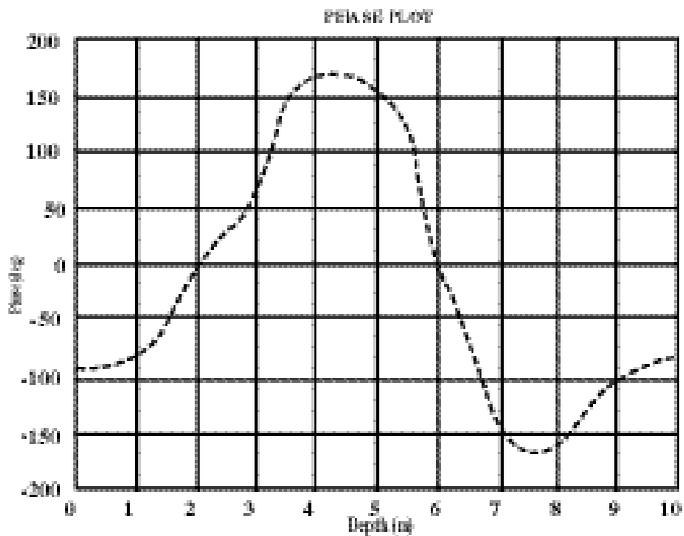
Figure 1.15: (a) Magnitude and (b) phase of the computed pressure field in a shallow water channel. Eighty images were taken into consideration for evaluating the net

field at a desired point. Note the highly complex structure of the field caused by interference.

$$\phi_n(z) = \sin(\gamma_n z) = \frac{e^{j\gamma_n z} - e^{-j\gamma_n z}}{2j}$$

where $\gamma_n = \dfrac{(n-\frac{1}{2})\pi}{H}$. For this special case the pressure field in the fluid channel with the rigid bottom is given by

$$p(r,z,t) =$$

$$\frac{e^{j\frac{\pi}{4}}}{2} \int_{-\infty}^{\infty} \sum_{n=0}^{\infty} \sqrt{\frac{1}{2k_n \pi r}} \sin(\gamma_n z_s)\left\{ e^{j(\gamma_n z + k_n r + \omega t)} - e^{j(-\gamma_n z + k_n r + \omega t)} \right\} P(\omega) d\omega$$

(1.47)

From (1.47) it is possible to infer that each mode is a sum of two plane wavefronts traveling in the vertical plane at angle $\pm\theta_n$, where $\tan\theta_n = \dfrac{\gamma_n}{k_n}$ with respect to the horizontal plane. In three dimensions the wavefront is a conical wavefront (see fig.1.16); the angle of the cone is $\dfrac{\pi}{2} - \theta_n$. Note that the direction of propagation of the wavefronts is solely dependent on the channel characteristics.

The comparison of ray and normal mode approaches to propagation in a shallow water channel is instructive. The ray approach is essentially a high frequency approximation to the exact solution of the wave equation. The normal mode approach on the other hand is a series approximation to the exact solution of the wave equation. The conical wavefronts obtained by decomposing each mode may be looked upon as a result of constructive interference of spherical wavefronts from the source and its images [8]. At low frequency the accuracy of the ray approach is enhanced if one were to use the concept of beam displacement. At a point of reflection the beam is found to be laterally displaced, a phenomenon first observed in optics [10]. Both numerical and experimental evidence in support of the fact that beam displacement does help to increase the accuracy of the ray approach in relation to the exact solution has been widely reported [11, 12].

1.3.4 Propagation Through Layered Medium: In seismic exploration a horizontally layered medium is often used as a model. Vertically incidentlongitudinal waves (p-waves) are preferred as there is no loss of energy through conversion into s-waves at each interface. Also, as s-waves arrive slightly later,
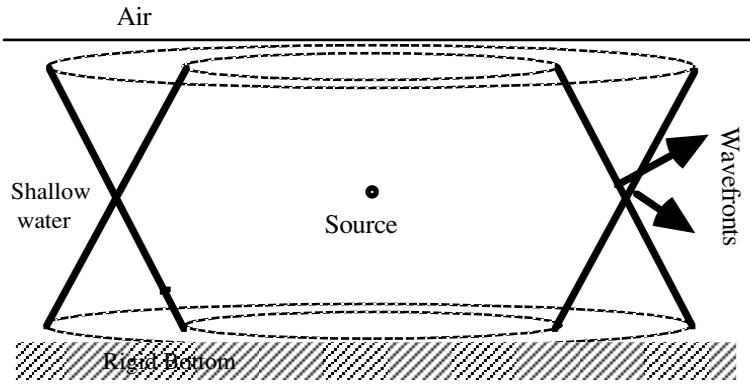
Figure 1.16: A mode can be decomposed into two conical wavefronts. In the vertical plane the wavefronts travel at the angle $\pm\theta_n$, where $\tan\theta_n = \dfrac{\gamma_n}{k_n}$.
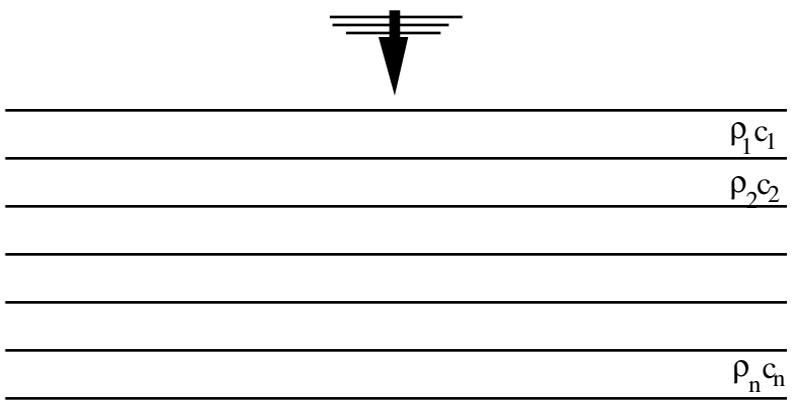
Figure 1.17: A stack of uniform layers. All layers are of the same thickness but with different impedances.

they would interfere with late arriving p-wave signals. A layered medium is modeled as uniform horizontal layers stacked one above the other (see fig. 1.17). A vertically propagating plane wavefront is repeatedly reflected and transmitted at each interface, thus producing a complex reverberation pattern which may be conveniently described within the frame work of filter theory [13].
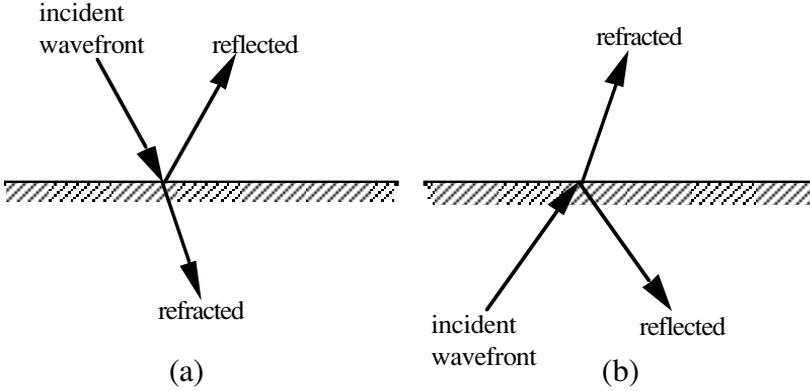
Figure 1.18: Reflected and transmitted waves at an interface. (a) Incident from above and (b) incident from below.

A plane wave of unit amplitude is vertically incident on an interface. It is split into two waves, the first one traveling upwards (reflected wavefront) and the second one traveling downwards (refracted wavefront). Similarly, a plane wave incident from below is split into two plane waves (see fig. 1.18). The amplitude of the reflected wave is equal to $-\hat{r}$ (reflection coefficient) and that of the refracted wave is equal to $\hat{t}$ (transmission coefficient). Now consider a layer sandwiched between two semi-infinite layers. The thickness of the layer is measured in units of return travel time Δt, that is, one unit corresponds to a physical thickness Δh where $\Delta h = \dfrac{c}{2}\Delta t$. Fig. 1.19 shows repeated reflections and refractions at two faces of the layer along with their respective amplitudes. r_0 and t_0 are respectively reflection and refraction coefficients at the top face of the layer; similarly r_1 and t_1 are those at the bottom face of the layer.

Let $f_{refl}(t)$ be the reflected waveform, which consists of a sum of all successively reflected wave components,

$$f_{refl}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_0(\omega) \frac{r_0 + r_1 e^{-j\omega}}{1 + r_0 r_1 e^{-j\omega}} e^{j\omega t} \qquad (1.48a)$$

and similarly the transmitted waveform is given by
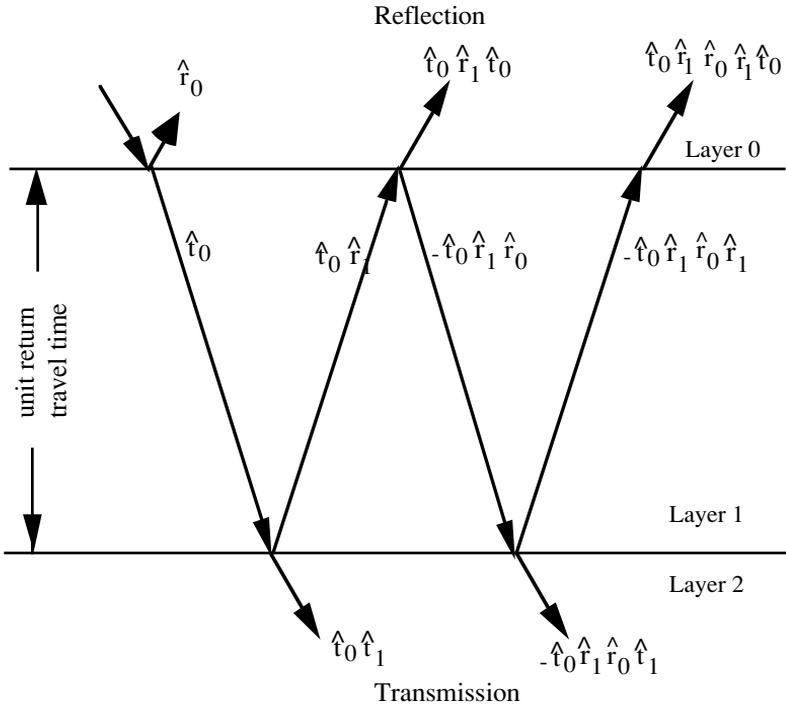
Figure 1.19: Repeated reflections at the two faces of a layer produce a sequence of reflected and transmitted waves. Some amount of energy is also trapped inside the layer. All waves travel vertically though for the sake of clarity the waves are shown as inclined.

$$f_{trans}(t) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} F_0(\omega) \frac{t_0 t_1 e^{-j\frac{\omega}{2}}}{1 + r_0 r_1 e^{-j\omega}} e^{j\omega t} \qquad (1.48b)$$

Let $z = e^{j\omega}$ and define the reflection and transmission response functions as

$$R(z) = \frac{r_0 + r_1 z^{-1}}{1 + r_0 r_1 z^{-1}}, \qquad T(z) = \frac{t_0 t_1}{1 + r_0 r_1 z^{-1}} z^{-\frac{1}{2}} \qquad (1.49)$$

We can now express the z-transforms of the reflected and the transmitted waveforms as

$$F_{refl}(z) = F_0(z)R(z)$$
$$F_{trans}(z) = F_0(z)T(z)$$

It is interesting to note that $T(z)$ has the form of a first order AR process filter and $R(z)$ has the form of a ARMA process filter of order (1,1) [14]. Since $r_0 r_1 \leq 1$, the only pole of the transfer function (1.49) lies within the unit circle, making $R(z)$ and $T(z)$ stable. The signal flow diagrams of transmission and reflection filters are shown in fig. 1.20.

The above characterization of a single layer model has been extended to a multilayer model. The basic structure of the transmission and reflection filters, however, remains unchanged. Further exposition of this approach may be found in [13].

§1.4 **Stochastic Wavefield**:
A signal emitted by a source, even if it is a simple sinusoid, may be modeled as a stochastic process because the signal may be controlled by a set of unknown random parameters, for example, a random phase in a sinusoid. The noise in the array output may consist of thermal noise, the wavefield emitted by numerous sources, either natural or man-made, and transmitted signal which has been scattered by numerous random scatterers or reflectors. In all these models there is an element of stochastic nature which imparts a stochastic character to the wavefield. In this and in the next section we shall study some of these stochastic models of the wavefield from the point of array signal processing. First, we consider the wavefield emitted by a large number of random sources which are distributed in two or three dimensional space, both in open space and bounded space. Basically we shall study the correlation and the spectral characteristics of the output of an elementary array of two sensors. Next, in §1.5, we shall study scattering or reflections produced by many random scatterers which act as obstacles in radio communication or micro-variations of speed and density in underwater detection communication.

1.4.1 Frequency-Wavenumber Spectrum: A stochastic wavefield observed over an infinite line or plane is best characterized in terms of a frequency wavenumber spectrum. For a stochastic wavefield, in place of Fourier representation (1.19), we have spectral representation (on z=0 plane)

$$f(x,y,t) = \frac{1}{8\pi^3} \int \int_{-\infty}^{\infty} \int dF(u,v,\omega)e^{j(\omega t - ux - vy)}$$

which is analogous to the spectral representation of a 1D stationary stochastic process [15]. $dF(u,v,\omega)$ is a differential of the generalized Fourier transform
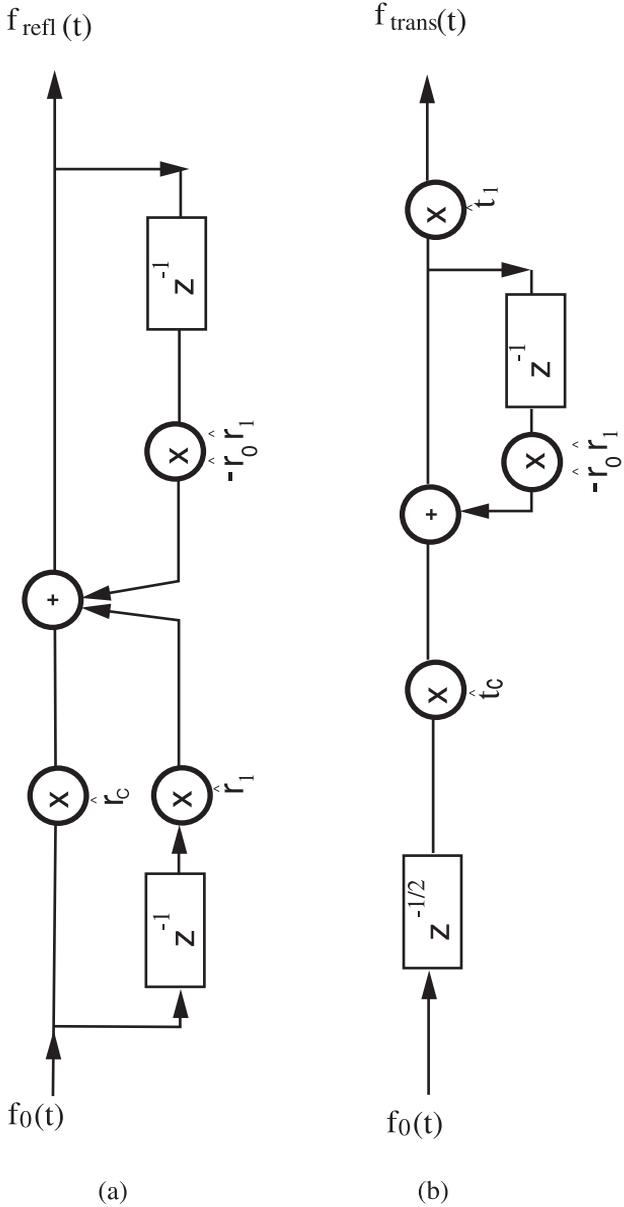
Figure 1.20: The response of a layer to an incident wave is described within the framework of filter theory. (a) reflection filter and (b) transmission filter.

of the stochastic wavefield having the property that $E\{|dF(u,v,\omega)|^2\} \propto S_f(u,v,\omega)$ which we shall use to relate the covariance function to the frequency wavenumber spectrum.

$$C_f(\Delta x, \Delta y, \tau) = \frac{1}{8\pi^3} \int \int\int_{-\infty}^{\infty} S_f(u,v,\omega)e^{j(\omega\tau - u\Delta x - v\Delta y)} du\, dv\, d\omega \qquad (1.50)$$

where $C_f(\Delta x, \Delta y, \tau) = E\{f(x,y,t)f(x+\Delta x, y+\Delta y, t+\tau)\}$ is the covariance function and $S_f(u,v,\omega)$ is the frequency wavenumber spectrum. In the spectrum analysis of time series (1.50) is known as the Wiener-Khinchin relation of great significance [14]. $S_f(u,v,\omega)$ represents power received at a given temporal frequency ω and spatial frequencies u and v. Since u and v are related to the direction cosines, $S_f(u,v,\omega)$ may be looked upon at a fixed temporal frequency as a function of the direction cosines.

1.4.2 <u>Open Space</u>: Consider noise sources in the far field region, distributed over a sphere or a circle (in two dimensions). Each point source emits a stationary stochastic waveform uncorrelated with all other sources. Let $f_i(t)$ be the stochastic waveform emitted by the ith source at angular distance $\varphi_0 + \theta_i$ where θ_i is a random variable uniformly distributed over an interval $\pm\theta_0$ (see fig. 1.21). The signal received at the upper sensor is given by $f_i(t - \frac{d}{2c}\sin(\varphi_0 + \theta_i))$ and that at the lower sensor is $f_i(t + \frac{d}{2c}\sin(\varphi_0 + \theta_i))$ where d is the sensor separation. The total signal obtained by summing over all sources is given by

$$g_1(t) = \sum_i f_i(t - \frac{d}{2c}\sin(\varphi_0 + \theta_i))$$

$$g_2(t) = \sum_i f_i(t + \frac{d}{2c}\sin(\varphi_0 + \theta_i))$$

$$(1.51a)$$

We shall now replace the random function in (1.51a) with its spectral representation [14] and rewrite it as

$$g_1(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \sum_i dF_i(\omega)e^{j(\omega(t - \frac{d}{2c}\sin(\varphi_0 + \theta_i)))}$$
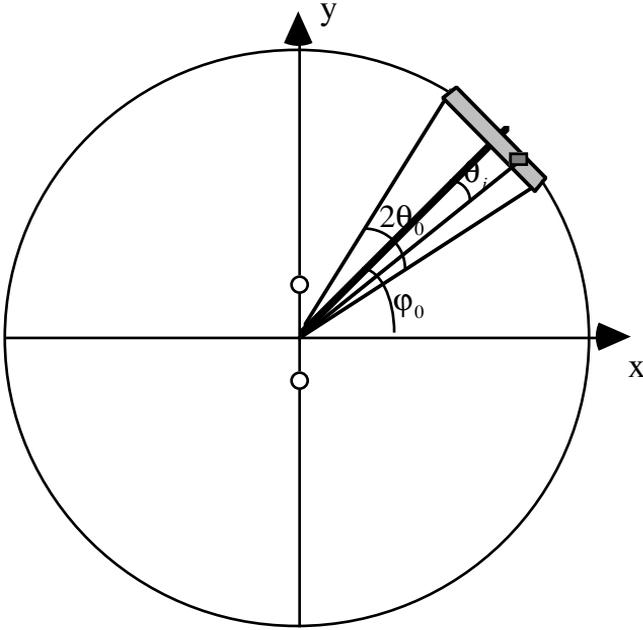
Figure 1.21: Point sources are uniformly distributed over an arc of a large circle in the x, y plane. An elementary array of two sensors is on the y-axis.

$$g_2(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \sum_i dF_i(\omega) e^{j(\omega(t+\frac{d}{2c}\sin(\varphi_0 + \theta_i)))} \tag{1.51b}$$

We will now compute the cross-covariance function between the two outputs,

$$E\{g_1(t)g_2(t+\tau)\}$$

$$= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_i \sum_i E\left\{ \begin{matrix} E\{dF_i^*(\omega)dF_k(\omega')\} \times \\ e^{-j(\omega(t-\frac{d}{2c}\sin(\varphi_0+\theta_i)))} \\ \times e^{j(\omega'(t+\tau+\frac{d}{2c}\sin(\varphi_0+\theta_k)))} \end{matrix} \right\} d\omega d\omega' \tag{1.52a}$$

Noting the properties of the generalized Fourier transforms of stationary processes [14] (1.52a) may be simplified to yield the following result:

$$c_{12}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\sum_{i} S_{ii}(\omega)E\left\{e^{j(\omega(\tau+\frac{d}{c}\sin(\varphi_0+\theta_i))}\right\}d\omega \qquad (1.52b)$$

where we have assumed that the sources are uncorrelated with an identical spectrum. Since θ_i is a uniformly distributed random variable the expected value in (1.52b) may be replaced by an integral

$$c_{12}(\tau) = c_{12}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_0(\omega)d\omega\frac{1}{2\theta_0}\int_{-\theta_0}^{\theta_0}e^{j(\omega(\tau+\frac{d}{c}\sin(\varphi_0+\theta)))}d\theta \quad (1.53a)$$

The integral over θ can be evaluated in the form of a series [16],

$$\int_{-\theta_0}^{\theta_0}e^{j(\omega(\tau+\frac{d}{c}\sin(\varphi_0+\theta)))}d\theta$$

$$= 2\left[\begin{Bmatrix}\cos(\omega\tau)\displaystyle\sum_{n=0,2,4...}\delta_n J_n(\frac{\omega d}{c})\frac{\sin(n\theta_0)\cos(n\varphi_0)}{n}+\\[2mm]\sin(\omega\tau)\displaystyle\sum_{n=1,3,5...}2J_n(\frac{\omega d}{c})\frac{\sin(n\theta_0)\sin(n\varphi_0)}{n}\end{Bmatrix}\\[8mm]j\begin{Bmatrix}\sin(\omega\tau)\displaystyle\sum_{n=0,2,4...}\delta_n J_n(\frac{\omega d}{c})\frac{\sin(n\theta_0)\cos(n\varphi_0)}{n}-\\[2mm]\cos(\omega\tau)\displaystyle\sum_{n=1,3,5...}2J_n(\frac{\omega d}{c})\frac{\sin(n\theta_0)\sin(n\varphi_0)}{n}\end{Bmatrix}\right] \qquad (1.53b)$$

where $\delta_0 = 1$ and $\delta_n = 2$ for all n. In the limiting case of circularly distributed noise sources, $\theta_0 = \pi$, (1.53b) reduces to $2\pi J_0(\frac{\omega d}{c})$.

$$c_{12}(\tau) = \frac{1}{\pi}\int_{0}^{\infty} S_0(\omega)J_0(\frac{\omega d}{c})\cos(\omega\tau)d\omega \qquad (1.53c)$$

The normalized spatial covariance function at zero lag as a function of sensor separation is shown in fig. 1.22. Notice that the sensor outputs become increasingly uncorrelated as the angular extent of the noise sources increases. In the limiting case of circularly distributed noise sources the correlation becomes
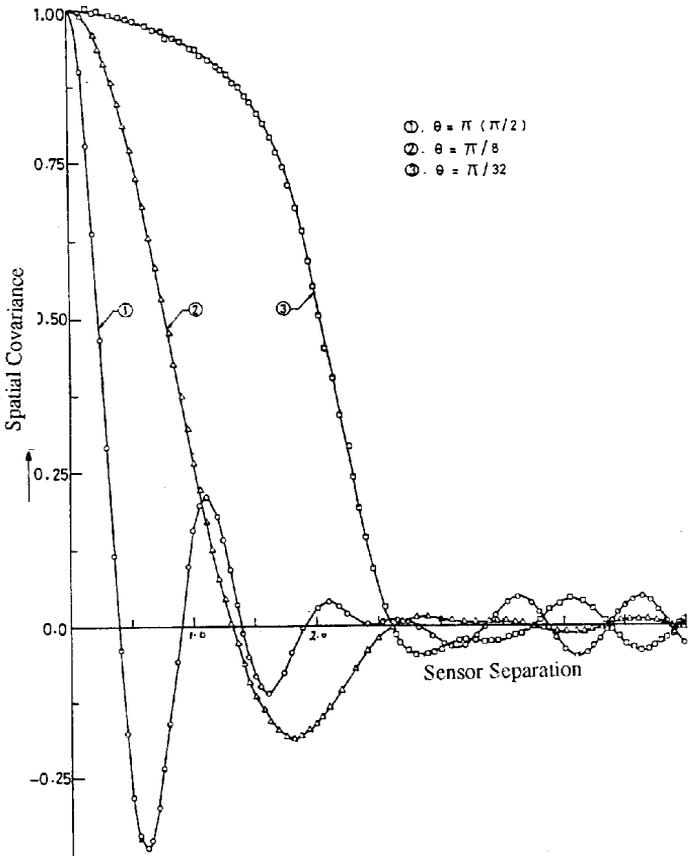
Figure 1.22: Spatial covariance function (at zero lag) as a function of sensor spacing, $\dfrac{d}{\lambda}$. The covariance function is shown for three different angular widths of the distant noise sources. (1) 180^0, (2) 22.5^0, (3) 5.6^0.

negligible even for a separation of the order of one wavelength. When the sources are uniformly distributed on a sphere of large radius, the spatial covariance function is given by $c_{12}(0) = \dfrac{\sin(kd)}{kd}$ [17].

1.4.3 <u>Channel</u>: In this model we shall consider a situation where the noise sources are uniformly distributed on one of the faces of a channel. This situation is close to an ocean channel where all noise sources are on or close to the ocean surface (fig. 1.23) [18].

Noise Sources

Array of Sensors
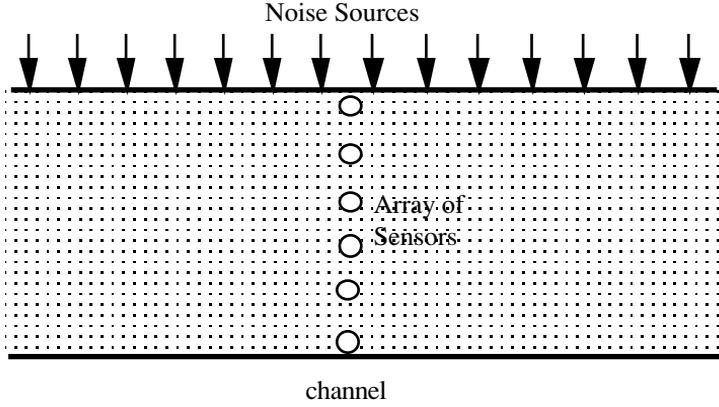
channel

Figure 1.23: The noise sources on the surface of a ocean channel. Correlation structure of wavefield in the vertical direction is investigated.

Let $\psi(r, \varphi, t)$, a homogeneous random function, represent the noise sources on the top face of the channel and $S_\psi(s, \alpha, \omega)$ be its spectrum where s is radial frequency and α is azimuth angle. The acoustic power radiated per unit area at a fixed temporal frequency is given by

$$S_\psi(\omega) = \frac{1}{4\pi^2} \int_0^\infty \int_0^{2\pi} s S_\psi(s, \alpha, \omega) d\alpha ds$$

For wind generated noise on the ocean surface the frequency wavenumber spectrum is approximately modeled as

$$S_\psi(s, \alpha, \omega) = \frac{2p}{k^2} \Gamma(m)(1 - \frac{s^2}{k^2})^{m-1} S_0(\omega) \qquad s < k$$
$$= 0 \qquad s \geq k$$

where p and m are constants $(2 \leq m \leq 3)$ [17, 19]. Note the spectral representation of noise sources, at a fixed point on the plane,

$$\psi(r, \varphi, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\Psi(r, \varphi, \omega) e^{j\omega t} \qquad\qquad (1.54a)$$

and over the entire plane

$$\psi(r,\varphi,t) = \frac{1}{(2\pi)^3} \int\!\!\!\int\!\!\!\int\limits_{-\infty}^{+\infty} d\Psi(u,v,\omega)e^{j(\omega t - ux - vy)} \qquad (1.54b)$$

The pressure field due to a small element of noise sources at a sensor at a depth z_1 is given by (1.45),

$$p_1(t) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \frac{\omega}{2H} d\Psi(r,\varphi,\omega)e^{j\omega t} \sum_m \phi_m(z_1)\phi_m^*(z_s)H_0^1(k_m r)r dr d\varphi \qquad (1.55)$$

where $d\Psi(r,\varphi,\omega)r dr d\varphi$ is the strength of the noise sources. The total pressure is obtained by integrating (1.55) over the entire plane. The spectral representation (1.54b) is used in carrying out the integration.

$$p_1(t) = \frac{1}{(2\pi)^2} \int\limits_{-\infty}^{\infty} \frac{\omega}{2H} e^{j\omega t} \sum_m \overset{\phi_m(z_1)\phi_m^*(z_s)}{\int\!\!\!\int\limits_{-\infty}^{+\infty}} d\Psi(u,v,\omega)\int\limits_0^{\infty} H_0^1(k_m r)J_0(sr)r dr \qquad (1.56)$$

Finally, we note that [17]

$$\int\limits_0^{\infty} H_0^1(k_m r)J_0(sr)r dr = \frac{2}{\pi}\frac{1}{k_m^2 - s^2}$$

and obtain

$$p_1(t) = \frac{1}{2\pi^3} \int\limits_{-\infty}^{\infty} \frac{\omega}{2H} e^{j\omega t} \int\!\!\!\int\limits_{-\infty}^{+\infty} d\Psi(u,v,\omega)\sum_m \frac{\phi_m(z_1)\phi_m^*(z_s)}{k_m^2 - s^2} \qquad (1.57a)$$

The pressure field at another sensor placed at a depth z_2 is given by

$$p_2(t) = \frac{1}{2\pi^3} \int\limits_{-\infty}^{\infty} \frac{\omega}{2H} e^{j\omega t} \int\!\!\!\int\limits_{-\infty}^{+\infty} d\Psi(u,v,\omega)\sum_m \frac{\phi_m(z_2)\phi_m^*(z_s)}{k_m^2 - s^2} \qquad (1.57b)$$

The cross-spectrum between these two sensor outputs may be evaluated as follows

$$\frac{1}{2\pi}S_{12}(\omega)d\omega = E\left\{\frac{1}{2\pi}dP_1(\omega)\frac{1}{2\pi}dP_2^*(\omega)\right\}$$

$$= \frac{\omega^2}{2\pi^3 H^2}\frac{1}{2\pi}d\omega\sum_m\sum_n\int_0^\infty \frac{\phi_m(z_1)\phi_m^*(z_s)\phi_n^*(z_2)\phi_n(z_s)}{(k_m^2 - s^2)(k_n^{*2} - s^2)} sS_\psi(s,\omega)\, ds \qquad (1.58)$$

where $p_1(t) = \dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty}dP_1(\omega)e^{j\omega t}$ and $p_2(t) = \dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty}dP_2(\omega)e^{j\omega t}$.

Simplifying (1.58) we obtain

$$S_{12}(\omega) = \frac{\omega^2}{2\pi^3 H^2}\sum_m\sum_n\int_0^\infty \frac{\phi_m(z_1)\phi_m^*(z_s)\phi_n^*(z_2)\phi_n(z_s)}{(k_m^2 - s^2)(k_n^{*2} - s^2)} sS_\psi(s,\omega)\, ds \qquad (1.59)$$

For spatially white noise sources, $S_\psi(s,\omega) = S_0(\omega)$, and the integral in (1.59) reduces to

$$\int_0^\infty \frac{sS_\psi(s,\omega)}{(k_m^2 - s^2)(k_n^{*2} - s^2)}\, ds = \frac{4}{\pi^2}\ln\left(\frac{k_m}{k_n^*}\right)\frac{S_0(\omega)}{(k_m^2 - k_n^{*2})}$$

For details on the derivation and numerical results the reader is urged to see [20] and [21] where a more general model is dealt with. As an illustration of the general variation of the spectrum and the coherence, the numerical results for a channel where $H = 4\lambda$ are shown in fig. 1.24. The noise sources are located on an annular ring with inner radius = 100λ and outer radius = 1000λ. The noise is presumed to have been generated by wind (m=1). For coherence calculation one sensor is kept fixed at 2λ and the other sensor is moved along the depth axis. We considered two types of bottoms, rigid and soft bottoms ($\rho_1 = 1$gms/cc, $c_1 = 1500$ m/sec, $\rho_2 = 2.0$ gms/cc, $c_2 = 1600$ m/sec). The results are shown in fig. 1.24. The y-axis in fig. 1.24a represents a ratio, $\dfrac{S_z(\omega)}{S_\psi(\omega)}$ where $S_z(\omega)$ is the spectrum of the pressure field at z. It is interesting to note that in the hard bottom channel, because of the trapping of energy, the spectrum is always greater than that in the soft bottom channel and it increases with depth. The coherence as a function of the sensor separation (fig.1.24b) is highly oscillatory, roughly following the interference pattern of vertically travelling modes.
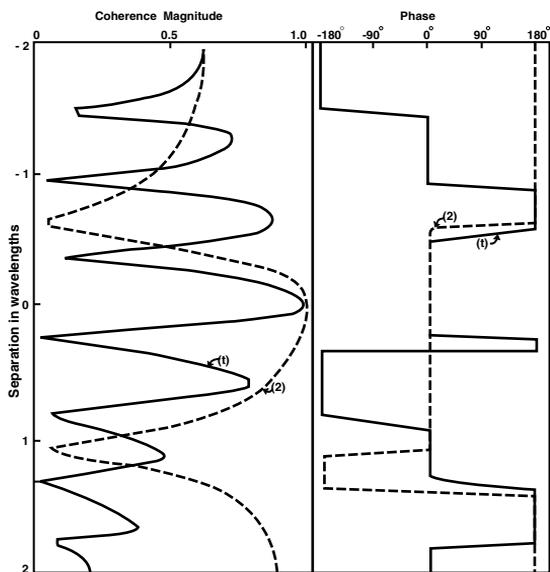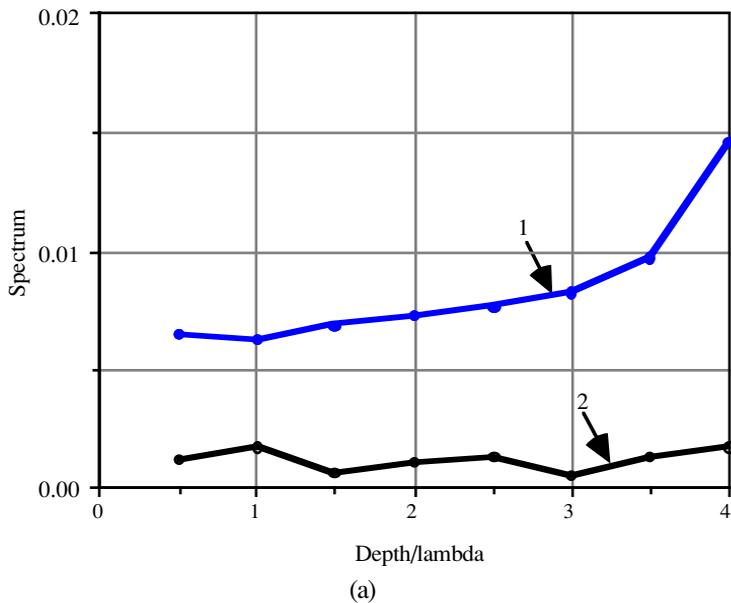
Figure 1.24: Spectrum and Coherence as a function of depth. (a) Spectrum. Empty circles represent soft bottom and the filled circles, hard bottom. (b) Coherence as a function of sensor separation. (1) Hard bottom. (2) Soft bottom. The magnitude is shown in the left panel and phase in the right panel.

The stochastic wavefield in a bounded medium is extremely complex largely due to the interference of multiply reflected waves. What we have described above is a simple example involving just two reflectors. Imagine the complexity of wavefield in a room with six reflectors! Often, in such acomplex situation, we model the wavefield as a diffused wavefield which we have briefly discussed in (1.4.2).

§1.5 **Multipath Propagation:**

When wave energy travels from point A to point B along more than one path, the propagation is said to be multipath propagation. Such a propagation regime is the result of local micro inhomogeneities or point scatterers. Since the local micro inhomogeneities or point scatterers are stochastic in nature it is appropriate to characterize the resulting multipath propagation under stochastic frame work. It often happens that the local inhomogeneities as well as the point scatterers are time varying, consequently multipath propagation is also time varying, sometimes very rapidly as in mobile wireless communication. The phenomenon of bending of a ray as it passes through a blob of inhomogeneity is illustrated in fig. 1.25. Scattering from a point scatterer produces a similar effect but is more pronounced as the wave energy is scattered in all directions.

The dominant effect of the multipath propagation is a large collection of rays, known as microrays [22] impinging on an array from different directions. In a weakly scattering medium a micro-ray follows a path very close to that of an unperturbed ray. Hence all micro-rays would have gone through similar macro variations but with slight path variations. A sensor will then receive a train of coherent signals arriving at different time intervals, possibly Doppler shifted or scaled versions of the transmitted signal, when the scatterers are in motion. The amplitude variation among the microrays is likely to be less and of lower significance. Multipath propagation can cause severe loss of signal fidelity particularly in those problems where the wave propagation is through natural channels, such as in underwater, solid earth, and urban environment, for wireless communication. We shall consider two types of multipath environments. First, we shall consider a random distribution of blobs of inhomogeneity as encountered in underwater channels and in subsurface channels. Next we shall consider a collection of scattering points around a transmitter typically encountered in wireless propagation. There are other possibilities but only the above two environments are well understood. The inhomogeneities in the medium are either stationary or very slowly moving; hence the Doppler effect can be ignored in the first instance.

1.5.1 <u>Random Filter Model</u>: The conceptual picture of many microrays closely following an unperturbed ray and reaching a sensor is mathematically described through the random filter model. The output of a sensor may be expressed as follows:

$$f_n(t) = \frac{1}{N_{rays}} \sum_k a_k g_0(t - t_n - \tau_k - \delta\tau_k(t)) \qquad (1.60)$$

where

$g_0(t)$: waveform transmitted by a source.

t_n: arrival time of unperturbed ray at n^{th} sensor.

τ_k: relative arrival time of kth micro-ray.

$\delta\tau_k(t)$: relative delay due to time varying scattering effect.

a_k: coefficient of attenuation for k^{th} ray.

N_{rays}: Number of rays.

In the frequency domain (1.60) may be expressed as

$$
\begin{aligned}
f_n(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dG_0(\omega) \left[\frac{1}{N_{rays}} \sum_k a_k e^{-j\omega(t_n + \tau_k + \delta\tau_k(t))} \right] e^{j\omega t} \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} dG_0(\omega) H(\omega, t) e^{j\omega t}
\end{aligned}
\qquad (1.61)
$$

where

$$H(\omega, t) = \frac{1}{N_{rays}} \sum_k a_k e^{-j\omega(t_n + \tau_k + \delta\tau_k(t))}$$

is the time-varying random filter representation of the multipath propagation effect. Here both τ_k and $\delta\tau_k(t)$ are modeled as random variables; hence the filter is known as a random time-varying filter. Sea experimental results as reported in [23] indicate that the delays may be approximated as a uniformly distributed random variable and that they are uncorrelated when the sensor separation is of the order of a few hundred feet. Using the above model of fluctuations we can compute the mean and the variance of the time-varying filter transfer function. The mean is given by
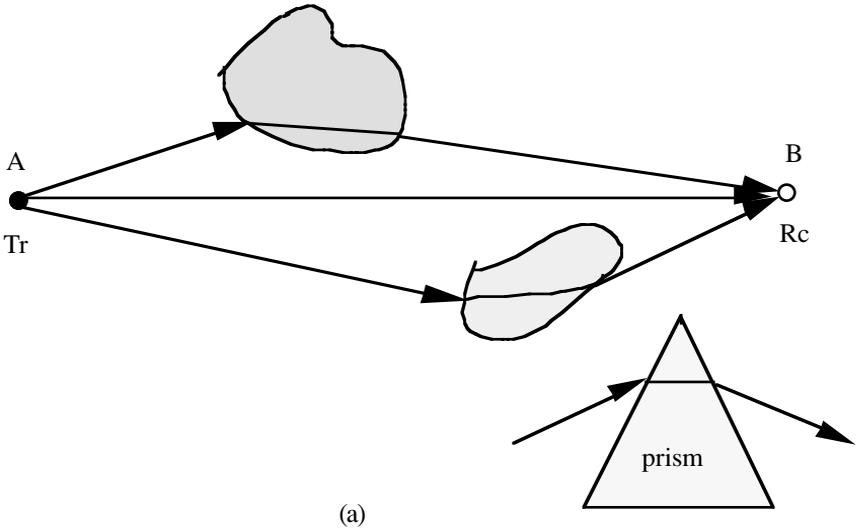
(a)
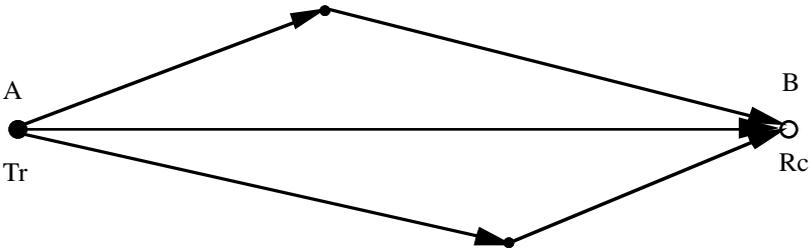
(b)

Figure 1.25: Phenomenon of a ray bending as it passes through a blob of inhomogeneity is shown above. For comparison a ray is shown to bend as it passes through a prism. Scattering by a point scatterer produces a similar effect.

$$E\{H_n(\omega,t)\} = E\left\{\frac{1}{N_{rays}}\sum_k a_k e^{-j\omega(t_n + \tau_k + \delta\tau_k(t))}\right\}$$

$$= \frac{1}{N_{rays}} e^{-j\omega t_n} \sum_k E\{a_k\} E\{e^{-j\omega\tau_k}\} E\{e^{-j\omega\delta\tau_k(t)}\}$$

$$= A_n e^{-j\omega t_n} \Phi_0(\omega)\Phi_1(\omega,t)$$

(1.62)

where $A_n = \dfrac{1}{N_{rays}} \sum_k E\{a_k\}$, $\Phi_0(\omega) = E\{e^{-j\omega\tau_k}\}$ and $\Phi_1(\omega,t) =$

$E\{e^{-j\omega\delta\tau_k(t)}\}$. For a uniformly distributed random variable in the interval

$\pm \dfrac{\Delta t}{2}$,

$$\Phi_0(\omega) = \frac{\sin(\omega\dfrac{\Delta t}{2})}{\omega\dfrac{\Delta t}{2}}$$

1.5.2 <u>Point Scatterers</u>: In radio communication, the electromagnetic waves travelling along straight line ray paths may encounter obstacles which would reflect or scatter the incident wavefield. Since frequency used is very high (800-1000 MHz, $\lambda \approx 1\,meter$) most of the obstacles are likely to be much larger than the wavelength. There will be significant reflections and corner diffractions. The air medium is however assumed to be homogeneous and free from any scattering. Furthermore, a transceiver used in a modern wireless communication system is likely to be in motion causing a significant Doppler shift. One consequence of the reflection or the scattering of waves is the possibility of more than one ray path connecting a transmitter and a receiver. Such multipath propagation in wireless communication is quite common. As a result of this, the different components of the wavefield reach a receiver at slightly different time instants, at different angles and with different Doppler shifts, but coherently. Thus, multipaths are characterized by the following attributes: (i) delay diversity (0-10 micro seconds), (ii) angular diversity (5-6 degrees) and (iii) Doppler shift (0 ± 50 Hz).

<u>Delay diversity</u>: The signals arrive at a receiver (usually a single sensor) at different time intervals, the delay being due to different path lengths. A long delay implies a weak signal due to multiple reflections, attenuation in the air and also by geometrical spreading. The quantity of great interest is the power received at a sensor as a function of delay, known as power delay profile. A typical power delay profile is sketched in fig. 1.26. Referring to this figure we define an excess delay spread τ_e as the delay within which 90 percent of the total power reaches the sensor. There are two other parameters commonly used to characterize a power delay profile,
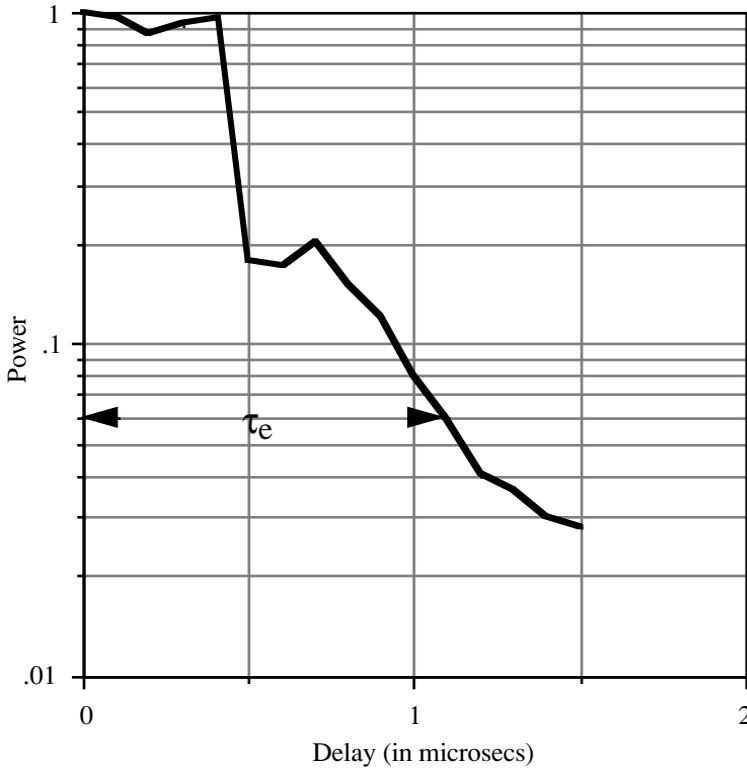
Figure 1.26: A sketch of the power delay profile. Excess delay parameter τ_e is shown.

$$\text{Mean delay } (\overline{\tau}) = \frac{\sum p_k \tau_k}{\sum p_k}$$

$$\text{rms delay spread} = \sqrt{\frac{\sum p_k \tau_k^2}{\sum p_k} - (\overline{\tau})^2} \qquad 1.63)$$

where p_k is the power in the kth path arriving with a delay τ_k. A cumulative plot of power, a plot of $\sum_{k=1}^{m} p_k$ versus $\sum_{k=1}^{m} \tau_k$, is useful in deciphering different groups of multipaths.

Angular diversity: The angle of arrival of different multipaths is a random variable. We define mean angle and rms angle spread as follows:

$$\text{Mean angle } (\overline{\varphi}) = \frac{\sum p_k \varphi_k}{\sum p_k}$$

$$\text{rms angle spread } = \sqrt{\frac{\sum p_k \varphi_k^2}{\sum p_k} - (\overline{\varphi})^2} \qquad (1.64)$$

Doppler Shift: It is common that a transceiver is moving, often at high speed, while the base station antenna is stationary. Consider a speeding car carrying a transceiver past a stationary scatterer (see fig 1.27). The component of car velocity in the direction of scatterer is $v\cos(\varphi)$ where v is the speed of the car moving along the road.

The doppler shift will be given by $\Delta\omega = \dfrac{v}{c}\omega_c \cos(\varphi)$ where ω_c is the carrier frequency. Plot of the Doppler shift as a function angle is shown in fig. 1.27 (inset). As a numerical example, let v=120 km/hour, i.e., 100/3 meters/sec, c=3x10^8 meters/sec, and $\omega_c = 18\pi\, 10^8$ the Doppler shift is equal to $\Delta\omega|_{\max} = 200\pi$.

It is interesting to compute a power delay profile for a simple model of a street lined with buildings which act as obstacles. For simplicity we have assumed regularly spaced building (spacing=15m and gap=5m) on either sides of a street, twenty meters wide (see fig. 1.28a). A stationary source in the middle of the street emits a spike waveform. A suite of rays starting from the source and ending at the base station antenna, which is 200 meters away from the street, were traced taking into effect all possible reflections at different buildings. A sample of five rays thus traced are shown in fig. 1.28a. Perfect reflection was assumed (reflection coefficient =1) but geometrical spreading was taken into account. The power delay profile was computed by noting the power received at computed delay. The computed power delay profile is shown in fig. 1.28b. Other parameters are mean delay=0.042 microsec, rms delay=0.149 microsec and excess delay spread=0.05 microsec.

1.5.3 Local Scattering: A signal emitted by a transceiver is often scattered by point scatterers in the immediate neighborhood of the transceiver. What reaches a distant array is a collection of plane wavefronts, differing in phase, angle of arrival and amplitude but all wavefronts remain correlated. Let $(\theta_0 + \delta\theta_k)$ be the direction-of-arrival (DOA) of the scattered wavefront from the kth scatterer and θ_0 is the nominal DOA of the direct wavefront from the source (see fig. 1.29). We
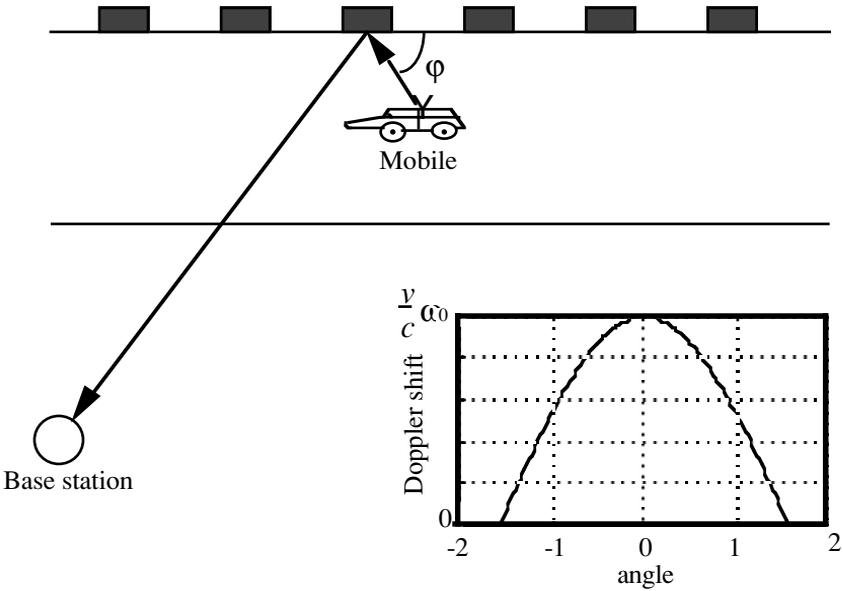
Figure 1.27: A moving car and stationary scatterer will produce a Doppler shift at the base station antenna. Inset shows the Doppler shift as a function of angle.

assume that there are L scatterers in the vicinity of the source. The array output may be expressed as

$$\mathbf{f}(t) = \left[\sum_{k=0}^{L-1} \alpha_k e^{-j\omega_c \delta t_k} \mathbf{a}(\theta_0 + \delta\theta_k) \right] f_0(t) \tag{1.65}$$

where $\mathbf{a}(\theta_0 + \delta\theta_k)$ is the array response vector defined as

$$\mathbf{a}(\theta_0 + \delta\theta_k) =$$

$$[1, e^{-j\frac{\omega_c d}{c}\sin(\theta_0+\delta\theta_k)}, e^{-j\frac{\omega_c 2d}{c}\sin(\theta_0+\delta\theta_k)}, \ldots, e^{-j\frac{\omega_c (M-1)d}{c}\sin(\theta_0+\delta\theta_k)}]$$

for a equispaced (spacing = d) linear array (see chapter 2 for more details on the array response), α_k is the coefficient of scattering (complex) and δt_k is the delay (with respect to direct wavefront) of a wavefront from the kth scatterer. Note that k=0 refers to the direct wavefront, for which $\alpha_0 = 1$ and $\delta t_0 = 0$. It
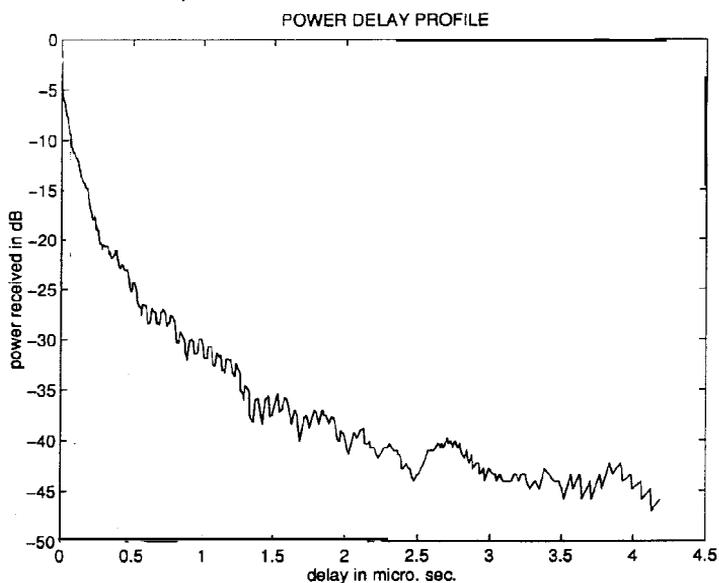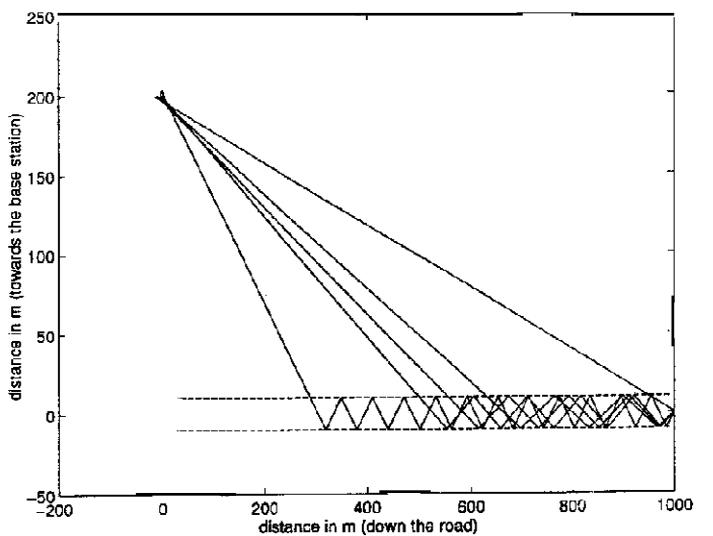
Figure 1.28: A simple model of a street lined with buildings which act as obstacles. (a) A sample of five rays is shown as they propagate through the street. (b) Computed power delay profile.

is assumed that the source emits a narrowband signal, that is, bandwidth $<< \dfrac{2\pi}{\delta t_k}$. We shall express $\mathbf{a}(\theta_0 + \delta\theta_k)$ in Taylor's series expansion,

$$\mathbf{a}(\theta_0 + \delta\theta_k) = \mathbf{a}(\theta_0) + \delta\theta_k \left.\frac{\partial \mathbf{a}(\theta)}{\partial \theta}\right|_{\theta=\theta_0} + \frac{(\delta\theta_k)^2}{2!} \left.\frac{\partial^2 \mathbf{a}(\theta)}{\partial \theta^2}\right|_{\theta=\theta_0} + ...$$

$$\approx \mathbf{a}(\theta_0) + \delta\theta_k \mathbf{a}'(\theta_0)$$

(1.66)

Using (1.66) with the first two terms only in (1.65) we obtain the following approximate result [24]

$$\mathbf{f}(t) \approx \left[\phi_0 \mathbf{a}(\theta_0) + \phi_1 \mathbf{a}'(\theta_0)\right] f_0(t) \qquad (1.67a)$$

and for uniform linear array

$$\mathbf{f}(t) \approx \phi_0 \left[1, (1 + j\frac{\phi_1}{\phi_0} \frac{2\pi d}{\lambda} \cos\theta_0)e^{j\frac{2\pi d}{\lambda}\sin\theta_0}, + ...\right] f_0(t) \qquad (1.67b)$$

where $\phi_0 = \displaystyle\sum_{k=0}^{L-1} \alpha_k e^{-j\omega_c \delta t_k}$ and $\phi_1 = \displaystyle\sum_{k=0}^{L-1} \alpha_k \delta\theta_k e^{-j\omega_c \delta t_k}$. The covariance matrix of the array output is of interest in the DOA estimation. Using the first order approximation in (1.67a), we obtain

$$\mathbf{c}_f \approx L\sigma_\theta^2 \sigma_\alpha^2 \left[\mathbf{a}(\theta_0)\mathbf{a}^H(\theta_0) + \sigma_\theta^2 \mathbf{a}'(\theta_0)\mathbf{a}'^H(\theta_0)\right] \qquad (1.68)$$

where we have assumed that α_k and $\delta\theta_k$ are independent random variables whose variances are σ_α^2 and σ_θ^2, respectively. Note $\mathbf{c}_f$ is a sum of two rank one matrices; hence its maximum rank will be two.

For uniform linear array we can derive a more specific result. For small $\delta\theta$, $\sin\delta\theta \approx \delta\theta$ and $\cos\delta\theta \approx 1$ we have $\left[\mathbf{a}(\theta_0 + \delta\theta)\right]_m = e^{j\frac{2\pi d}{\lambda}m(\sin\theta_0 + \delta\theta\cos\theta_0)}$. The coefficient of scattering is assumed to be uncorrelated. The covariance matrix is given by
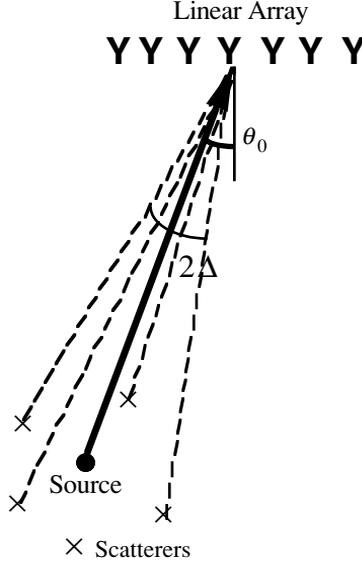
Figure 1.29: Local scattering model. A signal source (transceiver) is surrounded by scatterers. The signal reaching array consists of a suite of plane waves arriving with different DOAs and delays.

$$\left\{ \mathbf{c}_f \right\}_{mn} = L\sigma_{f_0}^2 \sigma_\alpha^2 E\left\{ \mathbf{a}(\theta_0 + \delta\theta)\mathbf{a}^H(\theta_0 + \delta\theta) \right\}_{mn}$$

$$= L\sigma_{f_0}^2 \sigma_\alpha^2 e^{j\frac{2\pi d}{\lambda}(m-n)\sin\theta_0} E\left\{ e^{j\frac{2\pi d}{\lambda}(m-n)\delta\theta\cos\theta_0} \right\} \quad (1.69a)$$

where $\sigma_{f_0}^2$ is the variance of the source signal and σ_α^2 is the variance of the coefficient of scattering. Assuming $\delta\theta$ is uniformly distributed over a range $\pm\Delta$ the expected value in (1.69a) may be shown to be

$$E\left\{ e^{j\frac{2\pi d}{\lambda}(m-n)\delta\theta\cos\theta_0} \right\} = \frac{\sin 2\pi\dfrac{d}{\lambda}\Delta(m-n)\cos\theta_0}{2\pi\dfrac{d}{\lambda}\Delta(m-n)\cos\theta_0} \quad (1.69b)$$

Using (1.69b) in (1.69a) we obtain

$$\mathbf{c}_f \approx L\sigma_{f_0}^2 \sigma_\alpha^2 \mathbf{D}(\theta_0)\mathbf{Q}\mathbf{D}^H(\theta_0) \quad (1.70)$$

where

$$\{\mathbf{Q}\}_{mn} = \frac{\sin 2\pi \dfrac{d}{\lambda} \Delta(m-n)\cos\theta_0}{2\pi \dfrac{d}{\lambda} \Delta(m-n)\cos\theta_0}$$

and

$$\mathbf{D} = diag\left[1, e^{j\frac{2\pi d}{\lambda}\sin\theta_0}, e^{j\frac{4\pi d}{\lambda}\sin\theta_0}, e^{j\frac{6\pi d}{\lambda}\sin\theta_0}, \dots \right].$$

Note that $\mathbf{Q}$ is a symmetric toeplitz matrix with real eigenvalues. Let $\mathbf{Q} = \sum\limits_{i=1}^{M} \lambda_i v_i v_i^H$ be the eigen decomposition of $\mathbf{Q}$ matrix. It is known that $\lambda_1 \approx \lambda_2 \approx \lambda_3 \approx \dots \lambda_r \approx 1$ where r is the rank of $\mathbf{Q}$ and the remaining eigenvalues are insignificant. The rank is approximately given by $r \approx \left[\dfrac{Md}{\lambda} 2\Delta\cos\theta_0\right]$ where [x] stands for the largest integer greater than x. The eigenvectors corresponding to the significant eigenvalues are known as discrete prolate spheroidal sequences (DSSP) [25]. We have computed the rank of the matrix $\mathbf{Q}$ as shown in table 1.1

§1.6 **Propagation through Random Medium**:
Imaging of a medium through which a wave is propagating is one of the important applications of the sensor arrays, in particular, in medical tomography and in seismic exploration where reflected or scattered field is used for imaging. When the inhomogeneous medium is made of simple layers as in subsurface imaging, the reflected field is all that is required for imaging purposes. This is by and large true in seismic imaging which we shall cover in chapter 8. But, when the medium is highly inhomogeneous, the scattering dominates. Currently, only weak scattering has been extensively used for imaging. In this section we would like to review the weak scattering and point out how the scattered field enables imaging of an inhomogeneous medium. Later in chapter 7 we shall describe the use of sensor arrays for imaging.
1.6.1 Acoustic Field: The acoustic pressure field in an inhomogeneous medium satisfies the following wave equation,

$$\nabla^2 f - \frac{1}{c(x,y,z)^2}\frac{\partial^2 f}{\partial t^2} = 0 \qquad (1.71)$$

Δcosθ₀	rank
0.03 rad	2
0.1	7
0.15	10

Table 1.1: Rank of **Q** matrix for different values of $\Delta\cos\theta_0$ and M=64 and λ=d/2.

where c is the acoustic wave speed which is a function of the space coordinates. We shall assume that the density remains constant. (Equation (1.2b) deals with both speed and density variations.) Let $c = c_0 + \delta c(x, y, z)$ where c_0 is the mean wave speed in the medium and $\delta c(x, y, z)$ is the fluctuation around the mean value. A medium is said to be weakly inhomogeneous if $|\delta c(x, y, z)| << c_0$. The wave equation for inhomogeneous medium (1.71) reduces to

$$\nabla^2 f - \frac{1}{c_0^2}\frac{\partial^2 f}{\partial t^2} = -\frac{2\delta c}{c_0^3}\frac{\partial^2 f}{\partial t^2}$$

(1.72)

where the term on the right hand side of (1.72) represents the contribution due to speed fluctuations in the medium. Let us represent $\dfrac{\delta c}{c_0} = \varepsilon\delta\tilde{c}$ where $\delta\tilde{c}$ is a normalized function with unit root mean square (rms) magnitude and ε (<<1) is a constant. Eq(1.72) may be expressed as

$$\nabla^2 f - \frac{1}{c_0^2}\frac{\partial^2 f}{\partial t^2} = -\frac{2\varepsilon}{c_0^2}\delta\tilde{c}\frac{\partial^2 f}{\partial t^2}$$

(1.73)

We shall now try to find a series solution of (1.73). Let the series solution be given by

$$f = f_0 + \varepsilon f_1 + \varepsilon^2 f_2 + ... + \varepsilon^P f_P$$

(1.74)

On substituting (1.74) in (1.73) and equating the coefficients of the powers of ε to zero we obtain the following system of partial differential equations,

$$\nabla^2 f_0 - \frac{1}{c_0^2}\frac{\partial^2 f_0}{\partial t^2} = 0$$

$$\nabla^2 f_1 - \frac{1}{c_0^2}\frac{\partial^2 f_1}{\partial t^2} = -\frac{2}{c_0^2}\delta\tilde{c}\,\frac{\partial^2 f_0}{\partial t^2}$$

... (1.75)

$$\nabla^2 f_p - \frac{1}{c_0^2}\frac{\partial^2 f_p}{\partial t^2} = -\frac{2}{c_0^2}\delta\tilde{c}\,\frac{\partial^2 f_{p-1}}{\partial t^2}$$

We like to solve these equations in an unbounded medium. $f_i(\mathbf{r})$, for i=1, 2, ..., must satisfy the radiation condition. Note that the Green's function is the same for all equations in (1.75); only the driving function on the right hand side differs. The solution of ith equation may be expressed as

$$f_i(\mathbf{r},t) = \frac{1}{4\pi}\int\!\!\!\int\!\!\!\int_{-\infty}^{+\infty} \frac{-\dfrac{2}{c_0^2}\tilde{c}(x',y',z')\dfrac{\partial^2 f_{i-1}}{\partial t^2}}{|\mathbf{r}-\mathbf{r}'|}\,e^{j\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}dx'dy'dz' \quad (1.76)$$

Consider the case of plane wave illumination,
$f_0(\mathbf{r},t) = A_0 e^{j(\omega t - u_0 x - v_0 y - w_0 z)}$, where u_0, v_0, and w_0 are related to the direction cosines of the traveling plane wave (see p. 9). Using the above illumination function in (1.76) we obtain the following expression for the first order term, also known as Born approximation,

$$f_1(\mathbf{r},t) = \frac{1}{4\pi}\int\!\!\!\int\!\!\!\int_{-\infty}^{+\infty} \frac{2k_0^2\delta\tilde{c}\,e^{j(k_0|\mathbf{r}-\mathbf{r}'|)}}{|\mathbf{r}-\mathbf{r}'|}\,f_0(\mathbf{r}',t)dx'dy'dz' \quad (1.77)$$

We shall now use the Sommerfeld formula [26]

$$\frac{e^{j(k_0|\mathbf{r}-\mathbf{r}'|)}}{|\mathbf{r}-\mathbf{r}'|} = \frac{e^{j(k_0\sqrt{\rho^2+z^2})}}{\sqrt{\rho^2+z^2}}$$

$$= \frac{1}{2\pi}\int_0^\infty\!\!\int_0^{2\pi} \frac{\lambda d\lambda}{\sqrt{\lambda^2 - k_0^2}}\,e^{j(\lambda(x-x')\cos\theta + \lambda(y-y')\sin\theta}e^{-\sqrt{\lambda^2-k_0^2}\,|z-z'|}d\theta$$

(1.78)

to simplify (1.77)

$$f_1(\mathbf{r}, t) =$$

$$\frac{k_0^2}{4\pi^2} \int_{-\infty}^{\infty} \int_0^{2\pi} \frac{\lambda d\lambda d\theta}{\sqrt{\lambda^2 - k_0^2}} \left[\int_\Gamma \frac{\delta\tilde{c}(x',y',z') f_0(\mathbf{r}', t)}{e^{j(\lambda(x-x')\cos\theta + \lambda(y-y')\sin\theta)} e^{-\sqrt{\lambda^2 - k_0^2}|z-z'|}} dx' dy' dz' \right]$$

(1.79)

where Γ represents the space occupied by the scattering medium (see fig. 1.30). Define $u' = \lambda\cos\theta$ and $v' = \lambda\sin\theta$. It follows that $du'dv' = \lambda d\lambda d\theta$. Equation (1.79) can be expressed as

$$f_1(x, y, \omega) =$$

$$\frac{-jk_0^2}{4\pi^2} \int_{-\infty}^{+\infty} \int_\Gamma \delta\tilde{c}(x',y',z') e^{-j((u'-u_0)x' + (v'-v_0)y' + (\sqrt{k_0^2 - u'^2 - v'^2} - w_0)z')} dx' dy' dz'$$

$$\frac{e^{-j(\sqrt{k_0^2 - (u'^2 + v'^2)} - w_0)l}}{\sqrt{k_0^2 - (u'^2 + v'^2)}} e^{j(u'x + v'y)} du' dv'$$

(1.80)

Notice that the inner integral over Γ represents a three dimensional Fourier transform of $\delta\tilde{c}(x', y', z')$. Hence, (1.80) can be written in the frequency domain as follows:

$$f_1(x, y, \omega)$$

$$= \frac{-jk_0^2}{4\pi^2} \int_{-\infty}^{+\infty} \frac{\Delta\tilde{c}(u' - u_0, v' - v_0, \sqrt{k_0^2 - u'^2 - v'^2} - w_0)}{e^{-j(\sqrt{k_0^2 - (u'^2 + v'^2)} - w_0)l}} e^{j(u'x + v'y)} du' dv'$$

(1.81)

$$\cdots \frac{e^{-j(\sqrt{k_0^2 - (u'^2 + v'^2)} - w_0)l}}{\sqrt{k_0^2 - (u'^2 + v'^2)}}$$

Note that the factor $e^{j\sqrt{k_0^2 - (u'^2 + v'^2)}l}$ rapidly decays for $(u'^2 + v'^2) > k_0^2$ and $l > 0$. Such waves correspond to evanescent waves, which will be significant only in the immediate neighborhood of the scattering object. The presence of density fluctuation merely introduces an extra term in (1.81). We will not go into the details but cite a reference where the density fluctuation is accounted for [27]. For a two dimensional object the scattered field has been obtained by Kak [28].
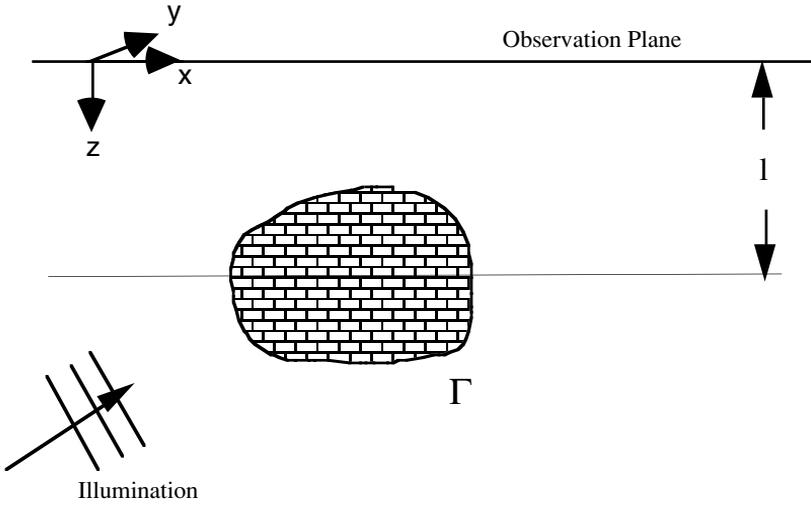
Figure 1.30: Γ represents the space occupied by the scattering medium. $\delta c = 0$ outside Γ. The scattered field is evaluated on observation plane l units above Γ.

$$f_1(x, \omega) = \frac{jk_0^2}{4\pi} \int_{-\infty}^{+\infty} \Delta\tilde{c}(u' - u_0, \sqrt{k_0^2 - u'^2} - v_0) \frac{e^{j(\sqrt{k_0^2 - u'^2} - v_0)l}}{\sqrt{k_0^2 - u'^2}} e^{ju'x} du'$$

(1.82)

1.6.2 <u>Far Field Approximation</u>: When the scattering object is finite and the array is placed quite far from the object we shall apply in (1.77a) the far field approximation, namely, $\dfrac{e^{jk \cdot (r - r')}}{|r - r'|} \approx \dfrac{e^{jk(r - r' \cdot \hat{r})}}{r}$ where $\hat{r}$ is the unit vector in the direction of r. The error due to this approximation, specially in the numerator, is illustrated in fig. 1.31. To assess the quantitative effect, consider the binomial expansion of

$$|r - r'| = \sqrt{r^2 + r'^2 - 2rr'\cos\theta} \approx r(1 + \frac{(\frac{r'}{r})^2 - 2\frac{r'}{r}\cos\theta}{2} + ...)$$

$$= (r + \frac{r'^2}{2r} - r'\cos\theta + ...) \approx (r - r'\cos\theta)$$
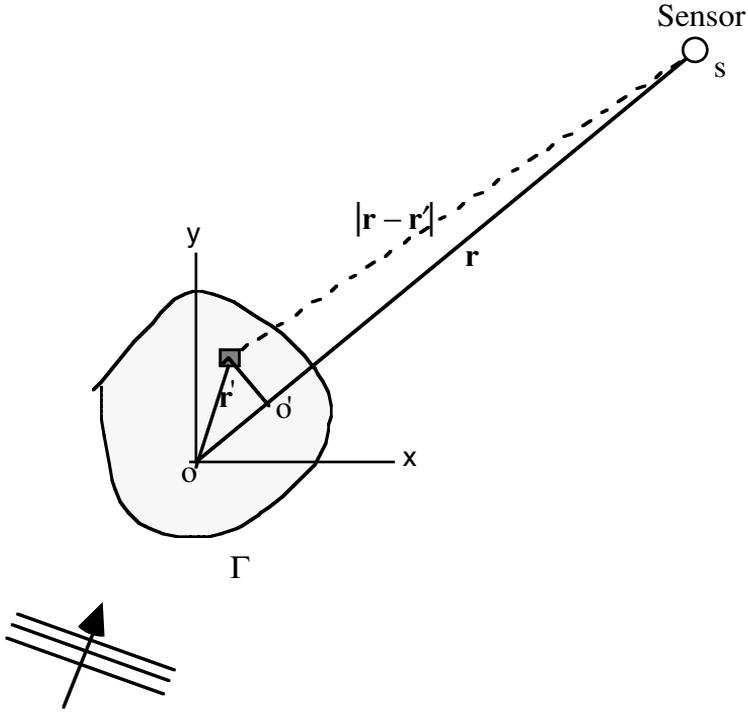
Figure 1.31: In far field approximation $\left|\mathbf{r} - \mathbf{r}'\right|$ is approximated by o's.

The error is of the order of $\dfrac{r'^2}{2r}$ in the first term of binomial expansion. This error will introduce a phase error, $\dfrac{\pi r'^2}{\lambda r}$. For the phase error to be small we must have $r'^2 \ll \lambda r$ or the largest dimension of the object must be much smaller than $\sqrt{\lambda r}$. Using this as far field approximation in (1.77a) we obtain

$$f_1(\mathbf{r}, t) =$$

$$\frac{e^{-j(\omega t - k_0 r)}}{4\pi r} \int_{\Gamma} 2k_0^2 \delta\tilde{c}(x', y', z') e^{-jk_0(\mathbf{r}' \cdot \hat{\mathbf{r}})} e^{-j(u_0 x' + v_0 y' + w_0 z')} dx' dy' dz' \qquad (1.83)$$

In (1.83) we note that

$$e^{-jk_0(\mathbf{r}'\cdot\hat{\mathbf{r}})} = e^{-jk_0(\alpha x' + \beta y' + \gamma z')}$$

$$= e^{-j(ux' + vy' + wz')}$$

where $u = k_0\alpha$, $v = k_0\beta$, $w = k_0\gamma$ and (α,β,γ) are the direction cosines of unit vector $\hat{\mathbf{r}}$. Using this simplification in (1.83) we obtain

$$f_1(\mathbf{r}, t) =$$

$$\frac{e^{-j(\omega t - k_0 r)}}{4\pi r} \int_\Gamma 2k_0^2 \delta\tilde{c}(x', y', z') e^{-j[(u-u_0)x' + (v-v_0)y' + (w-w_0)z']} dx'dy'dz'$$

$$= \frac{e^{-j(\omega t - k_0 r)}}{4\pi r} 2k_0^2 \Delta\tilde{c}(u - u_0, v - v_0, w - w_0) \qquad (1.84)$$

where $\Delta\tilde{c}(.)$ is the Fourier transform of $\delta\tilde{c}$. The result derived in (1.84) has some significance. In the far field region the first order scattered field has the form of a point scatterer (term outside the square brackets), that is, spherical waves. The term inside the square brackets depends on the Fourier transform of the speed fluctuations, evaluated at spatial frequencies determined by the direction of illumination and the direction of sensor. We shall later in chapter 6 exploit this result for reconstruction of speed fluctuations.

1.6.3 <u>Multisource Illumination</u>: The basic fact used in tomography is that when an object is illuminated from different directions the scattered field contains useful information for three dimensional reconstruction. This property of the wavefield is elaborated in this simple example. Consider a source and sensor array on opposite sides of an object to be imaged (see fig. 1.32). Let the mth source be fired and the scattered field be sensed by the sensor array. Let $\mathbf{r}_m^a$ be a vector to the mth source and $\mathbf{r}_n^b$ be a vector to the nth sensor. $\mathbf{r}'$ is vector to a scattering element. The scattered field due to the scattering element at nth sensor is given by

$$\Delta f_1(m, n) = \frac{1}{(4\pi)^2} \frac{2k_0^2 \delta\tilde{c} e^{jk_0|\mathbf{r}_n^b - \mathbf{r}'|}}{|\mathbf{r}' - \mathbf{r}_n^b|} \frac{e^{jk_0|\mathbf{r}' - \mathbf{r}_m^a|}}{|\mathbf{r}' - \mathbf{r}_m^a|} dx'dy'dz' \qquad (1.85)$$

Note that the source and sensor arrays are equispaced linear arrays in the y=0 plane. Therefore, the tips of the vectors $\mathbf{r}_m^a$ and $\mathbf{r}_n^b$ will lie at $[md, 0, -L_a, m = 0, \pm 1, \pm 2, \cdots]$ and $[nd, 0, L_b, n = 0, \pm 1, \pm 2, \cdots]$, respectively. Using Sommerfeld formula (p. 61) to express
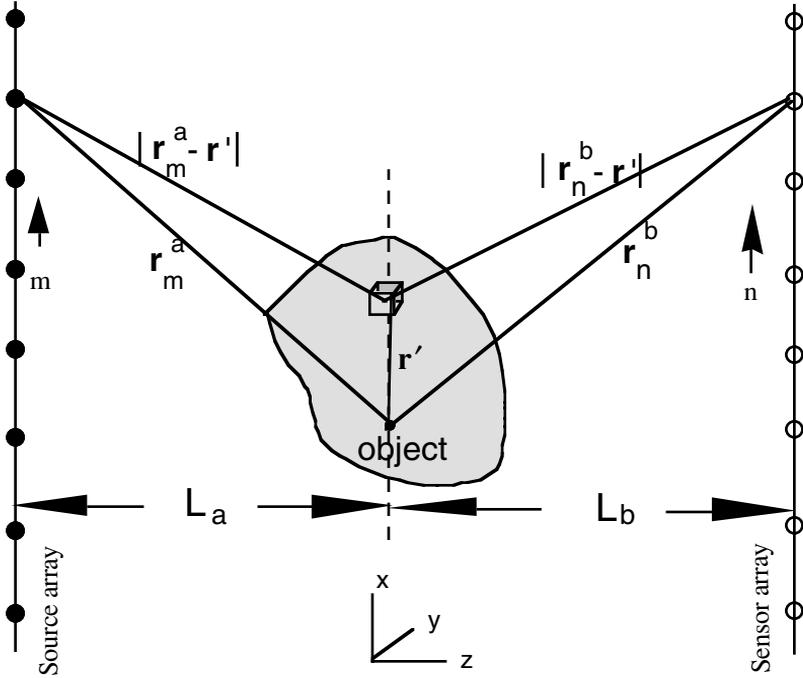
Figure 1.32: An array of sources is used to illuminate an object and the scattered wavefield is received by another array of sensors. The object Fourier transform can be directly related to the 2D Fourier transform of the received wavefield.

$$\frac{e^{jk_0\left|\mathbf{r}_m^a - \mathbf{r}'\right|}}{\left|\mathbf{r}_m^a - \mathbf{r}'\right|} = \frac{-j}{2\pi}\int\int_{-\infty}^{+\infty}\frac{e^{-j\sqrt{k_0^2 - u^2 - v^2}\,(z' + L_a)}}{\sqrt{k_0^2 - u^2 - v^2}}e^{j[(md - x')u - y'v]}\,du\,dv \qquad (1.86a)$$
$$|z'| < L_a$$

$$\frac{e^{jk_0\left|\mathbf{r}_n^b - \mathbf{r}'\right|}}{\left|\mathbf{r}' - \mathbf{r}_n^b\right|} = \frac{-j}{2\pi}\int\int_{-\infty}^{+\infty}\frac{e^{-j\sqrt{k_0^2 - u^2 - v^2}\,(-z' + L_b)}}{\sqrt{k_0^2 - u^2 - v^2}}e^{j[(nd - x')u - y'v]}\,du\,dv \qquad (1.86b)$$
$$|z'| < L_b$$

Next we compute the 2D discrete Fourier transform of $\Delta f_1(m, n)$. From (1.85) it may be seen that the discrete Fourier transform of $\Delta f_1(m, n)$ is equal to the product of discrete Fourier transforms of Green's functions (i.e., left hand side of (1.86)). These are given as follows:

$$DFT\left\{\frac{e^{jk_0\left|\mathbf{r}_m^a-\mathbf{r}'\right|}}{\left|\mathbf{r}_m^a-\mathbf{r}'\right|}\right\}_m$$

$$=\frac{-j}{2\pi}\int\limits_{-\infty}^{+\infty}\int\frac{e^{-j\sqrt{k_0^2-u^2-v^2}\,\left|z'+L_a\right|}}{\sqrt{k_0^2-u^2-v^2}}\delta(ud-u_1)e^{-j[x'u+y'v]}dudv$$

(1.87a)

$$DFT\left\{\frac{e^{jk_0\left|\mathbf{r}_n^a-\mathbf{r}'\right|}}{\left|\mathbf{r}'-\mathbf{r}_n^a\right|}\right\}_n$$

$$=\frac{-j}{2\pi}\int\limits_{-\infty}^{+\infty}\int\frac{e^{-j\sqrt{k_0^2-u^2-v^2}\,\left|-z'+L_a\right|}}{\sqrt{k_0^2-u^2-v^2}}\delta(ud-u_2)e^{-j[x'u+y'v]}dudv$$

(1.87b)

The subscripts m and n on the left hand side refer to discrete Fourier transforms with respect to index m and index n, respectively. Using (1.87) in (1.85) we obtain the 2D Fourier transform of the response of a scattering element

$$\Delta F_1(u_1,u_2)=$$

$$\frac{-k_0^2}{2\pi^2}\int\limits_{-\infty}^{+\infty}\int\left[\begin{array}{c}\dfrac{e^{-j\sqrt{k_0^2-(\frac{u_1}{d})^2-v'^2}\,(z'+L_a)}}{\sqrt{k_0^2-(\frac{u_1}{d})^2-v'^2}}\dfrac{e^{-j\sqrt{k_0^2-(\frac{u_2}{d})^2-v^2}\,(-z'+L_b)}}{\sqrt{k_0^2-(\frac{u_2}{d})^2-v^2}}\\[4mm]\delta\tilde{c}e^{-j[(\frac{u_1}{d}+\frac{u_2}{d})x'+(v'+v)y']}\,dv'dv\end{array}\right]dx'dy'dz'$$

(1.88)

Equation (1.88) is now summed over all scattering elements covering the entire object. We obtain

$$F_1(u_1,u_2)=$$

$$\frac{-k_0^2}{2\pi^2}\int\limits_{-\infty}^{+\infty}\int\Delta\tilde{c}\left(\frac{u_1}{d}+\frac{u_2}{d},\,v+v',\,\left(\sqrt{k_0^2-(\frac{u_1}{d})^2-v^2}-\sqrt{k_0^2-(\frac{u_2}{d})^2-v'^2}\right)\right)$$

$$\times \frac{e^{-j\sqrt{k_0^2-(\frac{u_2}{d})^2-v'^2}\,L_a}}{\sqrt{k_0^2-(\frac{u_1}{d})^2-v'^2}}\frac{e^{-j\sqrt{k_0^2-(\frac{u_1}{d})^2-v^2}\,L_b}}{\sqrt{k_0^2-(\frac{u_2}{d})^2-v^2}}\,dv'dv \qquad (1.89)$$

As a special case we consider an object having a slow variation in the y direction. Then, $\delta C(u,v,w) \approx \delta C(u,w)\delta(v)$ and (1.89) reduces to

$$F_1(u_1,u_2) =$$

$$\frac{k_0^2}{\pi}\int_{-\infty}^{\infty}\Delta\tilde{c}(\frac{u_1}{d}+\frac{u_2}{d},\left(\sqrt{k_0^2-(\frac{u_1}{d})^2-v^2}-\sqrt{k_0^2-(\frac{u_2}{d})^2-v^2}\right))$$

$$\times \frac{e^{-j\sqrt{k_0^2-(\frac{u_1}{d})^2-v^2}\,L_a}}{\sqrt{k_0^2-(\frac{u_1}{d})^2-v^2}}\frac{e^{-j\sqrt{k_0^2-(\frac{u_2}{d})^2-v^2}\,L_b}}{\sqrt{k_0^2-(\frac{u_2}{d})^2-v^2}}\,dv \qquad (1.90)$$

If we were to use a line source in place of a point source (1.90) reduces to, by letting v=0,

$$F_1(u_1,u_2)=$$

$$2k_0^2\Delta\tilde{c}(\frac{u_1}{d}+\frac{u_2}{d},\sqrt{k_0^2-(\frac{u_1}{d})^2}-\sqrt{k_0^2-(\frac{u_2}{d})^2}\ )\times$$

$$\frac{e^{-j\sqrt{k_0^2-(\frac{u_1}{d})^2}\,L_a}}{\sqrt{k_0^2-(\frac{u_1}{d})^2}}\frac{e^{-j\sqrt{k_0^2-(\frac{u_2}{d})^2}\,L_b}}{\sqrt{k_0^2-(\frac{u_2}{d})^2}} \qquad (1.91)$$

The above result is identical to that given in [29].

1.6.4 <u>Scattering of EM Field</u>: The electric properties of a medium are dielectric permitivity (ε), magnetic permeability (μ) and conductivity (σ). The external electromagnetic wavefield will induce electric current in inhomogeneous medium. The induced current will in turn create electromagnetic wavefields outside the medium. The induced current density at a point is given by

$$\mathbf{J}_1(x,y,z) = (\kappa_1^2(x,y,z)-\kappa_0^2)\mathbf{E}(x,y,z) \qquad (1.92)$$

where $\kappa_1^2(x, y, z) = \omega^2 \mu_0 \varepsilon_1(x, y, z) - j\omega\mu_0\sigma_1(x, y, z)$, $\quad \kappa_0^2 = \omega^2 \mu_0 \varepsilon_0$ and $\mathbf{E}(x, y, z)$ is an electric field which induces electric the current [27, 30]. It is assumed that the space outside the inhomogeneous medium is air, and hence $\sigma_0 = 0$. The electric field at a point outside the inhomogeneous medium is given by

$$\mathbf{E}_1(x, y, z, \omega) = j\omega \int_\Gamma \mathbf{J}_1(x', y', z', \omega) \frac{\exp(jk_0|\mathbf{r} - \mathbf{r}'|)}{4\pi|\mathbf{r} - \mathbf{r}'|} dx' dy' dz' \quad (1.93a)$$

and the magnetic field is given by

$$\mathbf{H}_1(x, y, z, \omega) = \int_\Gamma \nabla \times \mathbf{J}_1(x', y', z', \omega) \frac{\exp(jk_0|\mathbf{r} - \mathbf{r}'|)}{4\pi|\mathbf{r} - \mathbf{r}'|} dx' dy' dz' \quad (1.93b)$$

Under Born approximation $\mathbf{E}(x, y, z) \approx \mathbf{E}_0 e^{jk_0(\alpha x + \beta y + \gamma z)}$ where (α, β, γ) are direction cosines of the wave vector and $\mathbf{E}_0$ is the incident electric field. Note that $\mathbf{E}_1$ is in the same direction as $\mathbf{E}_0$ but $\mathbf{H}_1$ is $\perp$ to the incident vector. Let the incident electric field be in the z direction, then the scattered magnetic field will in the (x,y) plane. The x and y components of the magnetic field are given by

$$H_{1x}(x, y, z, \omega) = (1 + jk_0)E_0 \times$$

$$\int_\Gamma (\kappa_1^2(x', y', z') - \kappa_0^2)(y - y') \frac{\exp(jk_0|\mathbf{r} - \mathbf{r}'|)}{4\pi|\mathbf{r} - \mathbf{r}'|^2} dx' dy' dz' \qquad (1.94a)$$

and

$$H_{1y}(x, y, z, \omega) =$$

$$(1 + jk_0)E_0 \int_\Gamma (\kappa_1^2(x', y', z') - \kappa_0^2)(x - x') \frac{\exp(jk_0|\mathbf{r} - \mathbf{r}'|)}{4\pi|\mathbf{r} - \mathbf{r}'|^2} dx' dy' dz'$$

$$(1.94b)$$

Analogous to (1.81) we can write, in the frequency domain, an expression for the scattered electric and magnetic fields as

$$H_{1y}(x,y,z,\omega) = (1 + jk_0)E_0 \times$$

$$\int_{\Gamma} (\kappa_1^2(x',y',z') - \kappa_0^2)(x - x')\frac{\exp(jk_0|\mathbf{r} - \mathbf{r}'|)}{4\pi|\mathbf{r} - \mathbf{r}'|^2} dx'dy'dz' \qquad (1.95a)$$

and

$$H_{1x}(x,y,l,\omega)$$

$$= \frac{E_0}{4\pi^2} \int\int_{-\infty}^{+\infty} \left[\frac{K_1(u' - u_0, v' - v_0, \sqrt{k_0^2 - u'^2 - v'^2} - w_0)}{e^{j(\sqrt{k_0^2 - (u'^2 + v'^2)} - w_0)l}} (jv')e^{j(u'x + v'y)} \right] du'dv' \qquad (1.95b)$$

$$H_{1y}(x,y,l,\omega)$$

$$= \frac{E_0}{4\pi^2} \int\int_{-\infty}^{+\infty} \left[\frac{K_1(u' - u_0, v' - v_0, \sqrt{k_0^2 - u'^2 - v'^2} - w_0)}{e^{j(\sqrt{k_0^2 - (u'^2 + v'^2)} - w_0)l}} (ju')e^{j(u'x + v'y)} \right] du'dv' \qquad (1.95c)$$

where K_1 stands for the 3D Fourier transform of $(\omega^2\mu_0\varepsilon_1(x,y,z) - j\omega\mu_0\sigma_1(x,y,z) - \omega^2\mu_0\varepsilon_0)$. From equations (1.81) and (1.95) it may be inferred that a strong similarity between the scattered acoustic field and the scattered EM field exists.

§1.7 Exercises:

1. Show that if second order approximation is used in (1.66) an additional term will appear in (1.68), namely, $\frac{L}{4}\sigma_\alpha^2 v_\theta^4 \mathbf{a}''(\theta_0)\mathbf{a}''^H(\theta_0)$, where $v_\theta^4 = E\{\delta\theta^4\}$. The pdf of $\delta\theta$ must be symmetric.

2. In equation (1.53c) let the noise be a bandlimited process with a spectrum limited to $\omega_0 - \frac{\delta}{2} \leq \omega \leq \omega_0 + \frac{\delta}{2}$. Show that

$$c_{12}(\tau) = c_{11}(\tau)J_0(\omega_0 \frac{d}{c})$$

under the condition that $\dfrac{d}{c}\delta \ll 1$. What is the implication of this condition?[31].

3. Consider a diffused field in 3D space. One possible model is that point sources are assumed to be uniformly distributed over a large sphere of radius R. Each point source emits a stochastic waveform which is uncorrelated with the waveforms emitted by all other point sources. Show that the field at a sensor placed on the z-axis at a distance $\dfrac{d}{2}$ units from the origin of the coordinates is given by

$$f_1(t) \approx \frac{1}{2\pi} \frac{e^{-j\frac{\omega}{c}R}}{R} \int_{-\infty}^{\infty} dF(\omega,\theta,\varphi) e^{j\omega(t-\frac{d}{2c}\cos\theta)}$$

where θ, φ are respectively elevation and azimuth of the point source. Place another sensor, also on the z-axis, at a distance $-\dfrac{d}{2}$. Show that the cross correlation between the outputs is given by

$$c_{12}(\tau) \approx \frac{1}{2\pi} \int_{-\infty}^{\infty} S_f(\omega)\sin c(\frac{\omega d}{c}) e^{j\omega\tau} d\omega$$

4. Derive the reflection and transmission response of a single layer (shown in fig 1.19) when it is illuminated from below. Derive the reflection response of two layers illuminated from top.

5. In local scattering model show that the covariance matrix of the array output becomes diagonal, that is, the wavefields become spatially uncorrelated when the angular width subtended by the cloud of scatterers at the array is equal to $\dfrac{\lambda}{d\cos\theta_0}$. Consequently, it is not possible to estimate the DOA of the transceivers.

References

1. P. M. Morse and H. Feshbach, *Methods of Theoretical Physics*, vol. I, McGraw-Hill, New York, 1953.
2. J. Li and R. T. Compton, Angle and polarization estimation using ESPRIT with polarization sensitive array, IEEE Trans., AP-39, pp. 1376-1383, 1991.
3. S. A. Levin, Geophysics, vol. 49, No.5, pp. 581-583, 1984.

4. A. Nehorai, Advanced-sensor signal processing *in* Highlights of Statistical Signal and Array Processing, IEEE Signal Proc. Mag., pp.43-45, 1998

5. S. Treitel, P. R. Gutowski, and D. Wagner, Plane wave decomposition of seismograms, Geophy., vol. 47, pp. 1375-1401, 1982.

6. P. M. Morse, Vibrations of elastic bodies; Wave propagation in elastic solids, *in* Handbook of Physics, Eds. E. U. Condon and H. Odishaw, McGraw - Hill Book Co., New York, 1958.

7. K. V. Mackenzie, Long range shallow water transmission, J. Acoust. Soc. Am, vol. 33, pp. 1505-1514, 1961.

8. C. T. Tindle and G. E. J. Bold, Improved ray calculations in shallow water, J. Acoust. Soc. Am, vol. 70, pp. 813-819, 1981.

9. P. S. Naidu and P. G. Krishna Mohan, Signal subspace approach in localization of sound source in shallow water, Signal Proc., vol. 24, pp. 31-42, 1991.

10. Von F. Goos and H. Hanchen, Ein neuer und fundamentaler Versuch zur Totalreflexion, Annales der Physik, vol. 1, 1947, pp. 333-346.

11. C. T. Tindle, Ray calculations with beam displacement, J. Acoust. Soc. Am, vol. 73, pp. 1581-1586, 1983.

12. E. K. Westwood and C. T. Tindle, Shallow water time series simulation using ray theory, J. Acoust. Soc. Am, vol. 81, pp.1752-1761, 1987.

13. E. A. Robinson, Iterative leastsquares procedure for ARMA spectral estimation, *in Nonlinear Methods of Spectral Analysis*, Ed. S. Haykin, Springer Verlag, pp. 127-154, 1979.

14. P. S. Naidu, Modern spectrum analysis of time series, CRC Press, Boca Raton, Fl., USA, 1996.

15. A. H. Yaglom, Introduction to theory of stationary random functions, Prentice-Hall , Englewood Cliffs, 1962.

16. M. Subbarayudu, Performance of the eigenvector (EV) method in the presence of coloured noise, Ph. D. Thesis, Indian Institite of Science, Bangalore, 1985.

17. B. F. Cron and C. H. Sherman, Spatial-correlation functions for various noise models, J. Acoust. Soc. of Am, vol. 34, pp. 1732-1736, 1962.

18. W. A. Kuperman and F. Ingenito, Spatial correlation of surface generated noise field in a stratified ocean, J. Acoust. Soc. Am, vol. 67, pp. 1988-1996, 1980.

19. R. M. Hampson, The theoretical response of vertical and horizontal line arrays to wind generated noise in shallow water, J. Acoust Soc. of Am., vol. 78, pp. 1702-1712, 1985.

20. P. S. Naidu and P. G. Krishna Mohan, A study of the spectrum of an acoustic field in shallow water due to noise sources at the surface, J. Acoust Soc. of Am, vol. 85, pp. 716-725, 1989.

21. P. G. Krishna Mohan, Source localization by signal subspace approach and ambient noise modelling in shallow water, Ph. D. Thesis, Indian Institute of Science, Bangalore, 1988.

22. S. M. Flatte, Wave propagation through random media: Contribution from ocean acoustics, Proc. of the IEEE, vol. 71, pp.1267-1294, 1983.

23. H. W. Broek, Temporal and spatial fluctuations in single-path underwater acoustic wave fronts, I-III, J. Acoust. Soc. of Am, vol. 72, pp. 1527-1543, 1982.

24. D. Asztely, B. Ottersten, and A. L. Swindlehurst, Generalized array manifold model for wireless communication channels with local scattering, IEE Proc. Radar Sonar, Navigation, vol. 145, pp.51-57, 1998.

25. D. Slepian, H. O. Pollack, and H. J. Landau, Prolate spheroidal wave functions, Fourier analysis and uncertainty, Bell Systems Tech. J., vol. 40, pp. 43-84, 1961.

26. A. N. Tychonoff and A. A. Samarski, Differentialgleichungen der Mathematischen Physik, VEB Verlag der Wissenschaften, Berlin, 1959.

27. S. J. Norton, Generation of separate density and compressibility images in tissue, Utrasonic Imaging, vol. 5, pp 240-252 (1983).

28. A. C. Kak, Tomographic Imaging with diffracting and nondiffracting sources, *in Array Processing*, Ed. S. Haykin, Springer-Verlag, pp. 351-433, 1985.

29. Ru-Shan Wu and M. N. Toksoz, Diffraction tomography and multisource holography applied to seismic imaging, Geophy., vol. 52, pp. 11-25, 1987.

30. C. Pichot, L. Jofre, G. Peronnet, and J. C. Bolomey, Active microwave imaging of inhomogeneous bodies, IEEE Trans., AP-32, pp. 416-424, 1985.

31 N. T. Gaarder, The design of point detector arrays, I, IEEE Trans., IT-13, pp. 42-53, 1967.

Chapter Two

Sensor Array Systems

An array of sensors, distributed over a horizontal plane surface, is used to receive a propagating wavefield with the following objectives:

> 1) To localize a source.
> 2) To receive a message from a distant source.
> 3) To image a medium through which the wavefield is propagating.

In this chapter we shall study the basic structure of a sensor array system and in the sequel learn how the above objectives are achieved. The most commonly used array geometries are uniform linear array (ULA) and uniform circular array (UCA). A uniform planar array (UPA) where sensors are placed on an equispaced rectangular grid is more common in large military phased array systems. A wavefront which propagates across the array of sensors is picked up by all sensors. Thus, we have not one but many outputs which constitute an array signal. In the simplest case, all components of the array signals are simply delayed replicas of a basic signal waveform. In the worst case, individual sensor outputs are strongly corrupted with noise and other interference, leaving a very little resemblance among them. Array processing now involves combining all sensor outputs in some optimal manner so that the coherent signal emitted by the source is received and all other inputs are maximally discarded. The aperture of an array, that is, the spatial extent of the sensor distribution, is a limiting factor on resolution. However, the aperture can be synthetically increased by moving a source or sensor. The synthetic aperture concepts are extensively used in mapping radars and sonars. In this chapter we concentrate on sensor array systems which will form the basic material for the subsequent chapters.

§2.1 Uniform Linear Array (ULA):

2.1.1 <u>Array Response</u>: Consider a plane wavefront, having a temporal waveform f(t) incident on a uniform linear array (ULA) of sensors (see fig. 2.1) at an angle θ. In signal processing literature the angle of incidence is also known as direction of arrival (DOA). Note that the DOA is always measured with respect to the normal to array aperture, while another related quantity azimuth, which was introduced in chapter 1, is measured with respect to the x-axis, independent of array orientation. In this work θ stands for DOA and φ stands for azimuth. We shall assume that a source emits a stationary stochastic signal f(t). Let $f_m(t)$, m=0, 1, 2,..., M-1 be the outputs of the sensors. The signal arrives at successive sensors with an incremental delay. The output of the
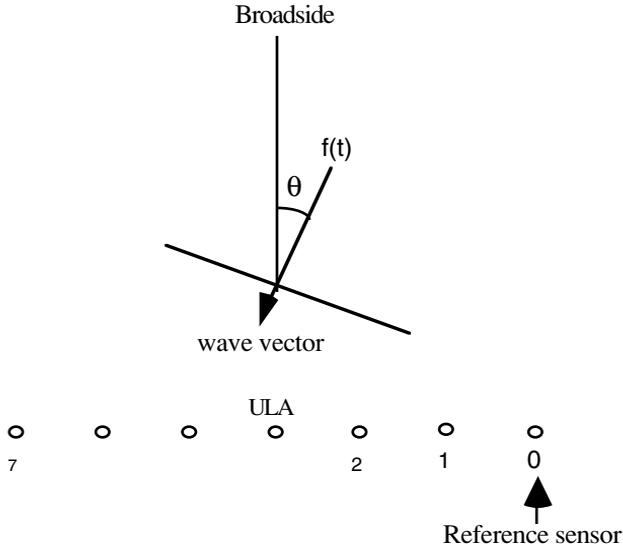
Figure 2.1: Uniform linear array of sensors. Note the convention of sensor indexing. The left most sensor is the reference sensor with respect to which all time delays are measured.

first sensor is $f_0(t) = f(t)$, the output of the second sensor is $f_1(t) = f(t - \Delta t)$ and so on. Thus, the output of the mth sensor is $f_m(t) = f(t - m\Delta t)$. Some times it is convenient to represent the sensor output in the frequency domain

$$f_m(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dF(\omega) e^{j\omega(t - \frac{md}{c}\sin\theta)} \tag{2.1}$$

where we have used the spectral representation of a stationary stochastic process [1]. The simplest form of array signal processing is to sum all sensor outputs without any delay.

$$g(t) = \frac{1}{M} \sum_{m=0}^{M-1} f_m(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dF(\omega) e^{j\omega t} \frac{1}{M} \sum_{m=0}^{M-1} e^{-j\omega \frac{md}{c}\sin\theta}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} dF(\omega) H(\omega\tau) e^{j\omega t} \tag{2.2}$$

where $H(\omega\tau)$ is the array response function, $\tau = \dfrac{d}{c}\sin\theta$, and d is sensor spacing. The array response function for a ULA is given by

$$H(\omega\tau) = \frac{1}{M}\sum_{m=0}^{M-1} e^{j\omega\frac{md}{c}\sin\theta} = \frac{\sin(\frac{M}{2}\omega\tau)}{M\sin\frac{\omega\tau}{2}} e^{j\frac{M-1}{2}\omega\tau} \qquad (2.3a)$$

When the sensor output is weighted with complex coefficients, a_m, m=0,1,...,M-1, the array response becomes

$$H(\omega\tau) = \frac{1}{M}\sum_{m=0}^{M-1} a_m e^{j\omega\frac{md}{c}\sin\theta} \qquad (2.3b)$$

A few samples of the frequency response function (magnitude only) are shown in fig. 2.2 for different values of M, that is, array size. The response function is periodic with a period 2π. The maximum occurs at $\omega\tau = 2n\pi$. The peak at n=0 is known as the main lobe and other peaks at $n = \pm1, \pm2, \ldots$ are known as grating lobes. Since the magnitude of the array response is plotted, the period becomes π as seen in fig. 2.2. The grating lobes can be avoided if we restrict the range of $\omega\tau$ to $\pm\pi$, that is, at a fixed frequency the direction of arrival must satisfy the relation $\dfrac{d}{\lambda}\sin\theta \le \dfrac{1}{2}$. For θ in the interval $\pm\dfrac{\pi}{2}$ this requirement is satisfied if $\dfrac{d}{\lambda} \le \dfrac{1}{2}$. If the range of θ is reduced it is possible to increase the sensor spacing, for example, for $-\dfrac{\pi}{4} \le \theta + \dfrac{\pi}{4}$ the sensor spacing need satisfies the constraint $\dfrac{d}{\lambda} \le \dfrac{1}{\sqrt{2}}$. The phase of the frequency response is a linear function of $\omega\tau$. This useful property of a ULA is lost when the sensors are nonuniformly spaced (see p. 94).

  The array response is a function of the product of frequency ω and delay τ or, more explicitly, $\omega\dfrac{d}{\lambda}\sin\theta$. The implication of this dependence is that two wavefronts whose waveform is a simple sinusoid but with different frequencies (ω_1, ω_2) arriving at different angles (θ_1, θ_2) will produce identical array response if $\omega_1\sin\theta_1 = \omega_2\sin\theta_2$. We shall discuss later such ambiguity issues when we look into the broadband beamformation. The response function has a main lobe which is surrounded by many sidelobes of decreasing magnitude just as we find in spectral windows. The first zero is at
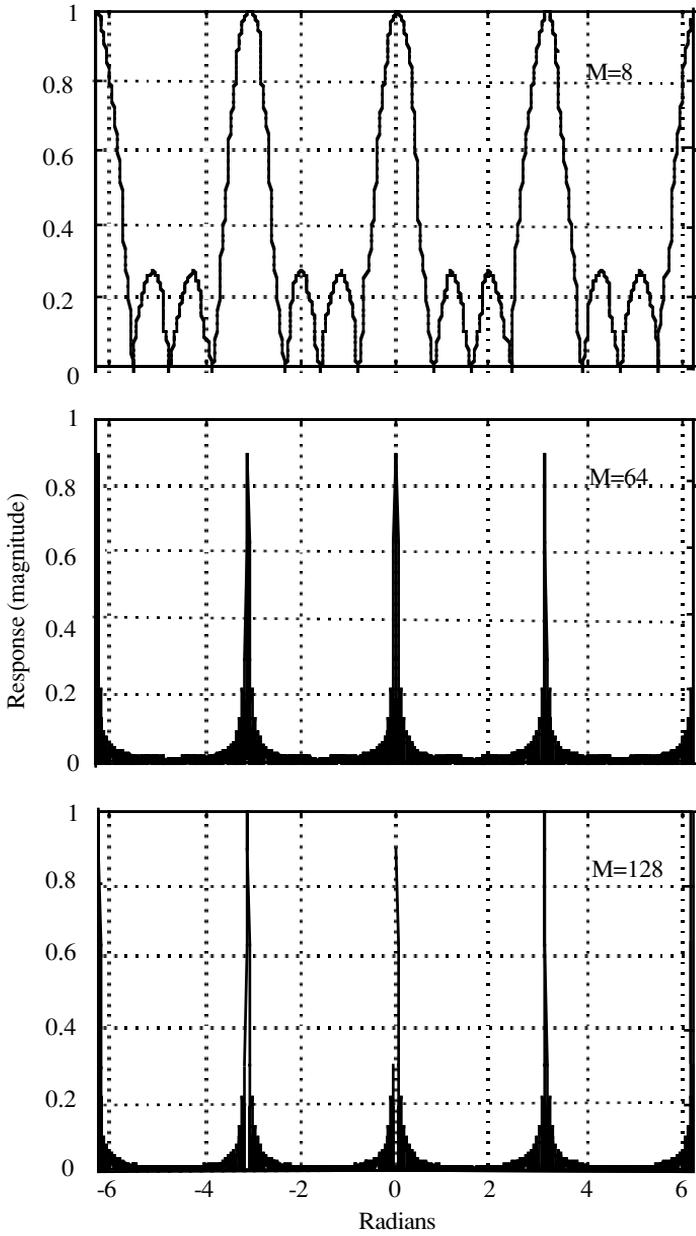
Figure 2.2: Array response function (magnitude) for different values of M. Notice that the main lobe becomes sharper as the array size is increased.

$$\theta_{zero} = \sin^{-1} \frac{\lambda}{Md} \qquad (2.4)$$

which, for large M, becomes inversely proportional to the array length expressed in terms of wavelength. The first sidelobe is 13.5 dB below the main lobe. It is well known that both width of the main lobe and magnitude of the sidelobes can be controlled by using a suitable weight function as in spectrum analysis [1].

2.1.2 <u>Array Steering</u>: We have seen that the array response is maximum when the direction of arrival (DOA) is on broad side ($\theta=0$). The maximum, however, can be changed to any direction through a simple act of introducing a time delay to each sensor output before summation. This is known as array steering. Let an incremental delay of τ per channel be introduced. The sum output of the array is now given by

$$\begin{aligned} g(t) &= \frac{1}{M} \sum_{m=0}^{M-1} f_m(t+m\tau) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dF(\omega)e^{j\omega t} \frac{1}{M} \sum_{m=0}^{M-1} e^{j(\tau-\frac{d}{c}\sin\theta_0)\omega m} \qquad (2.5) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dF(\omega)H((\tau-\frac{d}{c}\sin\theta_0)\omega)e^{j\omega t} \end{aligned}$$

where we have assumed that the DOA is θ_0. Let $\tau = \frac{d}{c}\sin\theta$. Then the array response is maximum whenever $\theta = \theta_0$. We say that the array is steered in the direction θ_0, that is, in the direction of arrival of the incident wavefront. The array response is now a function of DOA. This is demonstrated in fig. 2.3. It is interesting to note that the width of the main lobe increases with increasing DOA. To further understand this broadening effect we shall study the array response function around its maximum, that is, at $\tau = \frac{d}{c}\sin\theta_0$. The first zero will occur at

$$\frac{M}{2}\left[\omega\frac{d}{c}\sin\theta_0 - \omega\frac{d}{c}\sin(\theta_0 - \Delta\theta)\right] = \pi \qquad (2.6a)$$

Upon simplifying (2.6a) we get an equation

$$\sin\theta_0 - \sin(\theta_0 - \Delta\theta) = \frac{\lambda}{Md} \tag{2.6b}$$

whose solution is given by

$$\Delta\theta = \theta_0 - \sin^{-1}\left[\sin\theta_0 - \frac{\lambda}{Md}\right] \tag{2.6c}$$

The dependence of $\Delta\theta$ on the DOA for different array sizes is illustrated in fig. 2.4. The broadening of the main lobe is due to reduction in the array aperture for a wavefront which is incident away from the broadside. The response is maximum whenever

$$\omega\frac{d}{c}(\sin\theta_0 - \sin\theta) = 2\pi n$$

or

$$\frac{d}{\lambda}(\sin\theta_0 - \sin\theta) = n$$

For $\frac{d}{\lambda} \leq \frac{1}{2}$, the acceptable solution is $\theta = \theta_0$ for which n=0. For $\frac{d}{\lambda} > \frac{1}{2}$ there is more than one solution, one for each grating lobe. For example, let $\frac{d}{\lambda} = 1$; a solution of $\theta = \sin^{-1}(\sin\theta_0 - n)$ exists only for n=0 and ± 1.

Now, let $\tau = \frac{d}{c}\frac{k}{M}$. The array response function can be written as a discrete Fourier transform of a complex sinusoid, $\exp(-j\frac{2\pi d}{\lambda}\sin\theta_0 m)$,

$$H(k) = \frac{1}{M}\sum_{m=0}^{M-1} e^{-j\frac{2\pi d}{\lambda}\sin\theta_0 m}\, e^{j\frac{2\pi km}{M}\frac{d}{\lambda}} \tag{2.7}$$

Now $H(k)$ is the kth discrete Fourier transform coefficient which should correspond to the array response at a steering angle, $\sin^{-1}(\frac{k}{M})$. The array response is thus computed only at a set of discrete angles. Since M is finite, usually a few tens, the angular range of $\pm\frac{\pi}{2}$ is coarsely sampled. To overcome this limitation it may be necessary to pad zeros to the complex sinusoid before computing the discrete Fourier transform. However, it must be remembered that
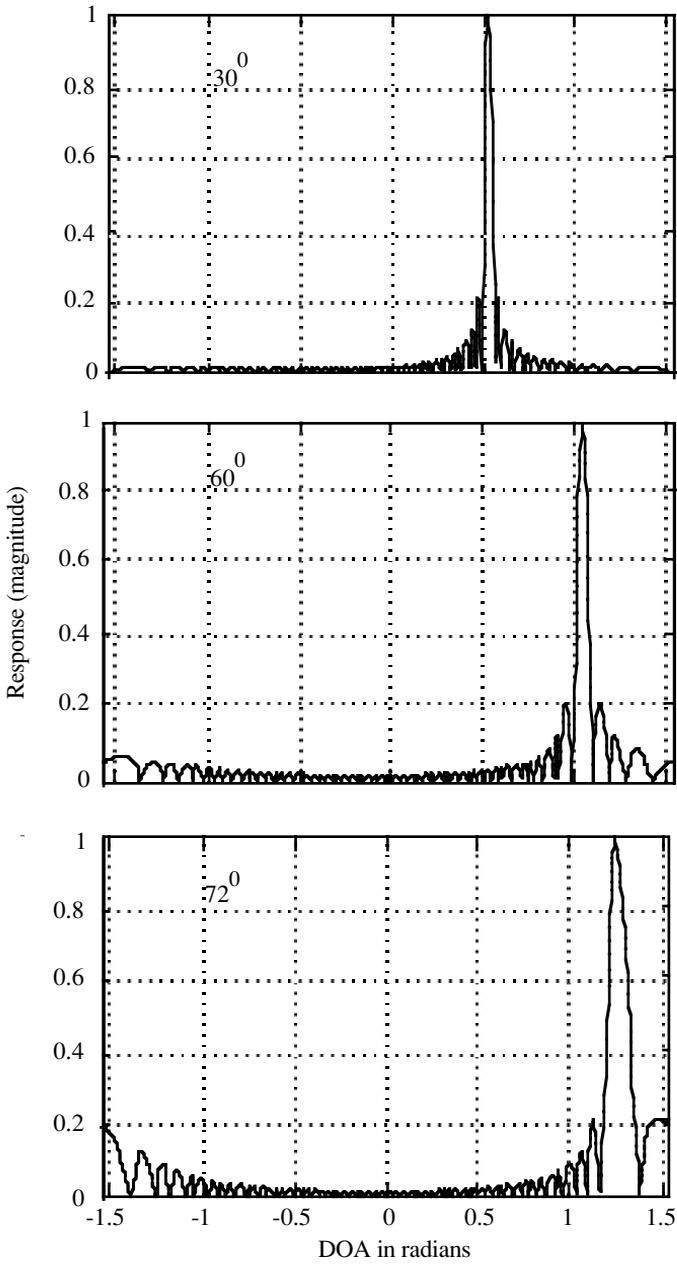
Figure 2.3: The effect of angle of arrival of a wavefront on the array response. The mainlobe broadens and the sidelobes become asymmetric.
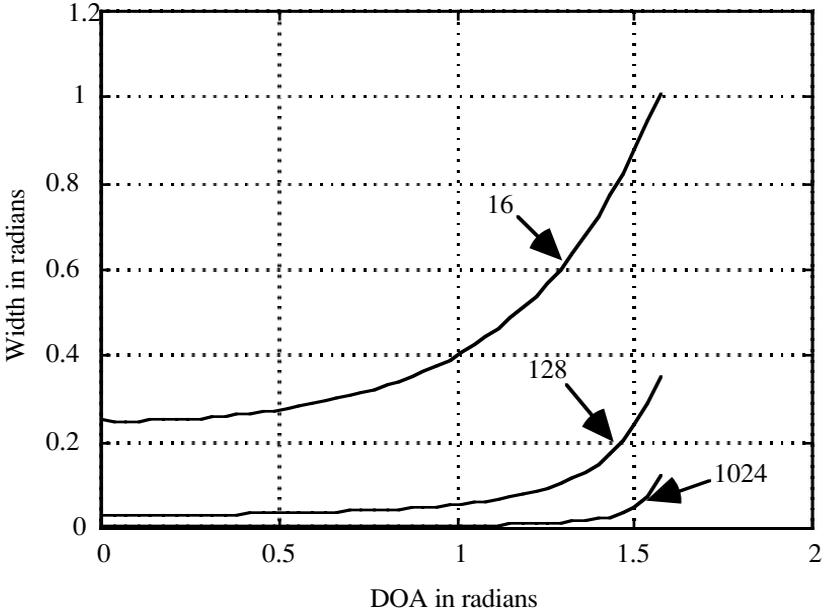
Figure 2.4 Width of the main lobe as a function of DOA for three different sizes of the sensor array (M=16, 128,1024). The sensor spacing is assumed to be λ/2.

this step does not enhance the resolution but only improves the sampling of the otherwise windowed Fourier transform. Use of discrete Fourier transform for beamformation was first suggested in [2].

2.1.3 <u>Broadband Source</u>: Often a remote source such as broadband radar, engine noise, or earthquake, etc. emits a broadband stochastic waveform. The simplest approach to DOA estimation in such a situation is to compute the spectrum of the sum of the sensor outputs. From (2.2) we have

$$g(t) = \frac{1}{M} \sum_{m=0}^{M-1} f_m(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dF(\omega) H(\omega\tau) e^{j\omega t} \qquad (2.8)$$

which may be considered as a spectral representation of the sum of the sensor outputs. Hence its spectrum is given by $S_g(\omega) = S_f(\omega)|H(\omega\tau)|^2$ [1] where $S_f(\omega)$ is the spectrum of the waveform emitted by the source. We can approximate $S_f(\omega)$ by the spectrum of the output of any one sensor. Thus, we obtain

$$|H(\omega\tau)|^2 = \frac{S_g(\omega)}{S_f(\omega)} \tag{2.9}$$

Now consider a plot of $|H(\omega\tau)|^2$ as a function of ω. There is always one peak at $\omega=0$ and a stream of peaks caused by the incident wavefront [3], at positions given by the equation below

$$\omega_{peak} = 2\pi n \frac{c}{d\sin\theta} \tag{2.10a}$$

where n=0, ± 1, ± 2,... We introduce a quantity called minimum array sampling frequency $\omega_{min} = 2\pi\frac{c}{d}$. An array of sensors may be considered as a waveform sampler which samples the waveform as it propagates across the array. The sampling interval is $\frac{d}{c}\sin\theta$ and the maximum interval or minimum sampling frequency occurs when $\theta = \frac{\pi}{2}$. In terms of the minimum array sampling frequency the peak may be written as

$$\omega_{peak} = \frac{\omega_{min}n}{\sin\theta_0} \tag{2.10b}$$

Evidently, ω_{peak} must be in the range $\omega_{min}n \le \omega_{peak} \le \infty$. For the sake of illustration, let the signal spectrum be of infinite width. Now, a plot of $|H(\omega\tau)|^2$ will show an infinite set of peaks spaced at an interval $\frac{\omega_{min}}{\sin\theta}$. For example, for $\theta = 45^0$ an idealised plot of $|H(\omega\tau)|^2$ is shown in fig. 2.5a. A numerical example is shown in fig. 2.5b where we have assumed a 16 sensor ULA with spacing d=15 meters. A broadband signal with bandwidth = (± 200 Hz) is incident at DOA angle equal to 45°. The average spacing of peaks is 42.0 Hz against the theoretical value of 42.43 Hz.

Angular Spectrum: Recall that the frequency wavenumber spectrum of a plane wave is a line passing through the origin with a slope inversely proportional to the direction cosines of the wave vector, in particular, on pages 15-17, chapter 1 we have

$$\omega = \frac{c}{\alpha}u, \quad \omega = \frac{c}{\beta}v \tag{2.11}$$
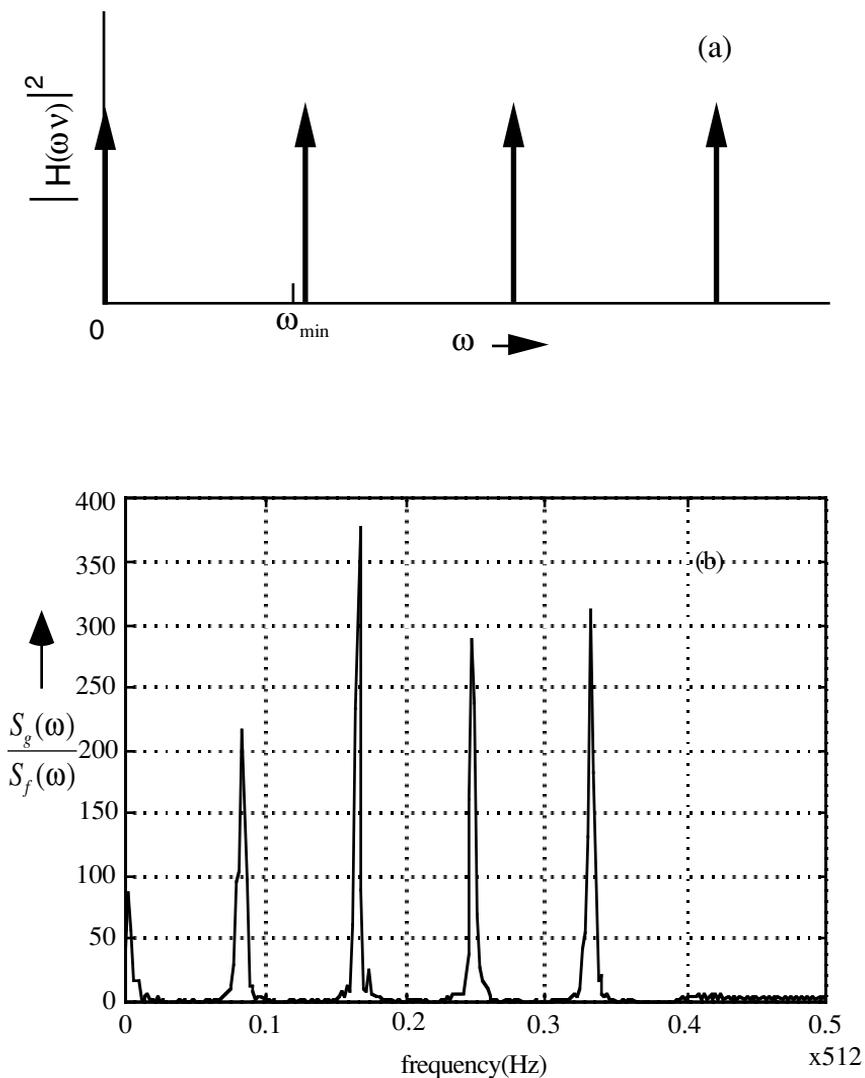
Figure 2.5: (a) A plot of $|H(\omega\nu)|^2$ for $\theta_0 = 41.81^0$. ω_{min} is the minimum array sampling frequency. The position of the first peak (excluding the one at $\omega = 0$) or the interval between the peaks can be used for estimation of DOA. (b) The ratio of the spectrum of the sum of all sensor outputs divided by the spectrum of the first sensor output. There are four peaks of $|H(\omega\nu)|^2$ within the available bandwidth. A 16 sensor ULA (d=15 m) is illuminated by a broadband signal (±200 Hz) incident at 45^0 DOA. Wave speed is 1500m/s.
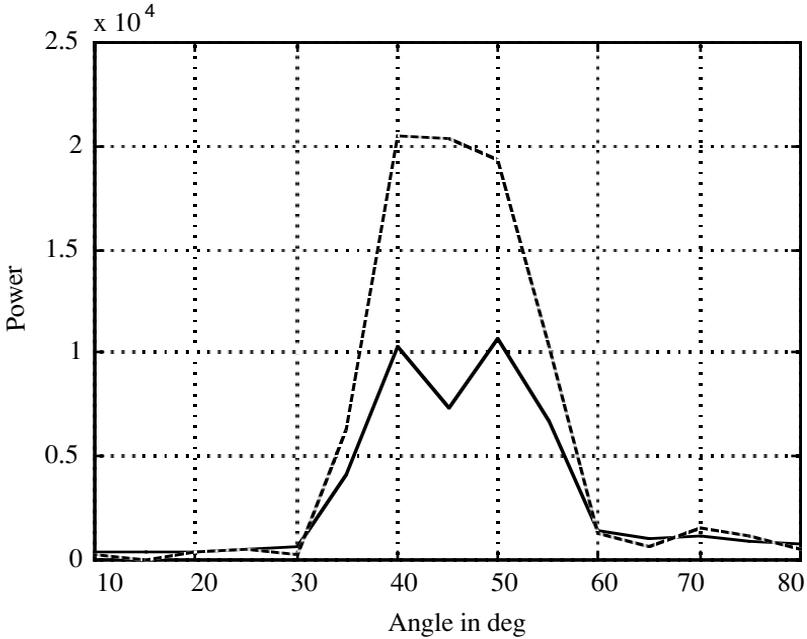
Figure 2.6: Angular spectrum obtained by averaging along the radial lines. Two broadband sources (solid line: 80-120 Hz) and two narrowband sources (dashed line: 98-102 Hz). 16 sensor array, 128 samples with sampling interval 0.005 sec. snr=0dB.

where α and β are direction cosines. Conversely, the spectrum on a line joining the origin with a point in the frequency wavenumber space corresponds to the power of a plane wave propagating with the direction cosines as in (2.11). Consider the frequency wavenumber spectrum of the output of a ULA and integrate the power along a series of radial lines. The integrated power thus obtained is plotted as a function of DOA θ. Such a plot will reveal the presence of plane waves incident on the ULA. The background incoherent noise will tend to average out giving constant or slowly varying power. As an example, we consider a ULA of 16 sensors along x-axis and two incoherent broadband (80 to 120 Hz) sources at DOAs 40° and 50°. The angular spectrum obtained by averaging over different radial lines clearly resolves two peaks but fails to resolve the narrowband signals (see fig. 2.6). Based on the above property of the frequency wavenumber spectrum a method of estimating the direction of arrival by projecting the spectrum along a radial line onto $\omega = \pi$ line has been proposed in [4]. It was mentioned earlier in chapter 1 that the frequency wavenumber spectrum may be considered as a directional spectrum.

Indeed, it is possible to estimate the direction of arrivals by integrating over the temporal frequency band

$$\tilde{S}_f(u,v) = \int_{freq\ band} S_f(u,v,\omega)d\omega \qquad (2.12)$$

Such a method was in fact proposed in [3] where the integrated wavenumber spectrum $\tilde{S}_f(u,v)$ was called zero-delay wavenumber spectrum.

<u>Slant Stacking</u>: Closely related to the idea of array steering (or beamforming) is slant stacking used extensively in seismic exploration. Stacking is also related to the radon transform [5]. Consider a wavefield, f(t,x), where we shall replace t by $\tau + p_x x$ where $p_x = \dfrac{1}{c_x} = \dfrac{u}{\omega}$. The stacking operation is defined as

$$\hat{f}(\tau, p_x) = \int_{-\infty}^{\infty} f(\tau + p_x x, x)dx \qquad (2.13)$$

Let $u = p_x \omega$ for a fixed ω

$$\hat{f}(\tau, p_x) = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty}\int \omega F(\omega, p\omega)e^{j\omega\tau}d\omega dp \int_{-\infty}^{\infty} e^{j(p_x\omega x - p\omega x)}dx$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega, p_x\omega)e^{j\omega\tau}d\omega \qquad (2.14a)$$

Taking the inverse Fourier transform we obtain from (2.14a)

$$\hat{F}(\omega, p_x) = F(\omega, p_x\omega) \qquad (2.14b)$$

Thus, 1D Fourier transform of the stacked output is equal to 2D Fourier transform of the wavefield. Beam steering assumes a plane wave model but stacking does not require this assumption. As shown in (2.14) the stacked output is directly proportional to the spatial Fourier transform of the wavefield, which is equal to array response function (2.7). When the incident wavefield is nonplanar the right thing to do is plane wave decomposition which is achieved through stacking (see (1.33)). Such a situation arises in seismic exploration where a low frequency (wavelengths on the order of a few hundred meters are common) source is used to energize rock strata. The wavefield observed on the surface is a function of two spatial coordinates and time. We shall first derive a result similar to that in (2.14) but for a two-dimensional wavefield.

The stacked output of a planar array is defined as

$$\hat{f}(\tau, p_x, p_y) = \int\limits_{-\infty}^{+\infty}\!\!\int f(\tau + p_x x + p_y y, x, y)dxdy \qquad (2.15)$$

In (2.15) we shall replace the integrand by its Fourier representation,

$$f(t, x, y) = \frac{1}{8\pi^3}\int\limits_{-\infty}^{+\infty}\!\!\int \omega^2 F(\omega, p_x\omega, p_y\omega)e^{j(\omega t - p_x \omega x - p_y \omega y)}d\omega dp_x dp_y$$

where we have used the relations $u = p_x\omega$ and $v = p_y\omega$ and obtain

$$\hat{f}(\tau, p_x, p_y) =$$

$$\int\limits_{-\infty}^{+\infty}\!\!\int dxdy \frac{1}{8\pi^3}\int\limits_{-\infty}^{+\infty}\!\!\int \omega^2 F(\omega, p_x\omega, p_y\omega)e^{j(\omega\tau + (p_x' - p_x)\omega x + (p_y' - p_y)\omega y)}d\omega dp_x' dp_y'$$

$$= \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} F(\omega, p_x'\omega, p_y'\omega)e^{j\omega\tau}d\omega$$

Hence,

$$\hat{F}(\omega, p_x, p_y) = F(\omega, p_x\omega, p_y\omega) \qquad (2.16a)$$

When there is axial symmetry, as it happens in horizontally stratified rocks, (2.16a) takes a different form [6]

$$\hat{F}(\omega, p) = \int\limits_{0}^{\infty} rdr J_0(\omega pr)F(\omega, r) \qquad (2.16b)$$

2.1.4 Matrix Formulation: When the incident signal is a narrowband signal the output of an array, in particular a ULA, may be conveniently represented in a matrix format which reveals some interesting properties. This is also true of a broadband signal but the processing has to be in the frequency domain.
Representation of Narrowband Signals: A narrowband signal $f_{nb}(t)$ may be represented as

$$f_{nb}(t) = s_0(t)\cos(\omega_c t + \varphi_0(t)) \qquad (2.17a)$$

where $s_0(t)$ is a slowly varying waveform, often called envelope and $\varphi_0(t)$ is also a slowly varying phase. $\cos(\omega_c t)$ is a rapidly varying sinusoid, often known as a carrier and ω_c is known as carrier frequency. Many active array systems radiate narrowband signals, for example, a phased array radar. Equation (2.17a) may be expressed as

$$f_{nb}(t) = f_i(t)\cos(\omega_c t) - f_q(t)\sin(\omega_c t)$$

where

$$f_i(t) = s_0(t)\cos(\varphi_0(t))$$
$$f_q(t) = s_0(t)\sin(\varphi_0(t))$$

$f_i(t)$ is known as an inphase component and $f_q(t)$ is a quadrature component. The inphase and quadrature components are uncorrelated. They have the same spectral density function. The inphase and quadrature can be uniquely recovered from a narrowband signal by a process known as mixing which involves multiplication with $2\cos(\omega_c t)$ and $-2\sin(\omega_c t)$ and low pass filtering [7]. A complex analytical signal is defined as $f_c(t) = f_i(t) + jf_q(t)$. Consider a narrowband signal delayed by one quarter period. Assuming that both inphase and quadrature components are slowly varying signals we get the following approximate result:

$$f_{nb}((t - \frac{\tau_0}{4})) = f_i(t)\cos(\omega_c(t - \frac{\tau_0}{4})) - f_q(t)\sin(\omega_c(t - \frac{\tau_0}{4}))$$
$$\approx f_i(t)\sin(\omega_c t) + f_q(t)\cos(\omega_c t)$$
$$= f_{nb}^{Hilb}(t)$$

We define a complex analytical signal as

$$f_{nb}(t) + jf_{nb}(t - \frac{\tau_0}{4}) = f_i(t)e^{j\omega_c t} + jf_q(t)e^{j\omega_c t}$$
$$= f_c(t)e^{j\omega_c t}$$

(2.17b)

The representation given by (2.17b) is useful in narrowband beamformation. The process described in (2.17b) is often referred to as quadrature filtering, which is illustrated in fig. 2.7. Note that the input to quadrature filter is real but the output is complex.
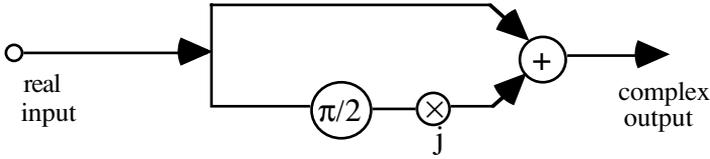
Figure 2.7. Quadrature filter structure. Since the phase change due to propagation appears in the complex sinusoid at the output it is easy to introduce phase adjustments for beamformation.

Consider the mth sensor of a ULA. The complex output of the quadrature filter is

$$f_m(t) = f_{nb}(t - m\frac{d}{c_x}) + jf_{nb}(t - m\frac{d}{c_x} - \frac{\tau_0}{4})$$

$$= f_c(t)e^{j\omega_c t - jm\omega_c \frac{d}{c_x}}$$

(2.17c)

The term representing propagation delay now appears in the complex sinusoid. Naturally, in order to form a beam, it is convenient to introduce phase adjustments. Let $w_0, w_1, \ldots, w_{M-1}$ be a set of complex weight coefficients for beamformation. The beam output will be given by

$$\text{output} = f_c(t) \sum_{m=0}^{M-1} w_m e^{j\omega_c t + jm\omega_c \frac{d}{c_x}} .$$

Through complex weight coefficients it is possible to adjust both amplitude and phase so that the resulting response is closest to any desired response.

Matrix: A snapshot is a vector representing the outputs of all sensors taken at the same time instant t. Let $\mathbf{f}(t) = col\{f_0(t), f_1(t), \ldots, f_{M-1}(t)\}$ be a snapshot, where $f_0(t), f_1(t), \ldots, f_{M-1}(t)$ stand for the sensor outputs at time t. When the incident signal is narrowband, the signal varies slowly with time (assume that the carrier has been removed). In the noise-free case a single time shot is adequate as it contains all available information. A snapshot vector for narrowband signal may be expressed using (2.17c) as

$$\mathbf{f}(t) = f_c(t)col\left\{1, e^{-j\omega_c \frac{d}{c_x}}, ..., e^{-j(M-1)\omega_c \frac{d}{c_x}}\right\} \qquad (2.17d)$$

$$= f_c(t)\phi(\theta_0)$$

where, it may be recalled that, the apparent speed $c_x = \dfrac{c}{\sin\theta_0}$. Further let the sensor response matrix be $\alpha(\theta_0) = diag\{\alpha_0(\theta_0), \alpha_1(\theta_0), ...\alpha_{M-1}(\theta_0)\}$, in which each element represents the response of a sensor as a function of the angle of incidence of the wavefront. $\phi(\theta_0)$ represents the propagation effect of the medium on a wavefront propagating across the array. $\phi(\theta_0)$ and $\alpha(\theta_0)$ together form a direction vector $\mathbf{a}(\theta_0) = \alpha(\theta_0)\phi(\theta_0)$ representing the response of an array to a wavefront incident at angle θ_0 (DOA). Finally, the array output may be expressed as follows:

$$\mathbf{f}(t) = f_c(t)\alpha(\theta_0)\phi(\theta_0)$$

$$= f_c(t)\mathbf{a}(\theta_0) \qquad (2.17e)$$

When there are P narrowband sources radiating simultaneously the array output may be expressed as a linear combination of P terms of the type shown in (2.17e)

$$\mathbf{f}(t) = \begin{bmatrix} \alpha(\theta_0)\phi(\theta_0), \alpha(\theta_1)\phi(\theta_1), ..., \alpha(\theta_{P-1})\phi(\theta_{P-1}) \\ (M \times P) \end{bmatrix} \begin{bmatrix} f_{c_0}(t) \\ f_{c_1}(t) \\ . \\ . \\ . \\ f_{c_{P-1}}(t) \\ (P \times 1) \end{bmatrix} + \eta(t) \qquad (2.18a)$$

where $\eta(t)$ is the noise vector assumed to be uncorrelated with the signal terms. Equationq (2.18a) may be written in a more compact form where P columns

$$\mathbf{f}(t) = \mathbf{As} + \eta(t) \qquad (2.18b)$$

of **A** matrix are P direction vectors pointing to P sources. The matrix representation of the array output model as in (2.18b) plays a very crucial role in the development of high resolution methods for DOA estimation.

The array steering can also be represented in terms of a matrix operation. To steer an array to a desired direction, θ, we form an inner product of the steering vector and the array snapshot

$$\mathbf{a}^H(\theta)\mathbf{f}(t) = f_c(t)\mathbf{a}^H(\theta)\mathbf{a}(\theta_0) \qquad (2.19a)$$

The output power is given by

$$\mathbf{a}^H(\theta)\mathbf{C}_f\mathbf{a}(\theta) = \sigma_{s_0}^2 E\left\{\left|\mathbf{a}^H(\theta)\mathbf{a}(\theta_0)\right|^2\right\}$$

$$= \sigma_{s_0}^2 M^2 \left|H(\omega\frac{d}{c}(\sin\theta - \sin\theta_0))\right|^2 \qquad (2.19b)$$

where $\mathbf{C}_f = E\left\{\mathbf{f}(t)\mathbf{f}^H(t)\right\}$ is the spatial covariance matrix (SCM). Whenever $\theta = \theta_0$, that is, when the steering angle is equal to the DOA, the left hand side of (2.19b) equals $\sigma_{s_0}^2 M^2$ giving the power of the source.

The M dimensional steering vector will span an M-dimensional space known as an array manifold. The tip of the steering vector traces a closed curve in the array manifold or a closed surface when the steering vector is a function of two variables, for example, azimuth and elevation. Consider the case of identical sensors, that is,

$$\alpha_0(\theta_0) = \alpha_1(\theta_0) = \ldots = \alpha_{M-1}(\theta_0) = \alpha(\theta_0)$$

In this case the direction vector is given by

$$\mathbf{a}(\theta_0) = \alpha(\theta_0)\text{col}\{1, e^{-j\omega\frac{d}{c}\sin\theta_0}, e^{-j\omega\frac{2d}{c}\sin\theta_0} \ldots e^{-j\omega\frac{(M-1)d}{c}\sin\theta_0}\}$$

In the event of sensors being omnidirectional, that is, $\alpha(\theta_0) = $ constant, the array manifold becomes a closed curve on a sphere (in M-dimensional space). For uniqueness the array manifold must not intersect; otherwise, at the point of intersection, the steering vector will point to two different directions, θ_1 and θ_2, such that $\mathbf{a}(\theta_1) = \mathbf{a}(\theta_2)$. Such a possibility exists only when $\frac{d}{\lambda} > 0.5$. To show this, consider the steering vector for omnidirectional sensors. Let θ_1 and θ_2 be two such directions for which $\mathbf{a}(\theta_1) = \mathbf{a}(\theta_2)$, that is, for all m

$$e^{j2\pi\frac{d}{\lambda}m\sin\theta_1} = e^{j2\pi\frac{d}{\lambda}m\sin\theta_2}$$

This is possible when

$$\frac{d}{\lambda}[\sin\theta_1 - \sin\theta_2] = 1$$

or

$$\sin\theta_1 = \frac{\lambda}{d} + \sin\theta_2 \qquad\qquad (2.20)$$

A solution of (2.20) exists only when $\frac{\lambda}{d} < 2$; for example, when $\frac{\lambda}{d} = 1.2$ the following pairs of directions are the possible solutions: $(36.87^o, -36.87^o)$, $(23.58^o, -53.13^o)$, and $(11.54^o, -90^o)$.

The steering vector satisfies the following properties:

(a) $\mathbf{a}(\theta) = \mathbf{a}(\pi - \theta)$

(b) $\mathbf{a}^*(\theta) = \mathbf{a}(-\theta)$

(c) $\mathbf{a}(\theta)$ is periodic with a period $\pm\frac{\pi}{2}$ only if $d = \frac{\lambda}{2}$.

Property (a) implies a wavefront coming from the north and another symmetrically opposite from the south (a and b in fig 2.8) cannot be distinguished (north-south ambiguity). Property (b) implies a wavefront coming from the east and another symmetrically opposite from the west (a and c in fig. 2.8) can be distinguished only if the signal is complex (east-west ambiguity). To show this recall (2.17a) and compare the outputs of a ULA for a real input signal incident at angle θ and $-\theta$. Let $\mathbf{f}_\theta(t)$ be output of a ULA for an incident angle, θ, and $\mathbf{f}_{-\theta}(t)$ be the output for an incident angle, $-\theta$. For a real signal $\mathbf{f}_\theta(t) = \mathbf{f}^*_{-\theta}(t)$ but for a complex signal $\mathbf{f}_\theta(t) \neq \mathbf{f}^*_{-\theta}(t)$. Property (c) implies that there is no grating lobe in the range $\pm\frac{\pi}{2}$ when the sensor spacing is $d \leq \frac{\lambda}{2}$.

The steering vector is closely related to the array response function. To show this we define a unit vector, $\mathbf{1}=\text{col}\{1, 1, 1,...,1\}$, and consider a dot product
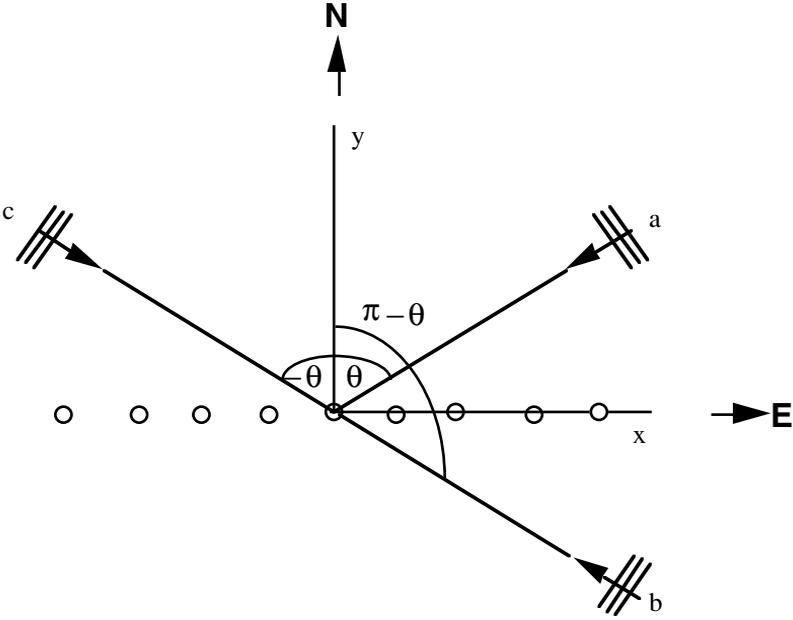
Figure 2.8: A ULA cannot distinguish wavefronts a and b (north-south ambiguity). However, it can distinguish wavefronts a and c if the signal is complex (east-west ambiguity).

$$\mathbf{a}(\theta)^H \mathbf{1} = \sum_{m=0}^{M-1} e^{j\omega m \frac{d}{c} \sin(\theta)} \qquad (2.21a)$$

which follows from (2.3). We have assumed for the sake of simplicity that all sensors are identical and omnidirectional. In real array, the individual sensor response is likely to be directional and varying. Then, the array response is given by

$$\mathbf{a}(\theta)^H \mathbf{1} = \sum_{m=0}^{M-1} \alpha_m(\theta) e^{j\omega m \frac{d}{c} \sin(\theta)} \qquad (2.21b)$$

An array is steered to a desired direction by introducing delays to each sensor output. The response of such a steered array is given by (from (2.5))

$$H(\tau - \omega\frac{d}{c}\sin\theta_0) = \frac{1}{M}\sum_{m=0}^{M-1}e^{j(\tau-\frac{d}{c}\sin\theta_0)\omega m}$$

$$= \frac{1}{M}\mathbf{a}^H(\theta)\mathbf{a}(\theta_0) \qquad (2.22)$$

where θ is the desired direction to which the array is steered and θ_0 is the direction of arrival of a wavefront. The response of the steered array is expressed as an inner product of the steering vector and direction vector as shown in (2.22).

2.1.5 <u>Nonuniform Linear Arrays</u> : There are reasons for having to consider non uniform linear arrays. These are: (a) Redundant sensors are removed and employed to increase the array aperture, (b) Certain sensors in a long ULA may fail as a result of factors beyond our control, and (c) Array spacing is intentionally made nonuniform in order to derive certain benefits, for example, there is no aliasing effect if the periodicity of a ULA is destroyed. We shall show how by means of nonuniform array the above objectives may be achieved. <u>Redundant sensors</u>: Let us first consider the case of redundant sensors. We shall rewrite (2.22) as

$$H((\tau - \frac{d}{c}\sin\varphi_0)\omega) = \frac{1}{M}\mathbf{b}^H\phi(r) \qquad (2.23)$$

where $\mathbf{b} = col\{\alpha_0(\theta_0)\alpha_0^*(\theta), \alpha_1(\theta_0)\alpha_1^*(\theta),\dots\alpha_{P-1}(\theta_0)\alpha_{P-1}^*(\theta)\}$ and $r = \omega\frac{d}{c}$ $\times(\sin\theta - \sin\theta_0)$. The power output of the array is simply proportional to the square of the transfer function

$$power \propto |H|^2 = \frac{1}{M^2}\mathbf{b}^H\phi(r)\phi^H(r)\mathbf{b} \qquad (2.24)$$

Let us expand matrix $\phi(r)\phi^H(r)$

$$\mathbf{H} = \phi(r)\phi^H(r) = \begin{bmatrix} 1 & e^{jr} & e^{j2r} & e^{j3r} & \dots e^{j(p-1)r} \\ e^{-jr} & 1 & e^{jr} & e^{j2r} & \dots e^{j(p-2)r} \\ e^{-j2r} & e^{-jr} & 1 & e^{jr} & \dots e^{j(p-3)r} \\ & & \dots & & \\ e^{-j(p-1)r} & e^{-j(p-2)r} & & \dots & 1 \end{bmatrix} \qquad (2.25)$$

Figure 2.9: A four sensor array spread over a 3 λ aperture array will produce all entries of the matrix on (2.25).

Array length in units of $\frac{\lambda}{2}$	Sensor locations for minimum redundancy
6	0, 2, 5,6
15	0, 1, 2, 3, 7, 11, 15 0, 1, 3, 6, 10, 14, 15 0, 1, 4, 8, 13, 14, 15 0, 2, 4, 5, 8, 14, 15
31	0, 1, 2, 3, 4, 5, 12, 18, 25, 31 0, 1, 2, 3, 6, 11, 18, 25, 27, 31 0, 1, 2, 3, 8, 14, 18, 22, 27, 31 0, 1, 2, 3, 12, 18, 22, 26, 29, 31

Table 2.1: Sensor locations for minimum redundancy. For array length of six there is only one arrangement but for array lengths 15 and 31 there are 77 and 888 arrangements, respectively.

It may be observed that **H** is a toeplitz matrix, that is, along any diagonal the entries are repeated even though they refer to different sensors in the array. For example, consider the second upper diagonal where the entries refer to a pair of sensors whose indices are m, and m-2, where m=2,3,...,M-2; explicitly, the pairs of sensors involved in creating the terms on this diagonal are (2,0), (3,1), (4,2), etc. Thus, insofar as the second diagonal is concerned there are several redundant pairs of sensors. This redundancy can be removed by selectively removing sensors [8]. For example, consider the seven sensors uniform array shown in fig. 2.9. All entries in the **H** matrix (2.25) can be obtained from just four sensors shown by filled circles; for example, the first diagonal may be obtained from sensors at position 5 and 6, the second diagonal from sensors at

position 0 and 2, the third diagonal from sensors 2 and 5, the fourth diagonal from sensors 2 and 6, the fifth diagonal from sensors 0 and 5 and finally the last diagonal from sensors 0 and 6. Thus we are able to generate a complete 7x7 matrix just from a four sensor nonuniformly spaced array by removing all redundant sensors. In general, it is possible to arrive at a distribution of the minimum number of sensors required to fill all entries in the matrix in (2.25) [9]. For a given array length the minimum number of sensors required to generate the **H** matrix is found through an exhaustive search. As the array size increases there exists more than one distribution of a zero redundancy array, that is, an array having sensors just enough to generate all diagonals of the **H** matrix. For example, for an array length of 15 (16 sensors spaced at $\frac{\lambda}{2}$) the minimum redundant array has seven sensors and there are 77 different distributions. For array length of 31 (32 sensors spaced at $\frac{\lambda}{2}$) the minimum redundant array has ten sensors and there are 888 different distributions. A few sample distributions are shown in table 2.1

The theoretical response of a minimum redundancy array is identical with that of a normal array (nonredundant array) but in the presence of background noise the response of a redundant array is considerably inferior. To demonstrate this we have considered an array of length 15 but having only seven sensors in the redundant array. Broadband array transfer was computed with background noise variance equal to one. First, the **H** matrix was computed. To simulate a uniform linear array, a complex noise of unit variance was added to each element of the matrix. But to simulate a minimum redundancy array, a complex noise of unit variance was added to each diagonal (same noise element to all elements in a given diagonal). The transfer functions are shown in fig. 2.10. The transfer functions show the main lobe and the first grating lobe at correct frequency in both cases; however, the noisy sidelobes are about ten times greater in the case of minimum redundancy array. The upper and lower bounds on the snr (signal to noise ratio) gain were theoretically derived in [10].

Missing Sensors: In a long chain of sensors such as in a towed array there may be a few sensors which have either failed or are malfunctioning. Such sensors are often skipped, leading to loss of fidelity in the array steering response. Here we shall analyse the effect of a missing sensor on the array performance. Let p be the probability that a sensor will be malfunctioning and it is independent of all other sensors. Let $x_l, l = 0, 1, 2, \ldots N-1$ be locations of live sensors. Note that x_l is a multiple of sensor spacing. Let $\Delta x_l = x_{l+1} - x_l$ be the spacing between the $l+1^{st}$ and l^{th} sensors. It may be expressed as $k \cdot d$, where d is a nominal sensor spacing and k -1 is a Bernoulli random variable so that

$$prob\{\Delta x_l = k \cdot d\} = (1-p)p^{k-1} \qquad (2.26)$$
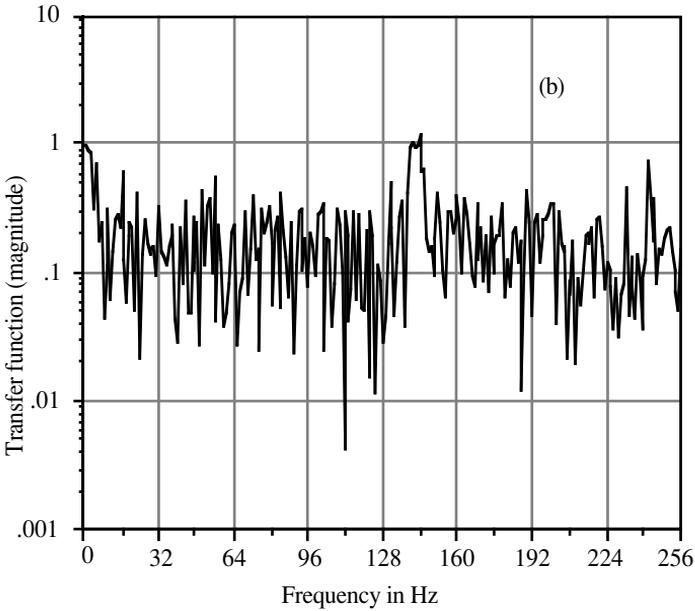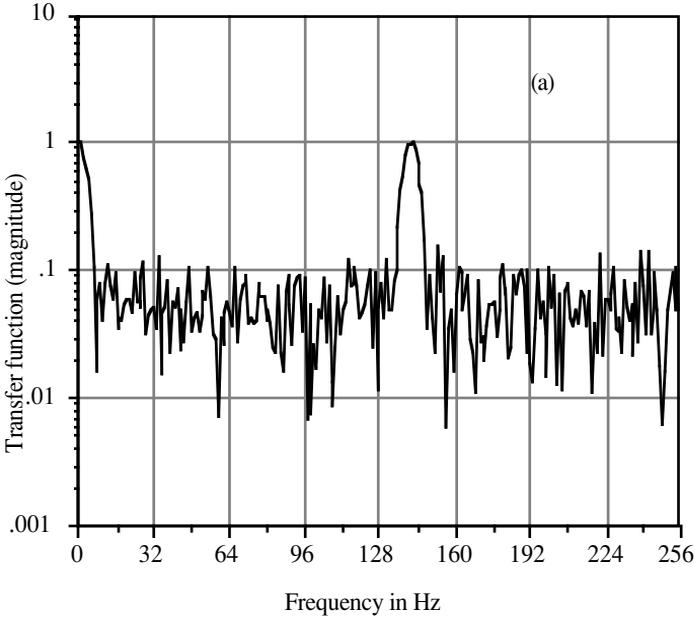
Figure 2.10: (a) Transfer function of a 16 sensor ULA and (b) Transfer function of a 7 sensor minimum redundancy array of same length as ULA in (a). Sensor spacing=15 meters, wave speed=1500 m/s and $\theta = 45^0$.

The output of the nth sensor may be expressed as

$$f_n(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dF(\omega) e^{j(\omega t - \frac{\omega}{c}\sin\theta_0 \sum_{l=0}^{n}\Delta x_l)}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} dF(\omega) e^{j(\omega t)} \prod_{l=0}^{n} e^{-j\frac{\omega}{c}\sin\theta_0 \Delta x_l} \tag{2.27}$$

where θ_0 is the direction of arrival of a wavefront. We further assume that $x_0 = 0$. We consider a delay and sum type of beam steering. Each sensor output is delayed by an amount equal to $\tau_n = \frac{\omega}{c} x_n \sin\theta$ where θ is the angle to which the array is steered

$$g(t) = \frac{1}{N} \sum_{n=0}^{N-1} f_n(t - \tau_n)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} dF(\omega) e^{j(\omega t)} \frac{1}{N} \sum_{n=0}^{N-1} \prod_{l=0}^{n} e^{-j\frac{\omega}{c}(\sin\theta_0 - \sin\theta)\Delta x_l} \tag{2.28}$$

We shall now evaluate the expected value of the delay and sum processor given by (2.28)

$$E\{g(t)\} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dF(\omega) e^{j(\omega t)} \frac{1}{N} \sum_{n=0}^{N-1} \prod_{l=0}^{n} E\left\{ e^{-j\frac{\omega}{c}(\sin\theta_0 - \sin\theta)\Delta x_l} \right\} \tag{2.29}$$

We need to evaluate the expectation of the expression inside curly brackets in (2.29). Using the distribution of the Bernoulli random variable (2.26) in (2.29) we obtain

$$E\left\{ e^{-j\frac{\omega}{c}(\sin\theta_0 - \sin\theta)\Delta x_l} \right\} = \sum_{k=1}^{\infty} (1-p) p^{k-1} e^{-j\frac{\omega}{c}(\sin\theta_0 - \sin\theta)k d}$$

$$= (1-p) e^{-j\frac{\omega}{c} d(\sin\theta_0 - \sin\theta)} \sum_{k=1}^{\infty} p^{k-1} e^{-j\frac{\omega}{c}(\sin\theta_0 - \sin\theta)(k-1)d}$$

$$= \frac{(1-p) e^{-j\frac{\omega}{c} d(\sin\theta_0 - \sin\theta)}}{1 - p e^{-j\frac{\omega}{c} d(\sin\theta_0 - \sin\theta)}} \tag{2.30}$$

Using the above result in (2.29) we obtain

$$E\{g(t)\} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dF(\omega)e^{j(\omega t)} \frac{1}{N} \sum_{n=0}^{N-1} \left[\frac{(1-p)e^{-j\frac{\omega}{c}d(\sin\theta_0 - \sin\theta)}}{1 - pe^{-j\frac{\omega}{c}d(\sin\theta_0 - \sin\theta)}} \right]^n$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} dF(\omega)e^{j(\omega t)} \frac{1}{N} \frac{1 - Q^N}{1 - Q} \tag{2.31}$$

where Q stands for the quantity inside the square brackets of (2.31). Note that for p=0, $Q = e^{-j\frac{\omega}{c}d(\sin\theta_0 - \sin\theta)}$ and (2.31) reduces to a known expression for the response function of a ULA. The response of an array with missing sensors, given by

$$H(\omega\tau) = \frac{1}{N} \frac{1 - Q^N}{1 - Q}$$

where $\tau = \frac{d}{c}(\sin\theta_0 - \sin\theta)$, is shown in for two different values of probability of malfunctioning. It is assumed that the total number of live sensors is the same in both cases, namely, 16. Notice that the magnitude response has fewer sidelobes but the phase characteristics appear to be grossly different from those of a ULA, which is a linear function of $\omega\tau$ (see (2.3)). A nonlinear phase response results in a distortion of the received waveform particularly when it is broadband.

Random Array: A third type of nonuniform array is one where the sensors are spaced at random intervals, in particular, an exponential distribution for which a closed form solution can be derived. Let $x_n, n = 0,1,\ldots N-1$ be the locations of the sensors; all of which, however, lie on a straight line. Let us assume an exponential distribution for the sensor spacing

$$pdf(\Delta x_n) = \beta e^{-\beta\Delta x} \qquad \Delta x \geq 0$$
$$= 0 \qquad \Delta x < 0 \tag{2.32}$$

where pdf stands for probability density function and β is a parameter in the exponential distribution. The output of the nth sensor may be written as in (2.27). The delay and sum type of processing would result in the following array output (from 2.29)
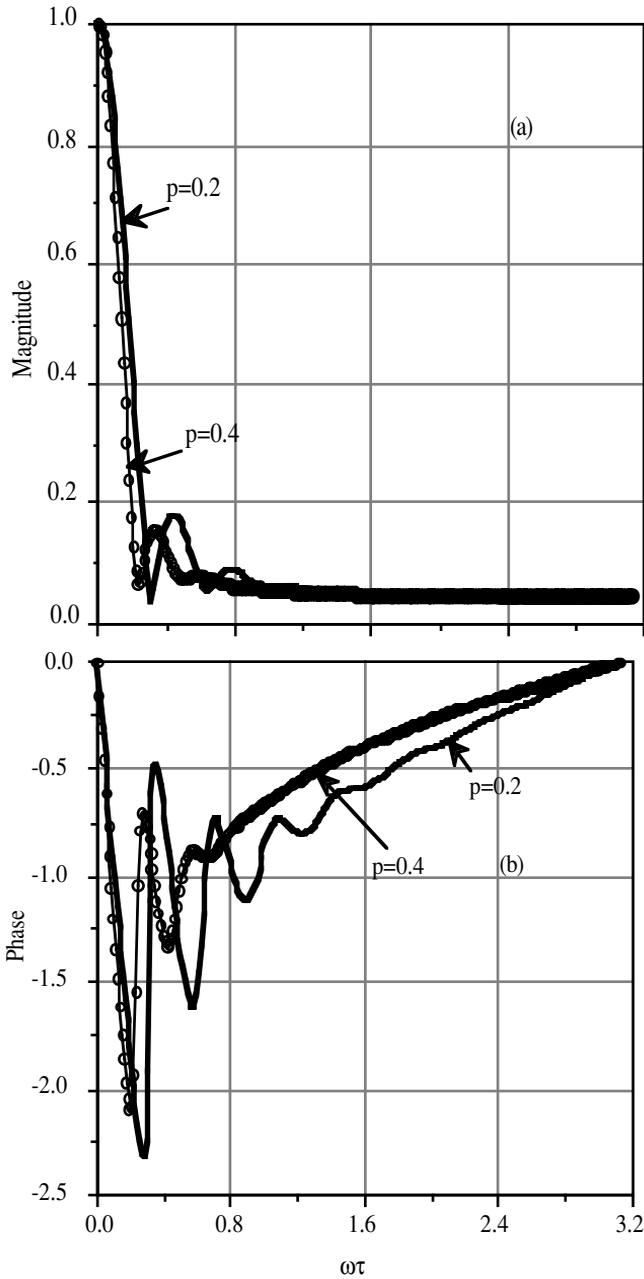
Figure 2.11: (a) Magnitude response and (b) Phase response of an array with malfunctioning sensors. The total number of live sensors is assumed to be the same in both cases, namely, 16.

$$E\{g(t)\} =$$

$$\frac{1}{2\pi}\int\limits_{-\infty}^{+\infty} dF(\omega)e^{j(\omega t)}\frac{1}{N}\sum_{n=0}^{N-1}\prod_{l=0}^{n}\int\limits_{0}^{\infty} pdf(\Delta x_l)e^{-j\frac{\omega}{c}(\sin\theta_0 - \sin\theta)\Delta x_l}d\Delta x_l \quad (2.33a)$$

Using the exponential distribution function (2.32) in (2.33a) we obtain

$$E\{g(t)\} = \frac{1}{2\pi}\int\limits_{-\infty}^{+\infty} dF(\omega)e^{j(\omega t)}\frac{1}{N}\frac{1 - \left[\dfrac{\beta}{j\dfrac{\omega}{c}(\sin\varphi_0 - \sin\varphi) + \beta}\right]^N}{1 - \dfrac{\beta}{j\dfrac{\omega}{c}(\sin\varphi_0 - \sin\varphi) + \beta}}$$

$$(2.33b)$$

The array response may be expressed in terms of the product of the wavelength and parameter β,

$$H(\nu,\beta\lambda) = \frac{1}{N}\frac{1 - \left[\dfrac{\beta\lambda}{j2\pi(\sin\theta_0 - \sin\theta) + \beta\lambda}\right]^N}{1 - \dfrac{\beta\lambda}{j2\pi(\sin\theta_0 - \sin\theta) + \beta\lambda}} \quad (2.33c)$$

where $\nu = 2\pi(\sin\theta_0 - \sin\theta)$. We have plotted the array response function for different values of $\beta\lambda$ in fig. 2.12. While the magnitude response is free from sidelobes, the phase response is highly nonlinear in the range, ν=0.0 to 1.6, where the magnitude response is significant. This behavior was also noticed in the case of an array with missing sensors.

2.1.6 <u>Flexible Array</u>: We have so far considered a sensor array which is rigidly fixed to the ground or to a platform. We now consider an array where the sensors are held in position by means of a flexible rope which allows a sensor to move over a circular arc of fixed radius. The sensor spacing, however, remains unchanged. An important consequence of this freedom allowed to the array is to alter the shape of the array when it is being towed or it is under the influence of ocean currents. A commonly used array model is that M sensors are separated by straight line segments of fixed length d [11] (see fig 2.13a). Let (x_m, y_m, z_m) represent coordinates of the mth sensor with respect to m+1st sensor.
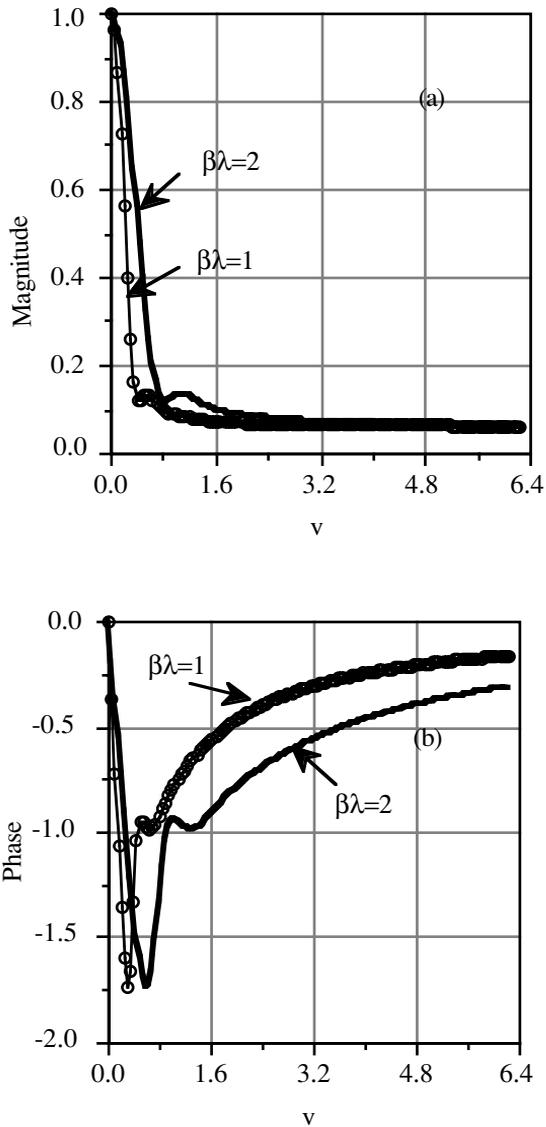
Figure 2.12: Frequency response of a random linear array. (a) magnitude response. (b) phase response. 16 sensors spaced at random intervals having an exponential distribution.

Since the distance between the sensors remains fixed (x_m, y_m, z_m) must satsify a
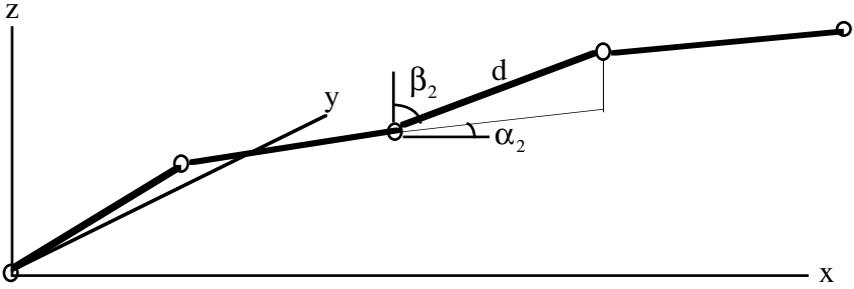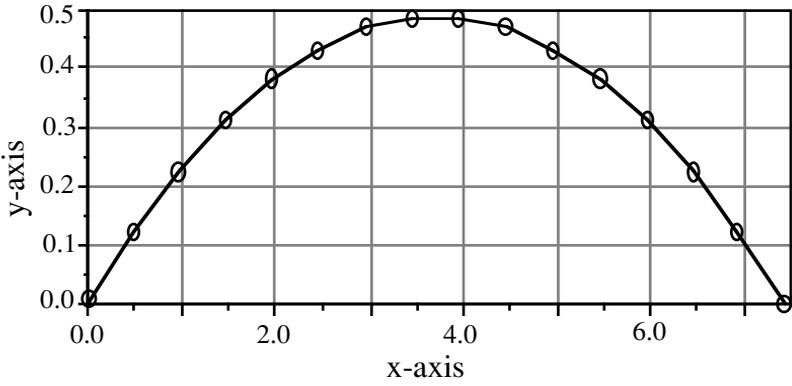
Figure 2.13: (a) Model of a flexible array where the adjacent sensors are held at a fixed distance but the azimuth and elevation of the line segment joining the sensors are random variables. (b) A bow shaped 16 sensor array in the x-y plane. The distance between the adjacent sensors is 0.5λ. All dimensions are in units of wavelength.

relation $\sqrt{x_m^2 + y_m^2 + z_m^2} = d$, alternatively,

$$
\begin{aligned}
x_m &= d\sin\beta_m \sin\alpha_m \\
y_m &= d\sin\beta_m \cos\alpha_m \\
z_m &= d\cos\beta_m
\end{aligned}
\tag{2.34a}
$$

where (α_m, β_m) are azimuth and elevation of a line segment joining m^{th} sensor with m+1st sensor (see fig 2.13a). It is assumed that $x_0 = y_0 = z_0 = 0$.

Let a plane wavefront be incident on the array shown in fig. 2.13. The output of mth sensor is given by

$$f_m(t) = s_0(t)\cos\left(\begin{array}{l}\omega_c t - \sum\limits_{i=0}^{m} x_i \dfrac{\omega_c}{c}\sin\theta_0 \sin\varphi_0 \\[2mm] - \sum\limits_{i=0}^{m} y_i \dfrac{\omega_c}{c}\sin\theta_0 \cos\varphi_0 - \sum\limits_{i=0}^{m} z_i \dfrac{\omega_c}{c}\sin\theta_0\end{array}\right) + \eta_m(t) \quad (2.34b)$$

where (φ_0, θ_0) are respectively azimuth and elevation of the incident wavefront. Using (2.34a) in (2.34b) we get

$$f_m(t) = s_0(t)\cos\left(\omega_c t - \omega_c \dfrac{d}{c}\left[\begin{array}{l}\sum\limits_{i=0}^{m}\sin\beta_i \cos\alpha_i \sin\theta_0 \sin\varphi_0 \\[2mm] + \sum\limits_{i=0}^{m}\sin\beta_i \sin\alpha_i \sin\theta_0 \cos\varphi_0 \\[2mm] + \sum\limits_{i=0}^{m}\cos\beta_i \sin\theta_0\end{array}\right]\right) + \eta_m(t)$$

Transforming into a complex analytical signal (2.17d) the array output may be expressed in a matrix form

$$\mathbf{f}(t) = f_c(t)\,col\left\{\begin{array}{l}1, e^{-j\omega_c \frac{d}{c}[\gamma_1 \sin\theta_0 \sin\varphi_0 + \varepsilon_1 \sin\theta_0 \cos\varphi_0 + \xi_1 \sin\theta_0]} \\[3mm] \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad, \cdots, \\[1mm] e^{-j\omega_c \frac{d}{c}[\gamma_{M-1} \sin\theta_0 \sin\varphi_0 + \varepsilon_{M-1} \sin\theta_0 \cos\varphi_0 + \xi_{M-1} \sin\theta_0]}\end{array}\right\} + \eta(t)$$

where $\gamma_m = \sum\limits_{i=0}^{m}\sin\beta_i \cos\alpha_i$, $\varepsilon_m = \sum\limits_{i=0}^{m}\sin\beta_i \sin\alpha_i$ and $\xi_m = \sum\limits_{i=0}^{m}\cos\beta_i$. These parameters may be expressed in a recursive form

$$\gamma_m = \gamma_{m-1} + \sin\beta_m \cos\alpha_m$$
$$\varepsilon_m = \varepsilon_{m-1} + \sin\beta_m \sin\alpha_m$$
$$\xi_m = \xi_{m-1} + \cos\beta_m$$

When the array is perfectly linear, $\gamma_0 = \gamma_1 =, \ldots, \gamma_{M-1} = \sin\beta_0 \cos\alpha_0$, $\varepsilon_0 = \varepsilon_1 =, \ldots, \varepsilon_{M-1} = \sin\beta_0 \sin\alpha_0$ and $\xi_0 = \sqrt{1 - \gamma_0^2 - \varepsilon_0^2}$.

We like to demonstrate the effect of array deformation on its response function. Consider a 16 sensor ULA which is deformed into a bow shaped curve in an x-y plane as shown in fig. 2.12b. The maximum displacement along the y-axis is half wavelength. The array responses are shown in fig. 2.14.
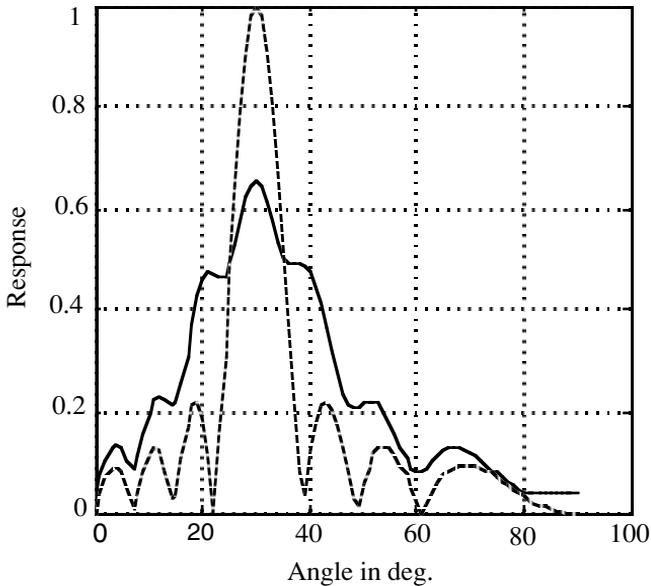
Figure 2.14: Response of a bow shaped array (solid curve) shown in fig. 2.13b and the dashed curve is for an undeformed array.

Notice the extremely broad main lobe which is fortunately located at the right position. If we further increase the deformation, for example, the bow height is increased to 2.3λ the array response is found to be totally distorted. Even the main lobe is found to be wrongly placed.

§2.2 **Planar Array:**
A planar array has its sensors distributed over a plane. When a wavefront has two unknown parameters, azimuth and elevation, we need a planar array for estimation of a pair of parameters. Since a plane has two dimensions, there are many possible array geometries; some of these are illustrated in fig. 2.15. The natural extension of a ULA in two dimensions is a square or rectangular array where the sensors are placed on a square or rectangular grid. Other geometries are essentially sparse versions of the square or the rectangular array. We shall first study the rectangular array and then look into the sparse arrays, in particular, the circular array which has found many applications.

2.2.1 Uniform Planar Array (UPA)**:** Sensors are placed on a rectangular grid where the nodes are spaced d_1 along x-axis and d_2 along y-axis (when $d_1=d_2=d$ we get a square grid). Let a plane wavefront be incident at azimuth angle φ and elevation θ (see fig. 2.16). A plane wavefront, given by

Uniform Square Array

Cross Array

Triangular Array
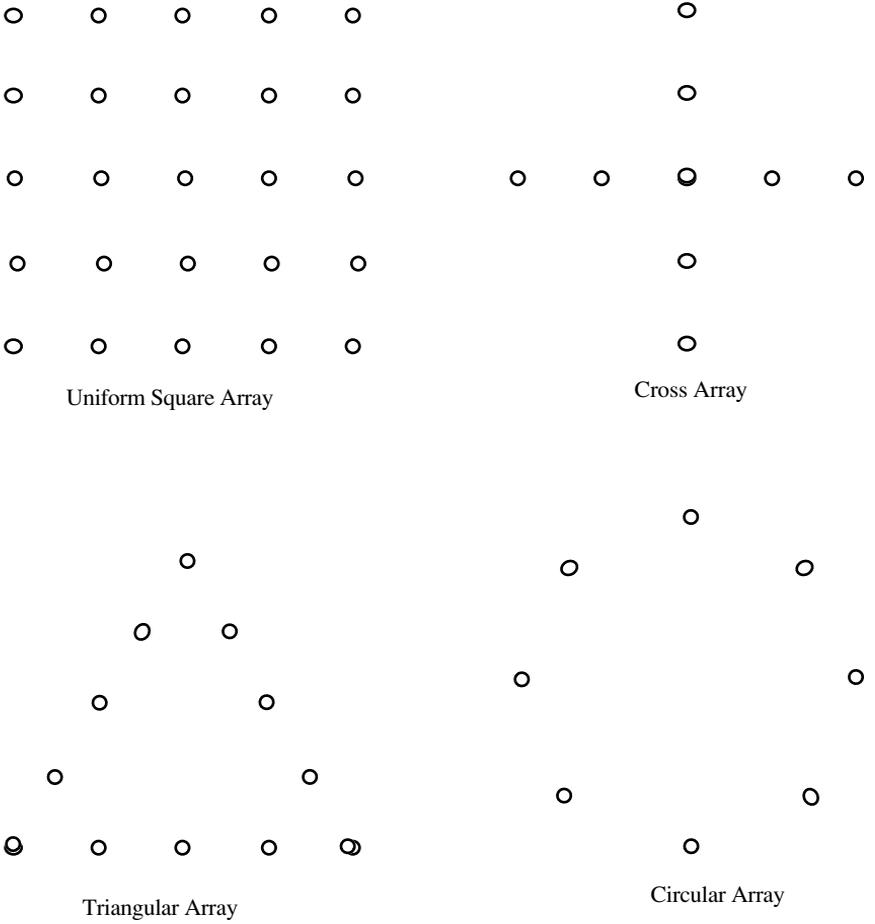
Circular Array

Figure 2.15: Planar array geometries are shown above. The square or rectangular array is a natural extension of ULA and other geometries are sparse versions of square or rectangular array.

$$f(t, x, y) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{j(\omega t - ux - vy)} d\omega \qquad (2.35)$$

where $u = \dfrac{\omega}{c} \sin\theta \cos\varphi$ and $v = \dfrac{\omega}{c} \sin\theta \sin\varphi$, is incident on a UPA. The output of the $(m,n)^{\text{th}}$ sensor is
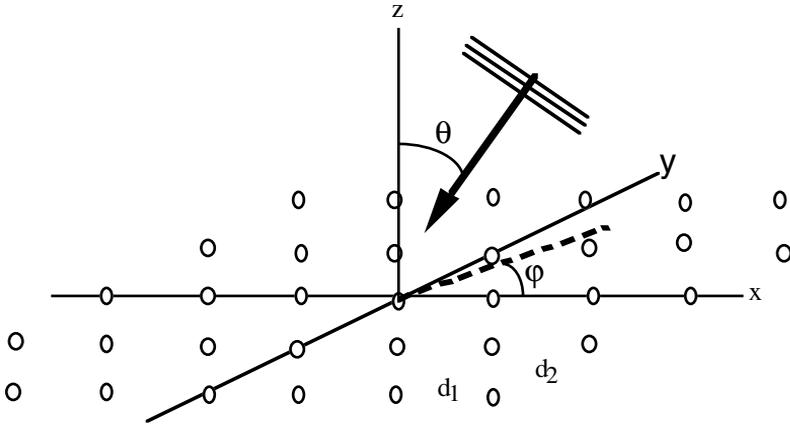
Figure 2.16: A plane wavefront incident on a UPA at azimuth φ and elevation θ.

$$f_{m_1 m_2}(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{j(\omega t - u m_1 d_1 - v m_2 d_2)} d\omega \qquad (2.36)$$

All sensor outputs are summed in phase, yielding

$$
\begin{aligned}
g(t) &= \frac{1}{M_1 M_2} \sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} a_{m_1 m_2} f_{m_1 m_2}(t) \\
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{j\omega t} \frac{1}{M_1 M_2} \sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} a_{m_1 m_2} e^{-j(u m_1 d_1 + v m_2 d_2)} d\omega \\
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{j\omega t} H(u d_1, v d_2) d\omega
\end{aligned}
\qquad (2.37)
$$

where

$$H(u d_1, v d_2) = \frac{1}{M_1 M_2} \sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} a_{m_1 m_2} e^{-j(u m_1 d_1 + v m_2 d_2)} \qquad (2.38a)$$

which, for constant weighting coefficients, becomes

$$H(ud_1, vd_2) = \frac{\sin(\frac{M_1 ud_1}{2})}{M_1 \sin\frac{ud_1}{2}} \frac{\sin(\frac{M_2 vd_2}{2})}{M_2 \sin\frac{vd_2}{2}} e^{-j\frac{[(M_1-1)u+(M_2-1)v]}{2}} \qquad (2.38b)$$

The frequency response function of a UPA given by (2.38b) can be written as a product of two frequency response functions of two ULAs, one in the x direction and the other in the y direction. This is also true in the weighted sum case provided the weighting coefficients can be expressed as a product of two coefficient sets, that is, $a_{mn} = \alpha_m \beta_n$ where $(\alpha_{m_1}, \; m_1 = 0,1,\ldots M_1 - 1)$ and $(\beta_{m_2}, m_2 = 0,1,\ldots M_2 - 1)$. This is the motivation behind the use of a cross array (fig. 2.15) in place of a UPA.

A UPA may be steered both in azimuth and elevation by means of appropriate delays introduced before summation. The delay to be introduced in the (m,n)$^{\text{th}}$ sensor is

$$\tau_{m_1, m_2} = \frac{m_1 d_1}{c} \sin\theta \cos\varphi + \frac{m_2 d_2}{c} \sin\theta \sin\varphi$$

where φ and θ are respectively azimuth and elevation to which the array is required to be steered. In place of (2.37) we have

$$
\begin{aligned}
g(t) &= \sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} f_{m_1 m_2}(t + \tau_{m_1, m_2}) \\
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{j\omega t} \sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} e^{-j[(u_0-u)m_1 d_1 + (v_0-v)m_2 d_2]} d\omega \qquad (2.39) \\
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{j\omega t} H((u_0-u)d_1, (v_0-v)d_2) d\omega
\end{aligned}
$$

where

$$
\begin{aligned}
&H((u_0-u)d_1, (v_0-v)d_2) \\
&= \frac{\sin(\frac{M_1(u_0-u)d_1}{2})}{M_1 \sin\frac{(u_0-u)d_1}{2}} \frac{\sin(\frac{M_2(v_0-v)d_2}{2})}{M_2 \sin\frac{(v_0-v)d_2}{2}} e^{-j\frac{[(M_1-1)(u_0-u)+(M_2-1)(v_0-v)]}{2}}
\end{aligned}
$$

where

$$u_0 = \frac{\omega}{c} \sin\theta_0 \cos\varphi_0,$$

$$v_0 = \frac{\omega}{c} \sin\theta_0 \sin\varphi_0$$

and (φ_0, θ_0) are respectively azimuth and elevation angles.

Let the steering angles φ and θ be chosen such that $\frac{d_1}{\lambda}\sin\theta\cos\varphi = \frac{k}{M_1}$ and $\frac{d_2}{\lambda}\sin\theta\sin\varphi = \frac{l}{M_2}$ where k and l are integers, $0 \le k \le M_1$ and $0 \le l \le M_2$. Equation (2.39) may be expressed as a 2D spatial discrete Fourier transform (DFT) of the wavefield incident on the array.

$$g(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} F(\omega)e^{j\omega t} \sum_{m_1=1}^{M_1}\sum_{m_2=1}^{M_2} e^{-j(u_0 m_1 d_1 + v_0 m_2 d_2)} e^{j(\frac{2\pi k m_1}{M_1} + \frac{2\pi l m_2}{M_2})} d\omega$$

(2.40a)

$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} F(\omega)e^{j\omega t} DFT\{e^{-j(u_0 m_1 d_1 + v_0 m_2 d_2)}\} d\omega$$

and

$$H((u_0 - u)d_1, (v_0 - v)d_2) = DFT\{e^{-j(u_0 m_1 d_1 + v_0 m_2 d_2)}\} \quad (2.40b)$$

This result is an extension of (2.7) for a ULA to a UPA. The spatial frequencies u and v, which appear in (2.40b), are linearly related to the frequency numbers, k and l, in DFT

$$u = \frac{2\pi k}{M_1 d_1}, \quad v = \frac{2\pi l}{M_1 d_1}$$

Given the frequency numbers, we can get the azimuth and elevation angles. For example, assuming $d_1 = d_2 = \frac{\lambda}{2}$ and $M_1 = M_2 = M$, $\varphi = \sin^{-1}\frac{l}{\sqrt{k^2 + l^2}}$ and $\theta = \sin^{-1}\frac{2\sqrt{k^2 + l^2}}{M}$. Note that $0 \le k, l \le \frac{M}{2}$ and $0 \le \sqrt{k^2 + l^2} \le \frac{M}{2}$. Thus, the acceptable domain for k and l is as shown in . For those values of k and l lying outside this domain we get nonpropagating waves (see chapter 1, page 11).

<u>Random Array</u>: An array with its sensors at random locations, modelled as an independent, identically distributed random variable, is known as a totally random array [12]. The root mean square (rms) error in the estimation of the
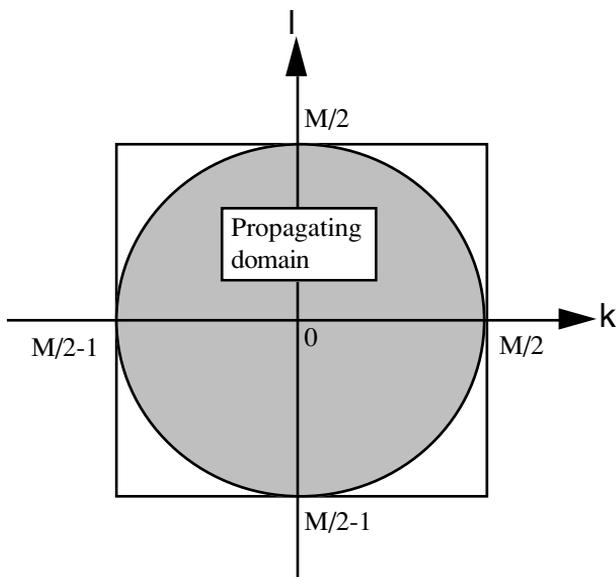
Figure 2.17: Acceptable domain for frequency numbers, k and l, is shown above. For those values of k and l lying outside this domain we get non propagating waves.

azimuth angle (measured with respect to the x-axis) is given in [12] and reproduced here without derivation

$$rms\ error = \sqrt{E\left\{(\hat{\varphi}_0 - \varphi_0)^2\right\}}$$

$$\approx \frac{\lambda_0}{2\pi} \frac{1}{\sqrt{2snrM}} \frac{1}{\sigma}$$

where σ is the standard deviation of the random variable representing the location of the sensors. The above approximate result is valid for large M.

Interestingly, the rms error decreases if the random array is spread out over a large area. The practical interest in a random array arises in the context of sonobuoy arrays used in submarine detection.

2.2.2 Uniform Circular Array (UCA): The sensors may be placed on a plane in a polar grid. For a fixed radial distance we have a circle on which the sensors are placed, forming a circular array. Consider a circular array of radius a with M sensors, symmetrically placed on the circumference (see fig. 2.18). Let a plane wavefront be incident on the array at angles φ and θ. The output of the mth sensor is given by
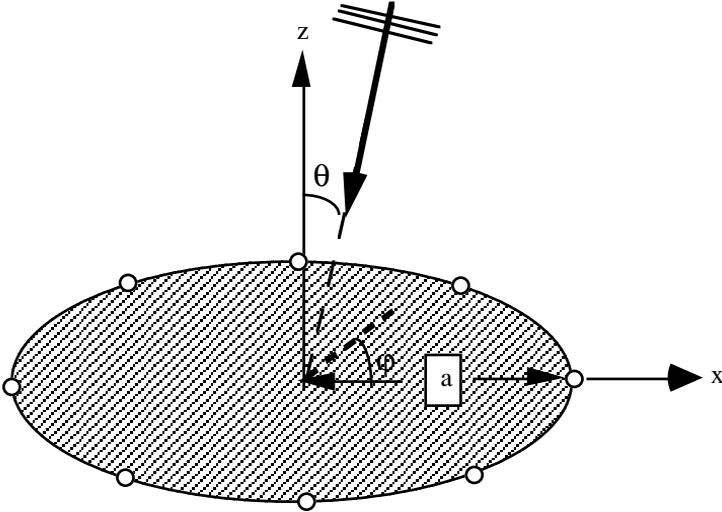
Figure 2.18: Sensors are uniformly spaced on the circumference of a circle of radius a. A plane wave is incident at an azimuth angle φ and an elevation angle θ.

$$f_m(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j[\omega t - \frac{\omega a}{c}(\cos\varphi\sin\theta\cos\frac{2\pi m}{M} + \sin\varphi\sin\theta\sin\frac{2\pi m}{M})]} \, d\omega$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j[\omega t - \frac{\omega a}{c}(\sin\theta\cos(\frac{2\pi m}{M} - \varphi))]} \, d\omega$$

(2.41)

Note that time is measured with respect to the time of arrival of the wavefront at the center of the array. First, we evaluate the frequency response function. The sum of all outputs in the frequency domain is given by

$$g(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega t} \frac{1}{M} \sum_{m=0}^{M-1} e^{-j[\frac{\omega a}{c}(\sin\theta\cos(\frac{2\pi m}{M} - \varphi))]} \, d\omega$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) H(\omega, \varphi, \theta) e^{j\omega t} \, d\omega$$

(2.42)

where the frequency response function $H(\omega, \varphi, \theta)$ is given by

$$H(\omega,\varphi,\theta) = \frac{1}{M}\sum_{m=0}^{M-1} e^{-j[\frac{\omega a}{c}(\sin\theta\cos(\frac{2\pi m}{M}-\varphi))]} \qquad (2.43a)$$

For large M (for example, M> 48 when a=6λ and M>32 when a=4λ) the summation in (2.43a) may be replaced by an integral and the result is

$$H(\omega,\varphi,\theta) \approx J_0(\frac{\omega a}{c}\sin\theta) \qquad (2.43b)$$

We shall call such a UCA a fully populated array. The most interesting property of a circular array is that the frequency response function is independent of φ. The property arises from (2.43b). Taking the distance to the first zero as the effective half width of the main lobe, the angular width will be equal to $\Delta\theta = \sin^{-1}(2.45\frac{c}{\omega a})$. The height of the first (largest) sidelobe is 0.4025 at $\theta = \sin^{-1}(0.8\frac{c}{\omega a})$.

   A circular array may be steered to any desired direction just like a ULA or a UPA. A delay τ_m is introduced at each sensor output before summation, where

$$\tau_m = [\frac{a}{c}(\cos\varphi\sin\theta\cos\frac{2\pi m}{M} + \sin\varphi\sin\theta\sin\frac{2\pi m}{M})]$$

and φ and θ respectively are the desired azimuth and elevation angles. The delayed outputs of all sensors are summed, for the time being without any weighting.

$$g(t) = \frac{1}{M}\sum_{m=0}^{M-1} f_m(t+\tau_m)$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{j\omega t}\frac{1}{M}\sum_{m=0}^{M-1} e^{-j[\frac{\omega a}{c}(\sin\theta_0\cos(\frac{2\pi m}{M}-\varphi_0)-\sin\theta\cos(\frac{2\pi m}{M}-\varphi))]} d\omega$$

$$(2.44)$$

where φ_0 and θ_0 respectively are the unknown azimuth and elevation angles of the incident wavefront. Let

$$H(\frac{\omega a}{c},\theta_0,\varphi_0,\theta,\varphi) = \frac{1}{M}\sum_{m=0}^{M-1} e^{-j[\frac{\omega a}{c}(\sin\theta_0\cos(\frac{2\pi m}{M}-\varphi_0)-\sin\theta\cos(\frac{2\pi m}{M}-\varphi))]}$$

The output power of the array, steered to any chosen direction φ and θ, is given by

$$\text{output power} = \left| F(\omega)H(\frac{\omega a}{c}, \theta_0, \varphi_0, \theta, \varphi) \right|^2$$

Earlier we had noted that steering of an array is equivalent to the spatial Fourier transform of the array output. This result holds in a slightly different form for a circular array. We will demonstrate how the spatial Fourier transform can be used for estimation of azimuth [13] and elevation angles. Consider the spatial discrete Fourier transform of the circular array output.

$$
\begin{aligned}
g_k(t) &= \frac{1}{M} \sum_{m=0}^{M-1} f_m(t) e^{-j\frac{2\pi m}{M}k} \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega t} \frac{1}{M} \sum_{m=0}^{M-1} e^{-j[\frac{\omega a}{c}(\sin\theta\cos(\frac{2\pi m}{M} - \varphi))]} d\omega \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) H(\omega, \varphi, \theta) e^{j\omega t} d\omega
\end{aligned}
\tag{2.45}
$$

Taking the temporal Fourier transform of (2.45) we obtain an important result,

$$G_k(\omega) \approx F(\omega) J_k(\frac{\omega a}{c} \sin\theta_0) e^{jk\frac{\pi}{2}} e^{-jk\varphi_0} \tag{2.46}$$

which is valid for $k < k_{max} \approx \frac{\omega a}{c}$ [15] and for sensor spacing approximately equal to $\frac{\lambda}{2}$ [14]. Consider the following quantity:

$$\frac{G_{k+1}(\omega)}{G_k(\omega)} = je^{-j\varphi_0} \frac{J_{k+1}(\frac{\omega a}{c}\sin\theta_0)}{J_k(\frac{\omega a}{c}\sin\theta_0)} \tag{2.47}$$

Referring to the recurrence relation of Bessel functions [15],

$$J_{k+1}(x) = \frac{2k}{x} J_k(x) - J_{k-1}(x),$$

we can write

$$\frac{J_{k+1}(x)}{J_k(x)} = \frac{2k}{x} - \frac{J_{k-1}(x)}{J_k(x)}$$

which we use in (2.47) and derive a basic result for the estimation of φ_0 and θ_0

$$-je^{j\varphi_0}\frac{G_{k+1}(\omega)}{G_k(\omega)} = \frac{2k}{\dfrac{\omega a}{c}\sin\theta_0} - je^{-j\varphi_0}\frac{G_{k-1}(\omega)}{G_k(\omega)} \qquad (2.48)$$

Equation (2.48) may be solved for φ_0 and θ_0 As an example, we consider a 16 sensor circular array of 3λ radius and a source in a far field emitting a bandlimited random signal. The center frequency is 100Hz and the bandwidth is 10Hz. The azimuth and the elevation angle of the source are respectively 10° (0.1745 rad) and 45° (0.7854 rad). The sampling rate was 500 samples/sec. The estimates were averaged over all frequency bins lying within the bandwidth. The results are shown in fig. 2.19. Notice that the standard deviation of the estimates decreases considerably when a reference sensor is used at the center. The decrease is more pronounced at a very low snr, e. g., at 0 dB the decrease is by a factor of three or more. An analysis of errors has shown that the standard deviation is dominated by a few outliers which are caused by random noise in the array output. Unless these outliers are eliminated the mean and the standard deviation of the estimate gets severely affected. To overcome this problem median in place of mean may be considered. It was observed through computer simulation that the median is a better estimate of the azimuth than the mean.

When there is more than one source (say, P sources), equation (2.46) takes a form

$$G_k(\omega) = \sum_{i=0}^{p-1} F_i(\omega)J_k(\frac{\omega a}{c}\sin\theta_i)e^{jk\frac{\pi}{2}}e^{-jk\varphi_i} + \eta_k(\omega) \qquad (2.49a)$$

We shall rewrite (2.49a) in a matrix form. For this purpose we define the following vectors amd matrices:

$$\mathbf{G}_r(\omega) = col\{G_0(\omega), G_1(\omega), \dots G_{r-1}(\omega)\}$$

$$\mathbf{D}_i(\omega) = diag\left\{J_0(\frac{\omega a}{c}\sin\theta_i), J_1(\frac{\omega a}{c}\sin\theta_i), \quad \dots \quad J_{r-1}(\frac{\omega a}{c}\sin\theta_i),\right\}$$
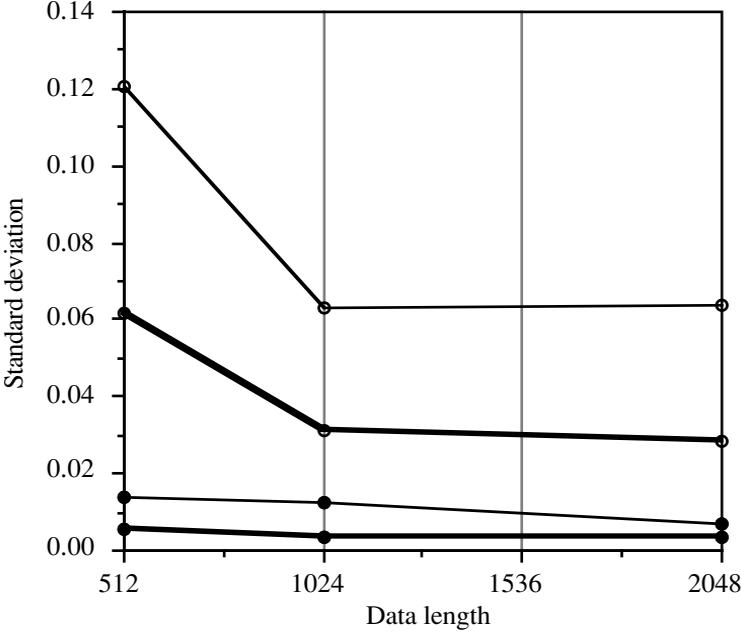
Figure 2.19: Standard deviation of azimuth and elevation estimates as a function of data length. Thick line: with a sensor at the center, thin line: without a sensor at the center, filled circle: azimuth and empty circle: elevation. snr=10 dB.

$$\mathbf{Z}_i = col\left\{1, e^{-j(\varphi_i - \frac{\pi}{2})}, e^{-j2(\varphi_i - \frac{\pi}{2})}, \ldots e^{-j(r-1)(\varphi_i - \frac{\pi}{2})}\right\}$$

$$\eta_r(\omega) = col\left\{\eta_0(\omega), \eta_1(\omega), \ldots \eta_{r-1}(\omega)\right\}$$

where r is an integer $(P \leq r \leq M)$.

$$\mathbf{G}_r(\omega) = \sum_{i=0}^{P-1} \mathbf{D}_i(\omega)\mathbf{Z}_i F_i(\omega) + \eta_r(\omega) \tag{2.49b}$$

Let us assume that all P sources emit stationary uncorrelated stochastic processes. Since the array output will also be a stationary stochastic process, the spectral matrix is given by $\frac{1}{2\pi}S_g(\omega)d\omega = E\left\{\frac{1}{2\pi}d\mathbf{G}_r(\omega)\frac{1}{2\pi}d\mathbf{G}_r^H(\omega)\right\}$ where we have used the generalized Fourier transform in place of the ordinary Fourier transform. From (2.49b) it follows that

$$S_g = \begin{bmatrix} \mathbf{D}_0\mathbf{Z}_0, \mathbf{D}_1\mathbf{Z}_1, \ldots \mathbf{D}_{p-1}\mathbf{Z}_{p-1} \end{bmatrix} \begin{bmatrix} S_0 & & & & \\ & S_1 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & & S_{p-1} \end{bmatrix} \begin{bmatrix} (\mathbf{D}_0\mathbf{Z}_0)^H \\ (\mathbf{D}_1\mathbf{Z}_1)^H \\ \cdot \\ \cdot \\ \cdot \\ (\mathbf{D}_{p-1}\mathbf{Z}_{p-1})^H \end{bmatrix} + \sigma_\eta^2 \mathbf{I}$$

(2.50)

(rxr)　　　　　　　(rxp)　　　　　　(pxp)　　　　　(pxr)

Exploiting the above signal structure a subspace approach has been developed in [14, 16] for the estimation of azimuth and elevation.

<u>Steering Vector</u>: For circular array we define a steering vector as

$$\mathbf{a}(\varphi, \theta)$$
$$= col\left[e^{-j[\frac{\omega a}{c}(\sin\theta\cos(\varphi))]}, e^{-j[\frac{\omega a}{c}(\sin\theta\cos(\frac{2\pi}{M}-\varphi))]}, \ldots e^{-j[\frac{\omega a}{c}(\sin\theta\cos(\frac{2\pi(M-1)}{M}-\varphi))]} \right]$$

(2.51)

Each term in the steering vector can be expanded in a series form [15]

$$e^{-j[\frac{\omega a}{c}(\sin\theta\cos(\frac{2\pi m}{M}-\varphi))]} = \sum_{k=-\infty}^{\infty}(-j)^k J_k\left(\frac{\omega a}{c}\sin\theta\right)e^{jk(\frac{2\pi m}{M}-\varphi)}$$

Define a matrix

$$\{\mathbf{W}\}_{km} = e^{j\frac{2\pi km}{M}} \qquad \begin{Bmatrix} m = 0,1,\ldots M-1 \\ k = 0,\pm1,\ldots \infty \end{Bmatrix}$$

and a vector

$$\{\mathbf{V}(\varphi,\theta)\}_k = J_k\left(\frac{\omega a}{c}\sin\theta\right)e^{-j(\frac{\pi}{2}+\varphi)k}, \quad k = 0,\pm1,\ldots\infty$$

In terms of matrix $\mathbf{W}$ and vector $\mathbf{V}$ we can express (2.51) as

$$\mathbf{a}(\varphi,\theta) = \mathbf{W} \ \mathbf{V}(\varphi,\theta)$$
$$M \times 1 \qquad M \times \infty \quad \infty \times 1$$

Observe that matrix $\mathbf{W}$ is independent of the azimuth and elevation angles, which are confined only to the vector $\mathbf{V}(\varphi,\theta)$. The size of $\mathbf{V}(\varphi,\theta)$ depends on the argument of the Bessel function which may be approximated, for large order, as [15]

$$J_k(\frac{\omega a}{c}\sin\theta) \approx \frac{1}{\sqrt{2\pi k}}(\frac{e^{\frac{\omega a}{c}\sin\theta}}{2k})^k$$

Hence, the size of vector $\mathbf{V}(\varphi,\theta)$ will be of the order of $e\frac{\omega a}{c}\sin\theta$. For example, for $a = 8\lambda$ and $\theta = \frac{\pi}{2}$ the size of the vector is about 140. The steering vector for a circular array possesses some interesting properties, namely,

(a) $\mathbf{a}(\varphi,\theta) \neq \mathbf{a}(-\varphi,\theta)$,

(b) $\mathbf{a}^*(\varphi,\theta) = \mathbf{a}(\pi - \varphi,\theta)$ and

(c) $\mathbf{a}(\varphi,\theta)$ periodic in φ with period 2π, and independent of sensor spacing.

Property (a) implies a wavefront coming from the north can be distinguished from one coming from the south (north-south ambiguity). Property (b) implies that a complex signal coming from the east can be distinguished from the one coming from the west (east-west ambiguity; see fig. 2.20 for an illustration) and property (c) implies that for any sensor spacing there is no grating lobe in the range of $\pm\pi$. A circular array differs from a linear array in respect of properties (a- c).

Boundary Array: An array of sensors for localization of an object in the near field region may take a form of a boundary array where the sensors are placed all around the object as its approximate location is known before hand. A circular array enclosing a source is an example of boundary array. With three coplanar sensors and accurate time delay measurements it is possible to localize a point target in the plane of array. There is a vast literature on time delay estimation but, on account of space limitation, the topic of time delay estimation will not be covered in this book. It is, on the other hand, possible for a boundary array to determine the time delays from the phase measurements. The basic idea is to consider a pair of sensors and estimate the phase difference at a fixed frequency ω. The source will fall on one of the phase trajectories drawn with a phase difference $\omega\Delta\tau = 2\pi s + \phi$ where $\Delta\tau$ is the unknown time delay, s is an integer and ϕ is the actual observed phase difference. To estimate the time delay $\Delta\tau$,
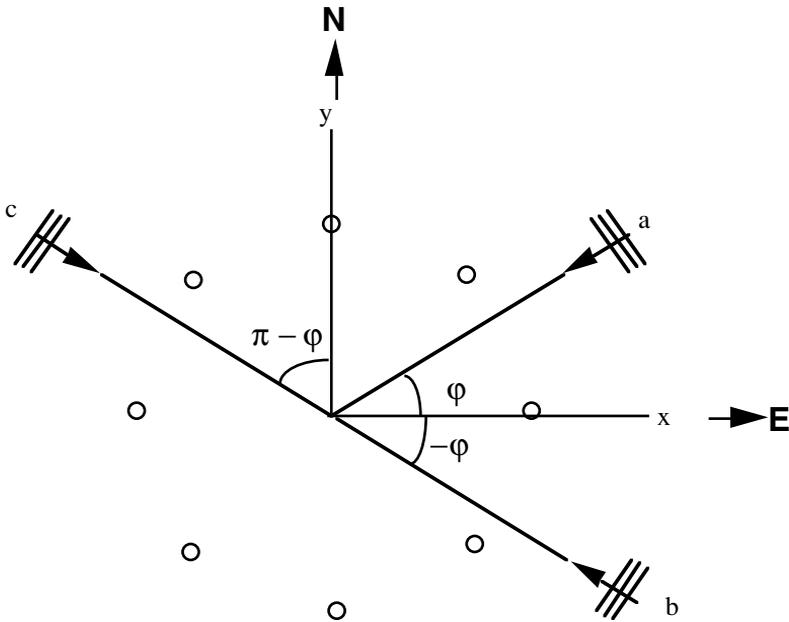
Figure 2.20: Circular array does not suffer from north-south ambiguity, that is, wavefronts a and b can be distinguished. There is no east-west ambiguity for complex signals, that is, wavefronts a and c can be distinguished.

we need to know the integer constant s in the range $\pm s_0$ where $s_0 = Int\left[\dfrac{\omega \Delta \tau}{2\pi}\right]$ and $Int[x]$ stands for the largest integer less than x. For a given ϕ, $\Delta \tau$ will assume a set of $2s_0 + 1$ values and for each value of s there corresponds a locus of points called the phase trajectory [17]. For every pair of adjacent sensors we can draw a suite of trajectories as shown in fig. 2.21a. The unknown source must lie on any one of the trajectories. Next, consider another sensor pair and draw another suite of trajectories. Any one of the points of intersection is a possible location of the unknown source (see fig. 2.21b). Since there are M pairs of sensors there will be M suites of trajectories. The true position of the unknown source is then given by the intersection of all M trajectories, one from each suite. At true source location all phase estimates obtained from different pairs of sensors must match with the theoretically evaluated phases. Let $\hat{\phi}_1^k$, $k = 0, 1, \ldots, M-1$ be the estimated phases from M pairs of adjacent sensors and ϕ_1^k, $k = 0, 1, \ldots, M-1$ be the theoretically computed phases. Define an error vector

(a)

(b)
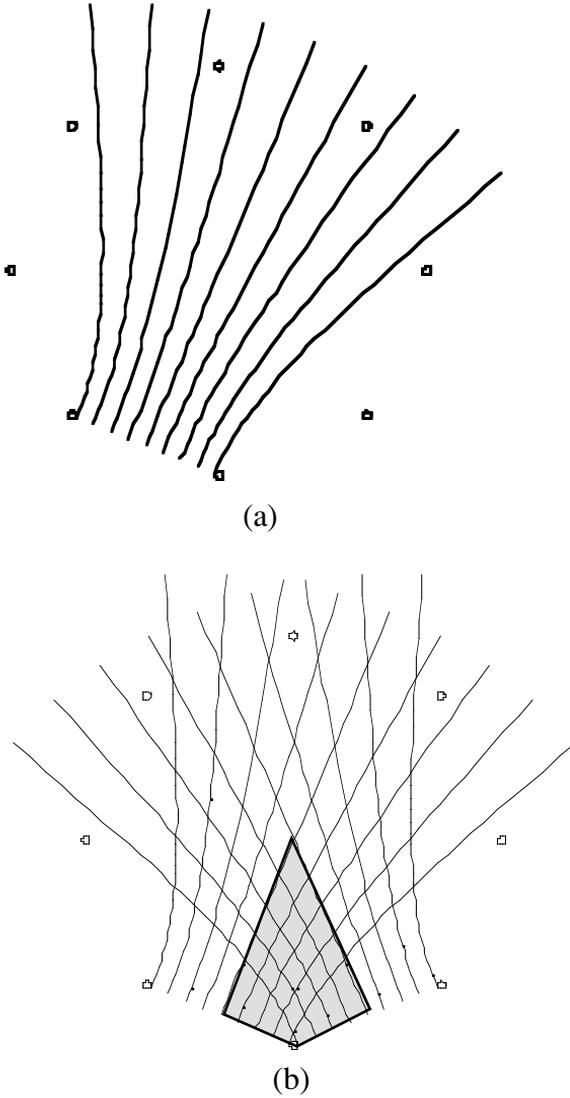
Figure 2.21: (a) For every pair of adjacent sensors we draw a suite of equiphase trajectories. (b) Intersection of two sets of equiphase trajectories. The unknown source must lie at one of the intersections. For every adjacent pair of sensors the search is carried out within the dotted quadrilateral.

$$\varepsilon = col\left\{(e^{j\phi_1^0} - e^{j\hat{\phi}_1^0}), (e^{j\phi_1^1} - e^{j\hat{\phi}_1^1}), ..., (e^{j\phi_1^{M-1}} - e^{j\hat{\phi}_1^{M-1}})\right\} \qquad (2.52)$$
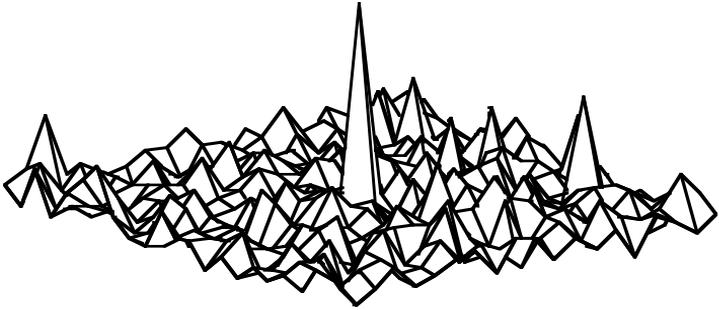
Figure 2.22 A plot of inverse error power. The source is at range 50 meters and azimuth -60$^{\text{o}}$. A circular array of radius 100 meters and having eight equispaced sensors is assumed.

and error power$= \varepsilon^{\text{H}} \varepsilon$. The error power will be zero at the true location of the source. This property may be utilized to spot the true location from all available intersections of any two suites of trajectories. Evaluate the error power at each and every intersection. That intersection which yields zero (or minimum) error power is the most likely location of the unknown source. Finally, to reconstruct the true phase we need to know the integers which may be obtained from the order of the trajectory for every pair of sensors passing through the source location. However, in any practical problem, this step may not be required as the interest is usually in localization and not in the true phase retrieval. The results of a numerical experiment are shown in fig. 2.22. The question of minimum size (number of sensors) of the array, required for unique localization, has not been answered. However, numerical experiments suggest a minimum array size of five sensors placed the circumference of a large circle. It is not necessary that the array be perfectly circular. Any closed curve will do, provided the phase trajectories are drawn for each pair.

2.2.3 Distributed Dipoles: A planar array with sensors randomly distributed in a plane constitutes a random planar array. Here we shall consider a random planar array of dipole sensors. A dipole sensor consists of a pair of *identical* sensors displaced by Δ (a vector). In fig. 2.23 an example of a dipole planar array is shown. Let $\mathbf{r}_i, i = 0,1,\dots M-1$ represent locations of M dipole sensors (the midpoint of the sensor pair). It may be emphasized that the spatial distribution of the dipoles can be quite arbitrary. Define the following data vector in frequency domain:
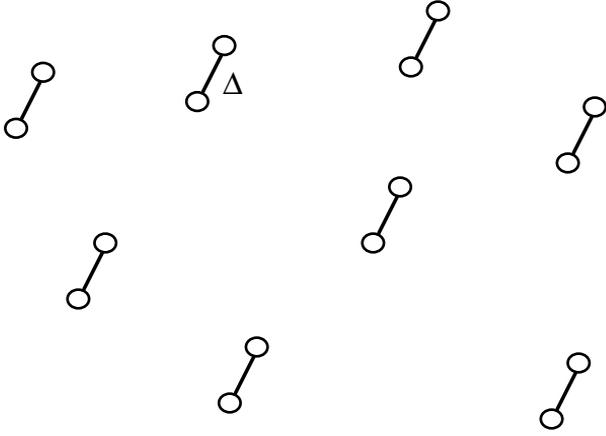
Figure 2.23: Randomly distributed planar dipole sensors. Each dipole consists of two identical sensors displaced by Δ (vector).

$$\tilde{\mathbf{F}} = \begin{bmatrix} \mathbf{F}_1 \\ \mathbf{F}_2 \end{bmatrix} \qquad (2.53)$$

where

$$\mathbf{F}_1 = col\left\{ F_0^+(\omega), F_1^+(\omega) \ldots F_{M-1}^+(\omega) \right\}$$

$$\mathbf{F}_2 = col\left\{ F_0^-(\omega), F_1^-(\omega) \ldots F_{M-1}^-(\omega) \right\}$$

where superscript + refers to the upper sensor and - refers to the lower sensor. Let P plane wavefronts from P sources be incident on the array.

$$\mathbf{F}_1 = \left[\mathbf{a}_0, \mathbf{a}_1, \ldots \mathbf{a}_{p-1} \right] \mathbf{F}(\omega) + d\eta_1(\omega) \qquad (2.54)$$

where

$$\mathbf{a}_i = col\left\{ e^{-j\frac{\Delta}{2}\cdot\delta_i}, e^{-j\frac{\omega}{c}(\mathbf{r}_1 + \frac{\Delta}{2})\cdot\delta_i}, \ldots, e^{-j\frac{\omega}{c}(\mathbf{r}_{M-1} + \frac{\Delta}{2})\cdot\delta_i} \right\}$$

$$= col\left\{ 1, e^{-j\frac{\omega}{c}\mathbf{r}_1\cdot\delta_i}, \ldots, e^{-j\frac{\omega}{c}\mathbf{r}_{M-1}\cdot\delta_i} \right\} e^{-j\frac{\Delta\cdot\delta_i}{2}}$$

where $\delta_i = col\{\sin\theta_i, \cos\theta_i\}$ and $\mathbf{F}(\omega) = [F_0(\omega), F_1(\omega)...F_{P-1}(\omega)]^T$ is a vector whose components are Fourier transforms of the waveforms emitted by P sources. Equation (2.54) may be expressed as follows

$$\mathbf{F}_1 = \mathbf{A}\Gamma_+\mathbf{F}(\omega) + d\eta_1(\omega) \qquad (2.55a)$$

where

$$\Gamma_+ = diag\left\{e^{-j\frac{\Delta\cdot\delta_0}{2}}, e^{-j\frac{\Delta\cdot\delta_1}{2}}, ..., e^{-j\frac{\Delta\cdot\delta_{p-1}}{2}}\right\}$$

and

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{-j\frac{\omega}{c}\mathbf{r}_1\cdot\delta_0} & e^{-j\frac{\omega}{c}\mathbf{r}_1\cdot\delta_1} & & e^{-j\frac{\omega}{c}\mathbf{r}_1\cdot\delta_{p-1}} \\ . & & & \\ . & & & \\ . & & & \\ e^{-j\frac{\omega}{c}\mathbf{r}_{M-1}\cdot\delta_0} & e^{-j\frac{\omega}{c}\mathbf{r}_{M-1}\cdot\delta_1} & \cdots & e^{-j\frac{\omega}{c}\mathbf{r}_{M-1}\cdot\delta_{p-1}} \end{bmatrix}$$

Similarly,

$$\mathbf{F}_2 = \mathbf{A}\Gamma_-F(\omega) + d\eta_2(\omega) \qquad (2.55b)$$

where

$$\Gamma_- = diag\left\{e^{j\frac{\Delta\cdot\delta_0}{2}}, e^{j\frac{\Delta\cdot\delta_1}{2}}, ..., e^{j\frac{\Delta\cdot\delta_{p-1}}{2}}\right\}$$

We shall now compute the spectral matrix of dipole array output. For this the incident signal will be treated as a stationary stochastic process. Then, in place of the ordinary Fourier transform we need to invoke the generalized Fourier transform of a stationary stochastic process (see p.40).

$$\frac{1}{2\pi}\mathbf{S}_{\tilde{f}}d\omega = E\left\{\frac{1}{2\pi}d\tilde{\mathbf{F}}\frac{1}{2\pi}d\tilde{\mathbf{F}}^{H}\right\} = E\left\{\frac{1}{2\pi}\begin{bmatrix}d\mathbf{F}_1\\d\mathbf{F}_2\end{bmatrix}\frac{1}{2\pi}\begin{bmatrix}d\mathbf{F}_1\\d\mathbf{F}_2\end{bmatrix}^{H}\right\}$$

$$= \frac{1}{2\pi}\begin{bmatrix}\mathbf{AS}_f\mathbf{A}^{H} & \mathbf{AS}_f\Gamma^{H}\mathbf{A}^{H}\\ \mathbf{A}\Gamma\mathbf{S}_f\mathbf{A}^{H} & \mathbf{AS}_f\mathbf{A}^{H}\end{bmatrix}d\omega + \frac{1}{2\pi}\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\sigma_{\eta_1}^2\\ & \sigma_{\eta_2}^2\end{bmatrix}d\omega$$

(2.56)

where

$$\Gamma = \Gamma_{-}\Gamma_{+}^{H} = diag\left\{e^{j\omega\Delta\cdot\delta_0}, e^{j\omega\Delta\cdot\delta_1},\ldots,e^{j\omega\Delta\cdot\delta_{p-1}}\right\}$$

$$\mathbf{S}_f = diag\left\{S_{f_0}, S_{f_1},\ldots,S_{f_{p-1}}\right\}$$

Equation (2.56) may be expressed as

$$\mathbf{S}_{\tilde{f}} = \begin{bmatrix}\mathbf{A}\\ \mathbf{A}\Gamma\end{bmatrix}\mathbf{S}_f\begin{bmatrix}\mathbf{A}\\ \mathbf{A}\Gamma\end{bmatrix}^{H} + \begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\sigma_{\eta_1}^2\\ & \sigma_{\eta_2}^2\end{bmatrix}$$

(2.57)

The DOA information is buried in Γ while the $\mathbf{A}$ matrix contains dipole location information, which is known. We shall show later in chapter 5 how Γ can be estimated using the subspace approach.

Electric Dipole Array: Now consider a ULA of electric dipoles. There are two possible arrangements: a dipole axis along an x-axis or along an y-axis as illustrated in fig. 2.24. Note that an electric dipole will measure the electric field component along its axis. Let a plane EM wave with its wave vector in the y-z plane be incident on both types of arrays (see figs. 2.24 a&b) of electric dipoles.

Thus, we are simultaneously measuring the x and y components of the electric field. The array outputs are given by

(a) for dipoles oriented in the x direction

$$\mathbf{f}_x = -\cos\gamma\left[1, e^{j\frac{\omega d}{c}\sin\theta}, e^{j2\frac{\omega d}{c}\sin\theta},\ldots,e^{j(M-1)\frac{\omega d}{c}\sin\theta}\right]^{T}Ee^{j(\omega t+\varphi)}$$

$$= -\cos\gamma\,\mathbf{a}Ee^{j(\omega t+\varphi)}$$

(b) for dipoles oriented in the y direction
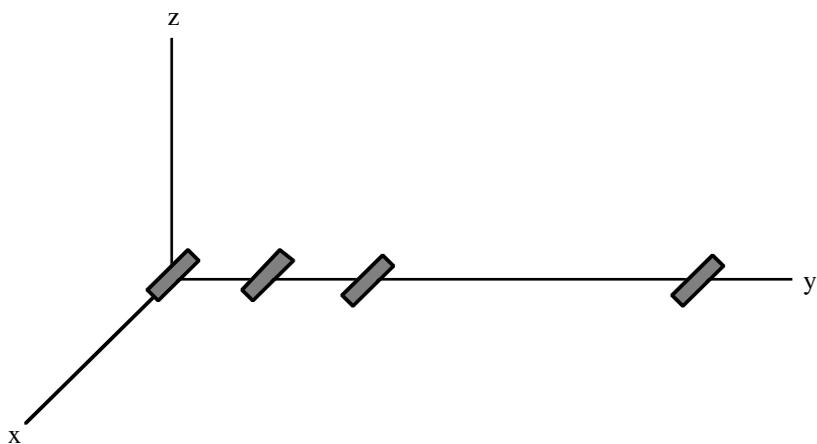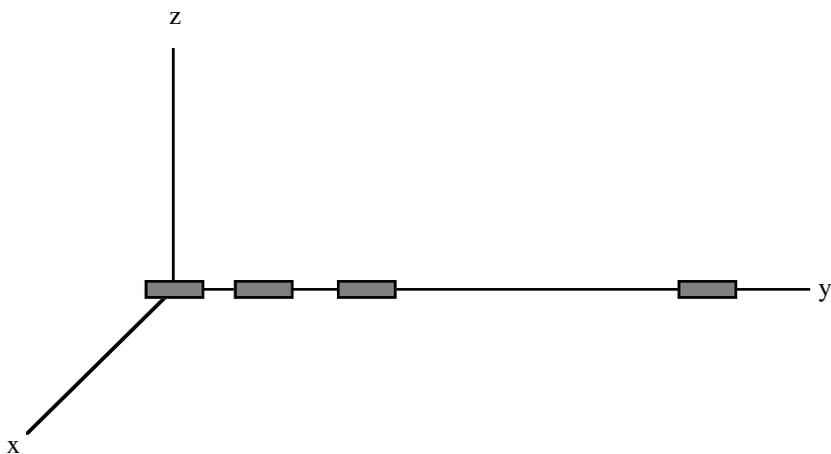
Figure 2.24: A ULA of electric dipoles oriented in the x direction (a) and in the y direction (b).

$$\mathbf{f}_y = \sin\gamma\cos\theta e^{j\kappa}\left[1, e^{j\frac{\omega d}{c}\sin\theta}, e^{j2\frac{\omega d}{c}\sin\theta}, \ldots, e^{j(M-1)\frac{\omega d}{c}\sin\theta}\right]^T Ee^{j(\omega t+\varphi)}$$

$$= \sin\gamma\cos\theta e^{j\kappa}\mathbf{a}\, Ee^{j(\omega t+\varphi)}$$

where θ is DOA and γ and κ are defined in terms of the polarization parameters (see p. 8). When there are P wavefronts incident on a ULA of dipoles, the array output may be expressed as follows:

$$\mathbf{f}_x = \mathbf{A}\Gamma_1 \mathbf{S}_0 e^{j\omega t} + \eta_1(t) \tag{2.58}$$

where

$$\Gamma_1 = -diag\{\cos\gamma_0, \cos\gamma_1, \ldots, \cos\gamma_{P-1}\}$$
$$\mathbf{A} = \{\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_{P-1}\}$$
$$\mathbf{S}_0 = col\{E_0 e^{j\varphi_0}, E_1 e^{j\varphi_1}, \ldots, E_{P-1} e^{j\varphi_{P-1}}\}$$

and $\eta_1(t) = col\{\eta_0'(t), \eta_1'(t), \ldots, \eta_{M-1}'(t)\}$ is the background noise vector. A similar expression for an array of dipoles oriented along the y-axis is given by

$$\mathbf{f}_y = \mathbf{A}\Gamma_2 \mathbf{S}_0 e^{j\omega t} + \eta_2(t) \tag{2.59}$$

where

$$\Gamma_2 = diag\{\sin\gamma_0 \cos\theta_0 e^{j\kappa_0}, \sin\gamma_1 \cos\theta_1 e^{j\kappa_1}, \ldots, \sin\gamma_{P-1} \cos\theta_{P-1} e^{j\kappa_{P-1}}\}$$

$$\eta_2(t) = col\{\eta_0''(t), \eta_1''(t), \ldots, \eta_{M-1}''(t)\}$$

§2.3 Broadband Sensor Array:

Very often an array of sensors is required to receive broadband signals, which include both natural signals (e.g., seismic and sonar signals) or man-made signals (e.g., communication signals). Not only must the sensors be broadband but also the special processing technique must be devised to exploit the broadband property of the signal. The bandwidth of the array response function depends upon the time taken by the wavefront to sweep across the array aperture, hence on the DOA for ULA but only on the diameter of UCA. A broadband signal may be considered as an aggregate of many narrowband signals covering the entire bandwidth. Since every narrowband signal is capable of determining the source parameters, we have considerable redundant information which may be used to fight against the noise and the model uncertainties. Alternatively, a broadband signal may also be represented by a collection of independent temporal samples (sampled at Nyquist rate) or snapshots, each carrying information about the source parameters. In this section we shall introduce the concept of array bandwidth, and frequency and time snapshots.

2.3.1 <u>Bandwidth of an Array</u>: The bandwidth of an array is the bandwidth of its transfer function which for a ULA is given by (2.3a)

$$H(\omega\tau) = \frac{1}{M}\sum_{m=0}^{M-1} e^{j\omega\frac{md}{c}\sin\theta} = \frac{\sin(\frac{M}{2}\omega\tau)}{M\sin\frac{\omega\tau}{2}} e^{j\frac{M-1}{2}\omega\tau}$$

A plot of $H(\omega\frac{d}{c}\sin\theta)$ with θ held fixed but ω varied will be same as in fig. 2.2 except that the x-axis variation is now linear in ω, in place of a noninear variation in θ. If we define the bandwidth as one half of the distance between the first nulls, we obtain

$$\Delta\omega = \frac{2\pi}{M}\frac{c}{d\sin\theta} = \underbrace{\frac{2\pi}{Md}}_{c_x} = \frac{2\pi}{\tau} \qquad (2.60)$$

where τ is the time required to sweep across the array aperture. Note that the bandwidth is infinite when the wavefront is incident on the broadside, that is, the array is steered in the direction of source; but it is equal to $\frac{2\pi c}{Md}$ when the wavefront is incident from the endfire. For a circular array, however, the wavefront has to sweep a constant aperture equal to the diameter of the circle, independent of the azimuth angle. This essential difference between a ULA and a UCA is illustrated in fig. 2.25. The bandwidth of a long ULA or a large diameter UCA is very small and hence much of the energy of a broadband source will be lost, unless the array is steered in the direction of the desired source. Further, as pointed in §2.1, when there is more than one source radiating at different frequencies there is a possibility for ambiguity. Let P narrowband sources with DOAs $\theta_i, i = 0,1,...,P-1$ and center frequencies ω_i, $i = 0,1,...,P-1$ be incident on a ULA. Further, we assume that the center frequency and DOA pair satisfies the following relation:

$$\omega_i \frac{d}{c}\sin\theta_i = \tau_0(cons\tan t), \quad i = 0,1,...,P-1 \qquad (2.61)$$

where τ_0 is the delay per sensor introduced by all sources. Now, through a process of sum and delay, the array is steered *simultaneously*, to all sources; in other words, the array will "see" all P sources at the same time. When the sources are broadband with overlapping spectrum we can always find a set of frequencies which satisfies (2.61). As a result, when the array is steered to one of the sources the output may be contaminated with the power derived from other sources. Such an interference is unacceptable particularly when waveform
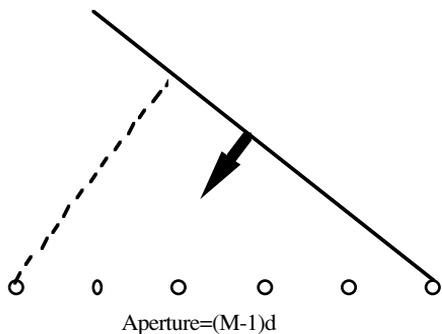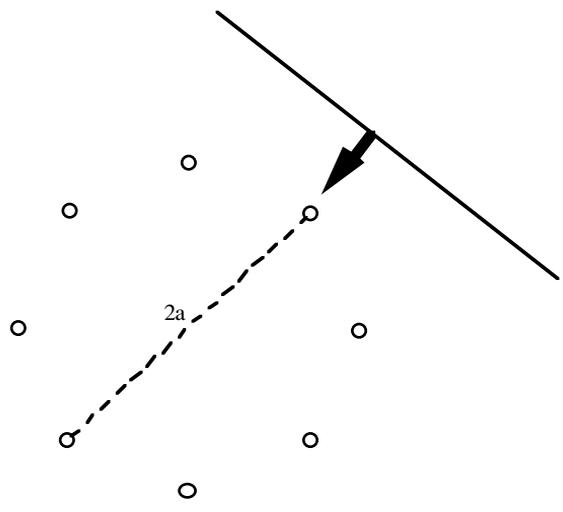
Figure 2.25: The essential difference between a ULA and UCA is that the effective aperture (dashed line) is azimuth dependent for a ULA but independent of the azimuth for a UCA. (a) Linear array, the effective aperture= $(M-1)d \sin \theta$. (b) Circular array, the effective aperture = diameter of the circle.

estimation is the goal. Some of these problems can be overcome by introducing temporal processing over and above the delay and sum type of spatial processing.

A snapshot may be expressed in the frequency domain as

$$\mathbf{f}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{a}(\omega, \theta_0) dF_0(\omega) e^{j\omega t} \qquad (2.62)$$

where

$$\mathbf{a}(\omega, \theta_0) = col\left\{1, e^{-j\omega \frac{d \sin \theta_0}{c}}, \dots, e^{-j\omega(M-1)\frac{d \sin \theta_0}{c}}\right\}$$

On account of the properties of the steering vector listed on page 92, we can conclude the following about the broadband array output:

(a) $\mathbf{f}_{\theta_0}(t) = \mathbf{f}_{\pi-\theta_0}(t)$. Follows from property (a) on page 92.

(b) $\mathbf{f}_{\theta_0}(t) \neq \mathbf{f}_{-\theta_0}(t)$. From property (b) and (2.61) we can write

$$\mathbf{f}_{-\theta_0}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{a}^*(\omega, \theta_0) dF_0(\omega) e^{j\omega t} \neq \mathbf{f}_{\theta_0}(t)$$

except in the unlikely event of $dF_0(\omega)$ being real. Thus, the east-west ambiguity shown for the narrowband complex signal does not apply to the broadband signals.

2.3.2 <u>Broadband Signals</u>: In the case of thebroadband signal, since a snapshot may vary rapidly, it is necessary that many snapshots must be collected at different time instants in the past; for example, $\mathbf{f}(t), \mathbf{f}(t-\Delta t), \mathbf{f}(t-2\Delta t), \dots,$ $\mathbf{f}(t-(N-1)\Delta t)$ are N past or delayed snapshots (see fig. 2.26). An alternate approach, in the case of broadband signal, is to go over to the frequency domain (temporal frequency). The output of each sensor, consisting of N samples, is subjected to Fourier analysis (DFT). A collection of the Fourier coefficients, one from each sensor at a fixed frequency, constitutes a *frequency* snapshot. The array signal processing of broadband signals using the frequency snapshots closely follows the time domain approach for narrowband signals (after removing the carrier), widely used in radar signal processing. In place of a covariance matrix we use a spectral matrix which indeed is a spatial covariance matrix (SCM). In this book we shall emphasize the frequency domain approach as the time domain approach for wideband signals turns out to be conceptually a bit more involved. We begin with time domain approach. First, let us introduce

some new quantities required in the processing. All delayed snapshots are stacked one above the other to form one large vector of size $M \cdot N$,

$$\mathbf{f}_{MN}(t) = col\left\{\mathbf{f}^T(t), \mathbf{f}^T(t-\Delta t), \mathbf{f}^T(t-2\Delta t),...,\mathbf{f}^T(t-(N-1)\Delta t)\right\}$$

(2.63)

We define a covariance matrix known as spatio-temporal covariance matrix, STCM,

$$\mathbf{C}_{STCM} = E\left\{\mathbf{f}_{ML}(t)\mathbf{f}_{ML}^H(t)\right\}$$

(2.64)

As an example, consider a two sensor array (M=2) and two delayed snapshots (N=2). The STCM is given by

$$\mathbf{C}_{STCM} =$$

$$\begin{bmatrix} C_f(0) & C_f(\tau_0) & C_f(\Delta t) & C_f(\Delta t + \tau_0) \\ C_f(-\tau_0) & C_f(0) & C_f(\Delta t - \tau_0) & C_f(\Delta t) \\ C_f(\Delta t) & C_f(\Delta t - \tau_0) & C_f(0) & C_f(\tau_0) \\ C_f(\Delta t + \tau_0) & C_f(\Delta t) & C_f(-\tau_0) & C_f(0) \end{bmatrix}$$

where $\tau_0 = \dfrac{d \sin \theta_0}{c}$ is a propagation delay per sensor. We express the stacked vector by using (2.62) in (2.63),

$$\mathbf{f}_{stacked}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \zeta(\omega) \otimes \mathbf{a}(\omega, \theta_0) dF_0(\omega)e^{j\omega t}$$

(2.65)

where

$$\zeta(\omega) = col\left\{1, e^{-j\omega \Delta t},..., e^{-j\omega(L-1)\Delta t}\right\}$$

and $\otimes$ stands for the Kronecker product. Define a direction vector $\mathbf{h}(\omega, \theta_0) = \zeta(\omega) \otimes \mathbf{a}(\omega, \theta_0)$ and rewrite (2.65) as
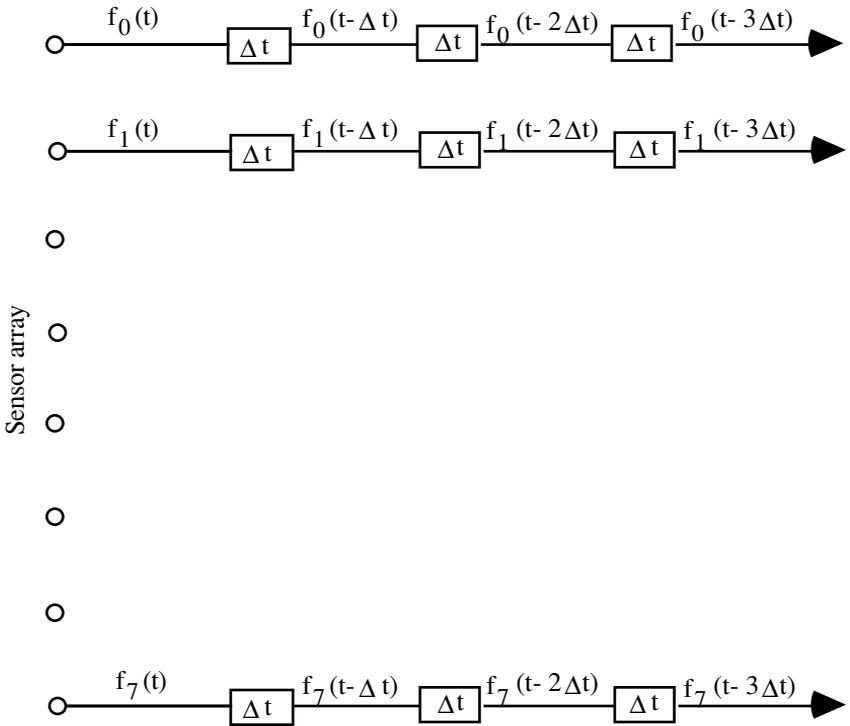
Figure 2.26: The output of an eight sensor ULA is sampled at different time instants in the past. All outputs taken at the same time instant are grouped into a vector called a *temporal* snapshot. We have as many *temporal* snapshots as the number of samples or taps taken from each sensor output.

$$\mathbf{f}_{stacked}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{h}(\omega, \theta_0) dF_0(\omega) e^{j\omega t} \qquad (2.66)$$

which may be considered as an extended output of the array (stacked vector) due to a source at azimuth angle θ_0. Using (2.66) we can express STCM as

$$\mathbf{C}_{STCM} = E\{\mathbf{f}_{stacked}(t)\mathbf{f}_{stacked}^H(t)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{h}(\omega, \theta_0) \mathbf{h}^H(\omega, \theta_0) S_{f_0}(\omega) d\omega$$

$$(2.67)$$

Let $d\mathbf{F}_{stacked}(\omega)$ be the generalized Fourier transform of the extended vector output, $\mathbf{f}_{stacked}(t)$,

$$d\mathbf{F}_{stacked}(\omega) = \mathbf{h}(\omega, \theta_0) dF_0(\omega) \tag{2.68}$$

Premultiply on both sides of (2.68) with the steering vector in some desired direction and evaluate the expected value of the magnitude square. Dividing by $(M \cdot N)^2$ we get the output power

$$\frac{1}{(M \cdot N)^2} \mathbf{h}^H(\omega, \theta) \mathbf{S}_{f_{stacked}}(\omega) \mathbf{h}(\omega, \theta) =$$

$$\left| \frac{\mathbf{h}^H(\omega, \theta) \mathbf{h}(\omega, \theta_0)}{M \cdot N} \right|^2 S_0(\omega) = \left| \frac{\mathbf{a}^H(\omega, \theta) \mathbf{a}(\omega, \theta_0)}{M} \right|^2 S_0(\omega) \tag{2.69}$$

where $\mathbf{S}_{f_{stacked}}(\omega)$ is the spectral matrix of an extended array signal. Interestingly, in spite of delayed snapshots used in deriving (2.69), the output power remains the same as in (2.19b). However, the STCM is different from the spatial covariance matrix (SCM). For example, reconsider the frequency-direction ambiguity due to the fact that if there is a set of DOAs and frequencies such that

$$\omega_i \frac{d}{c} \sin\theta_i = \tau_0 (cons \tan t), i = 0, 1, \ldots, P - 1,$$

the steering vector $\mathbf{a}(\omega_i, \theta_i)$ remains unchanged. This, however, is not true for a ULA with a tapped delay line, as the steering vectors, $\mathbf{h}(\omega_i, \theta_i) = \zeta(\omega_i) \otimes \mathbf{a}(\omega_i, \theta_i)$, $i = 0, 1, \ldots, P - 1$, will be different because of $\zeta(\omega_i)$. It may be noted that for a circular array there is no frequency-direction ambiguity of the type described above even when no tapped delay line is used [18].

§2.4 Source and Sensor Arrays:
In nonoptical imaging systems such as in radar, sonar, seismic, and biomedical imaging systems, we often use an array of sources and an array of sensors. The main idea is to illuminate an element to be imaged from different directions and collect as much of the scattered radiation as possible for the purpose of imaging. To achieve this goal in a straightforward manner would require an impractically large array. It is known that a large aperture array can be synthesized from a small array of transmitters and receivers (transceivers). This leads to the concept of coarray used in radar imaging. In tomographic imaging the element to be imaged has to be illuminated from all directions. We,

therefore, need a boundary array of transceivers or a single source and a line of sensors going round the object as in some biomedical imaging systems. In seismic exploration a source array is often used to provide the additional degree of freedom required to combat the high noise and the interference level, and also to estimate the wave speed in different rock layers.

2.4.1 <u>Coarray</u>: An imaging system is basically concerned with faithful mapping of a point target in the object space into a point in the image space. The point spread function (PSF) of the imaging system describes the mapping operation. Ideally, it is a delta function, but in practice a point may be mapped into a small area, the size of which will depend upon the aperture of the array used in the imaging system. The coarray is required to synthesize an arbitrary PSF using a finite array of sensors and sources. A source array (transmit array) is weighted with a complex function t(x) and the sensor array (receive array) is also weighted with another complex weight function r(y). The response function of a source array in the direction of the scatterer (in far field) is given by

$$T(u,v) = \int\int_{-\infty}^{+\infty} t(x)\delta(y)e^{j(ux+vy)}dxdy \qquad (2.70a)$$

and the response function of the sensor array in the direction of the scatterer is given by

$$R(u,v) = \int\int_{-\infty}^{+\infty} r(y)\delta(x)e^{j(ux+vy)}dxdy \qquad (2.70b)$$

where $u = k\sin\theta\cos\varphi$, $v = k\sin\theta\sin\varphi$ and $k = \dfrac{2\pi}{\lambda}$ is the wavenumber. Let $\rho(u,v)$ be the reflection coefficient as a function of the azimuth angle φ and the elevation angle θ. The total response of the source/sensor array is given by

$$O(u,v) = T(u,v)\rho(u,v)R(u,v)$$

This is the image of a point scatterer in the frequency domain. If we assume that the scatterer is omnidirectional, that is, $\rho(u,v) = \rho_0$ (constant), the PSF of the array will be given by the inverse Fourier transform of $T(u,v)R(u,v)$

$$PSF = \sum_x\sum_y t(x)\delta(y)r(q-y)\delta(p-x)$$

$$= t(p)r(q)$$

Thus, we note that the combined PSF of a source array along the x-axis and a sensor array along the y-axis is equal to the product of two weighting functions;

for example, when the source and sensor arrays are uniformly weighted, the PSF of the combined system is a rectangular function [19]. This is illustrated in fig. 2.27 for an L-shaped array. For an arbitrary distribution of sources and sensors the PSF is given by a convolution of the source and sensor weighting functions,

$$PSF = \int\int_{-\infty}^{+\infty} t(x,y)r(p-x,q-y)dxdy \qquad (2.71)$$

The PSF at point (p,q) is obtained by summing the product of two weight functions over two lines p=x+x´ and q=y+y´ where (x,y) refer to a point in the source array and (x´,y´) refer to a point in the sensor array. Such a coarray is also known as a sum coarray. Since we have a discrete set of point sources and sensors, the discrete version of equation (2.71) is given by $\mathbf{R}^T\mathbf{T}$ where $\mathbf{T} = col\{t_0, t_1, \ldots t_{M-1}\}$ and $\mathbf{R} = col\{r_0, r_1, \ldots r_{M-1}\}$.

An arbitrary PSF may be synthesized through eigen-decomposition. Let $\mathbf{P}$ be the desired PSF given in the form of a matrix which we shall assume to have a hermitian symmetry. We can express its eigen-decomposition as $\mathbf{P} = \sum_{l=0}^{M-1} \lambda_l \mathbf{u}_l^H \mathbf{u}_l$ where λ_l is lth eigenvalue (real) and $\mathbf{u}_l$ is the corresponding eigenvector. We let $\mathbf{T} = \mathbf{R} = \mathbf{u}_l$, that is, both source and sensor arrays are weighed by the same eigenvector of the desired PSF matrix, and thus an image is formed. This step is repeated over all significant eigenvectors. All such images are linearly combined after weighting each with the corresponding eigenvalue [20].

Let us consider an example of circular array of eight transceivers uniformly distributed on the circumference. A transceiver consists of a source and a sensor physically placed at the same place. The transceivers are located at

$$\left[\begin{array}{l} (a,0), (\frac{a}{\sqrt{2}},\frac{a}{\sqrt{2}}), (0,a), (-\frac{a}{\sqrt{2}},\frac{a}{\sqrt{2}}), (0,-a), \\ (-\frac{a}{\sqrt{2}},-\frac{a}{\sqrt{2}}), (-a,0), (\frac{a}{\sqrt{2}},-\frac{a}{\sqrt{2}}) \end{array} \right]$$

where a stands for radius of the circle. The sum coarray consists of 64 locations whose coordinates may be found by summing the 1st column and the 1st row entries from table 2.2. The actual coordinates, thus obtained, are also shown in the table. There are in all $\frac{N^2}{2}+1$ distinct nodes. The coarray nodes lie on a set of concentric circles of radii (0, 0.76a, 1.141a, 1.85a, 2a) as shown in fig. 2.28.
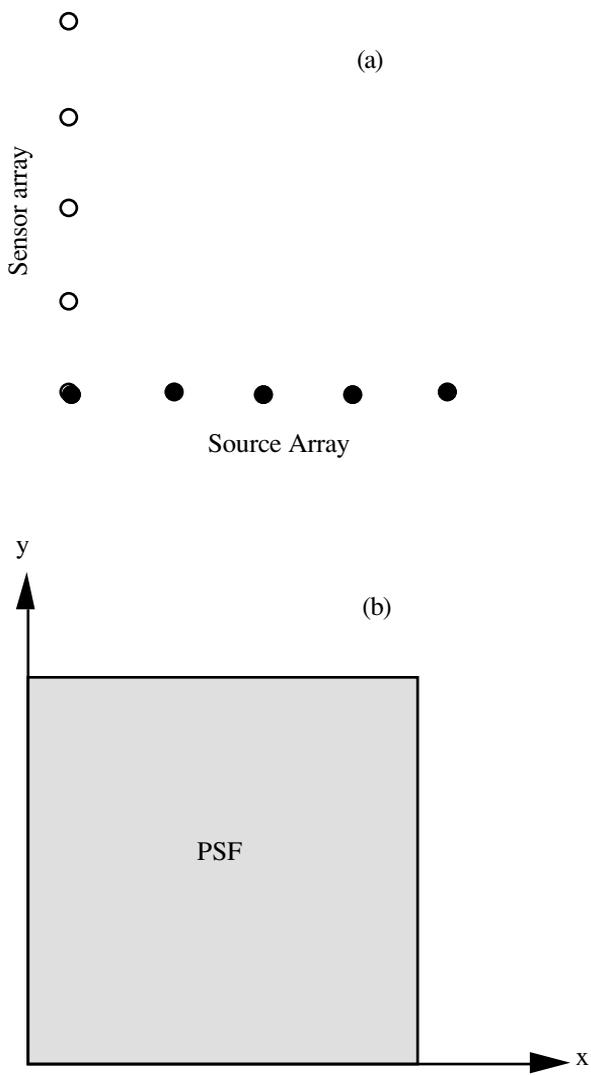
Figure 2.27: (a) Source and sensor array. (b) Point spread function (PSF) on a grid, known as coarray.

The synthesized aperture at (m,n)th node of the circular coarray is given by

$$t(a\cos(m\frac{2\pi}{M}),a\sin(m\frac{2\pi}{M}))$$

Column entry

$\longrightarrow$ m

	$(a,0)$	$(\frac{a}{\sqrt{2}}, \frac{a}{\sqrt{2}})$	$(0,a)$	$(-\frac{a}{\sqrt{2}}, \frac{a}{\sqrt{2}})$	$(0,-a)$	$-(\frac{a}{\sqrt{2}}, \frac{a}{\sqrt{2}})$	$(-a,0)$	$(\frac{a}{\sqrt{2}}, -\frac{a}{\sqrt{2}})$
$(a,0)$	$(2a,0)$	$(a+\frac{a}{\sqrt{2}}, \frac{a}{\sqrt{2}})$	(a,a)	$(a-\frac{a}{\sqrt{2}}, \frac{a}{\sqrt{2}})$	$(a,\text{-}a)$	$(a-\frac{a}{\sqrt{2}}, -\frac{a}{\sqrt{2}})$	$(0,0)$	$(a+\frac{a}{\sqrt{2}}, -\frac{a}{\sqrt{2}})$
$(\frac{a}{\sqrt{2}}, \frac{a}{\sqrt{2}})$	$(a+\frac{a}{\sqrt{2}}, \frac{a}{\sqrt{2}})$	$(\sqrt{2}a, \sqrt{2}a)$	$(\frac{a}{\sqrt{2}}, a+\frac{a}{\sqrt{2}})$	$(0, \sqrt{2}a)$	$(\frac{a}{\sqrt{2}}, \frac{a}{\sqrt{2}}-a)$	$(0,0)$	$(\frac{a}{\sqrt{2}}-a, \frac{a}{\sqrt{2}})$	$(\sqrt{2}a, 0)$
$(0,a)$	(a,a)	$(\frac{a}{\sqrt{2}}, a+\frac{a}{\sqrt{2}})$	$(0,2a)$	$(-\frac{a}{\sqrt{2}}, a+\frac{a}{\sqrt{2}})$	$(0,0)$	$(-\frac{a}{\sqrt{2}}, a-\frac{a}{\sqrt{2}})$	$(\text{-}a,a)$	$(\frac{a}{\sqrt{2}}, a-\frac{a}{\sqrt{2}})$
$(-\frac{a}{\sqrt{2}}, \frac{a}{\sqrt{2}})$	$(a-\frac{a}{\sqrt{2}}, \frac{a}{\sqrt{2}})$	$(0, \sqrt{2}a)$	$(-\frac{a}{\sqrt{2}}, a+\frac{a}{\sqrt{2}})$	$(-\sqrt{2}a, \sqrt{2}a)$	$(-\frac{a}{\sqrt{2}}, -a+\frac{a}{\sqrt{2}})$	$(-\sqrt{2}a,0)$	$(-a-\frac{a}{\sqrt{2}}, \frac{a}{\sqrt{2}})$	$(0,0)$
$(0,-a)$	$(a,\text{-}a)$	$(\frac{a}{\sqrt{2}}, \frac{a}{\sqrt{2}}-a)$	$(0,0)$	$(-\frac{a}{\sqrt{2}}, -a+\frac{a}{\sqrt{2}})$	$(0,\text{-}2a)$	$(-\frac{a}{\sqrt{2}}, -a-\frac{a}{\sqrt{2}})$	$(\text{-}a,\text{-}a)$	$(\frac{a}{\sqrt{2}}, -a-\frac{a}{\sqrt{2}})$
$-(\frac{a}{\sqrt{2}}, \frac{a}{\sqrt{2}})$	$(a-\frac{a}{\sqrt{2}}, -\frac{a}{\sqrt{2}})$	$(0,0)$	$(-\frac{a}{\sqrt{2}}, a-\frac{a}{\sqrt{2}})$	$(-\sqrt{2}a,0)$	$(-\frac{a}{\sqrt{2}}, -a-\frac{a}{\sqrt{2}})$	$(-\sqrt{2}a, -\sqrt{2}a)$	$(-a-\frac{a}{\sqrt{2}}, -\frac{a}{\sqrt{2}})$	$(0, -\sqrt{2}a)$
$(-a,0)$	$(0,0)$	$(\frac{a}{\sqrt{2}}-a, \frac{a}{\sqrt{2}})$	$(\text{-}a,a)$	$(-a-\frac{a}{\sqrt{2}}, \frac{a}{\sqrt{2}})$	$(\text{-}a,\text{-}a)$	$(-a-\frac{a}{\sqrt{2}}, -\frac{a}{\sqrt{2}})$	$(\text{-}2a,0)$	$(-a+\frac{a}{\sqrt{2}}, -\frac{a}{\sqrt{2}})$
$(\frac{a}{\sqrt{2}}, -\frac{a}{\sqrt{2}})$	$(a+\frac{a}{\sqrt{2}}, -\frac{a}{\sqrt{2}})$	$(\sqrt{2}a,0)$	$(\frac{a}{\sqrt{2}}, a-\frac{a}{\sqrt{2}})$	$(0,0)$	$(\frac{a}{\sqrt{2}}, -a-\frac{a}{\sqrt{2}})$	$(0, -\sqrt{2}a)$	$(\frac{a}{\sqrt{2}}-a, -\frac{a}{\sqrt{2}})$	$(\sqrt{2}a, -\sqrt{2}a)$

$\uparrow$ Row entry

Table 2.2: The sum coarray consists of 64 locations whose coordinates are obtained by summing the 1st column (left shaded column) and the respective entries from the 1st row (top shaded row). The coordinates are shown in clear cells.

$$r(a\cos(n\frac{2\pi}{M}), a\sin(n\frac{2\pi}{M}))$$

The PSF at (p,q) where $p = x_m + x_n$ and $q = y_m + y_n$ may be obtained from the discrete equivalent of (2.70). Note that the coordinates of the mth point are (x_m, y_m) and those of the nth point are (x_n, y_n).

$$PSF(p,q)$$

$$= t(a\cos(m\frac{2\pi}{M}), a\sin(m\frac{2\pi}{M}))r(a\cos(m\frac{2\pi}{M}), a\sin(m\frac{2\pi}{M}))$$

$$m = n$$

$$= \sum_{i=0}^{\frac{M}{2}-1} \left[\begin{array}{l} t(a\cos(i\frac{2\pi}{M}), a\sin(i\frac{2\pi}{M}))r(a\cos((i+\frac{M}{2})\frac{2\pi}{M}), \\ a\sin((i+\frac{M}{2})\frac{2\pi}{M})) + \\ r(a\cos(i\frac{2\pi}{M}), a\sin(i\frac{2\pi}{M}))t(a\cos((i+\frac{M}{2})\frac{2\pi}{M}), \\ a\sin((i+\frac{M}{2})\frac{2\pi}{M})) \end{array} \right]$$

$$|m - n| = \frac{M}{2}$$

$$= t(a\cos(m\frac{2\pi}{M}), a\sin(m\frac{2\pi}{M}))r(a\cos(n\frac{2\pi}{M}), a\sin(n\frac{2\pi}{M}))$$

$$+ r(a\cos(m\frac{2\pi}{M}), a\sin(m\frac{2\pi}{M}))t(a\cos(n\frac{2\pi}{M}), a\sin(n\frac{2\pi}{M}))$$ (2.72)

$$m \neq n$$

$$|m - n| \neq \frac{M}{2}$$

As an illustration, consider uniform transmitter strength and receiver sensitivity; let $t = t_0$ and $r = r_0$. Clearly, since the array has a circular symmertry, PSF will also be radially symmetric. It is straightforward to show that

$$PSF(0,0) = 8t_0 r_0$$

$$PSF(\sqrt{p^2 + q^2}) = 2t_0 r_0$$ (2.73)

The PSF has a tall peak at the center, four times the background level.
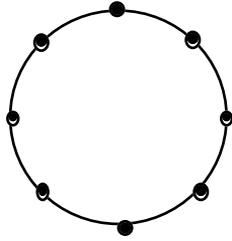
(a)
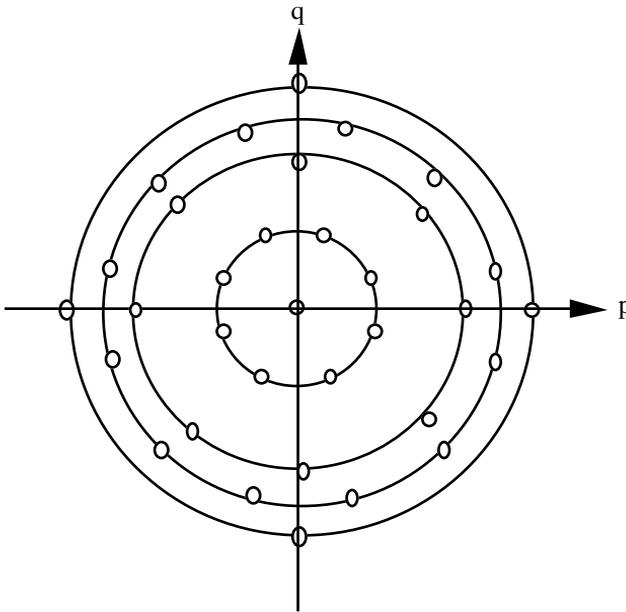
(b)

Figure 2.28 (a) A circular array (radius=1.5cm) of eight transceivers. (b) Coarray: Nodes are shown by empty circles. There are 33 nodes. The diameter of the outermost circle is 3.0cm. The PSF is defined at the nodes. For uniform array, with constant source strength and sensor sensitivity, the PSF at the center is equal to $8\,t_0 r_0$ and elsewhere it is equal to $2\,t_0 r_0$.

In general, with M sensors $PSF(0,0) = Mt_0 r_0$ but $PSF(\sqrt{p^2 + q^2}) = 2t_0 r_0$. Thus, for large M the PSF tends to a delta function, the ideal condition for a perfect imaging system. A reverse problem of synthesizing a circular array of transceivers given the desired PSF is proposed in [19].

2.4.2 <u>Passive Imaging</u>: In passive imaging the object to be imaged is itself a source of energy, for example, a distant star or an earthquake deep inside the earth. It is of some interest to estimate the distribution of the energy as a function of the azimuth and elevation or the spatial frequencies u and v ($u = \dfrac{2\pi}{\lambda}\sin\theta\cos\varphi$ and $v = \dfrac{2\pi}{\lambda}\sin\theta\sin\varphi$). The imaging system consists of two receiving arrays each with its own weighting functions. Let $r_1(x,y)$ and $r_2(x,y)$ be the weighting functions. Let $P(u,v)$ be the source distribution. The outputs of the two arrays, in frequency domain, can be written as

$$O_1(u,v) = P(u,v)R_1(u,v)$$
$$O_2(u,v) = P(u,v)R_2(u,v)$$

(2.74)

We form a cross-correlation of the two outputs. Using (2.74) we obtain

$$I(u,v) = O_1(u,v)O_2^*(u,v)$$
$$= |P(u,v)|^2 R_1(u,v)R_2^*(u,v)$$

(2.75a)

Or

$$I(u,v) = |P(u,v)|^2 \iint_{p\ q}\left[\int\int_{-\infty}^{+\infty} r_1(x,y)r_2(x+p,y+q)dxdy\right]e^{j(up+vq)}dpdq$$

(2.75b)

The quantity inside the square brackets in (2.75b) represents the cross correlation of two weighting functions.

$$C_{r_1 r_2}(p,q) = \int\int_{-\infty}^{+\infty} r_1(x,y)r_2(x+p,y+q)dxdy$$

(2.76)

Let $x' = x + p$ and $y' = y + q$. Then the cross-correlation may be looked upon as an integral of the product of weight functions at $r_1(x,y)$ and $r_2(x',y')$ for fixed p and q; alternatively, for a fixed difference, $p = x' - x$ and $q = y' - y$. Such a coarray is also known as a difference array. For example, for a L-shaped receiver array (sensors on both arms) the difference array is on a square grid in the fourth quadrant (see fig. 2.29). Recall that the sum coarray is also on a square grid but in the 1st quadrant (see fig. 2.27).

2.4.3 <u>Synthetic Aperture</u>: A long aperture is the primary requirement for achieving high spatial resolution; however there is a limit on the size of a physical array that one can deploy. This is particularly true in imaging systems

where it is uneconomical to use a large physical array required to cover a large target area with high resolution. The main idea in synthetic aperture is that by moving a transceiver, preferably at a constant speed and in a straight line, an array of large aperture can be synthesized through subsequent processing. A simple illustration of this principle is shown in fig. 2.30. A transceiver is moving at a constant speed c_s in a straight line parallel to the ground. Both transmit and receive beams are wide angle, as shown in the figure. At position p_1 the beam just begins to illuminate a scattering particle on the ground. The particle remains under illumination until the transceiver reaches position p_2. The energy scattered by the particle is received over an interval p_1p_2; hence the effective array aperture is equal to p_1p_2. This is the synthesized aperture. Let us now look into how to process the echoes received over p_1p_2. These echoes reach the receiver with different delays depending upon the position of the transceiver. Secondly, since the transceiver is moving, there will be a Doppler shift proportional to the component of transceiver speed in the direction of the radius vector joining the transceiver and the scattering particle. From fig. 2.30b it can be shown that $p_1p_2 = 2l_0 \tan\theta_0$. Let us assume that the transceiver is a directional sensor with effective aperture L; hence the beam width is $2\theta_0$ where $\sin\theta_0 = \lambda/L$. The beamwidth is measured between the first two nulls of the array response function (see p. 77-82). Therefore, the synthesized aperture may be expressed as

$$p_1p_2 = 2l_0 \frac{\lambda}{\sqrt{L^2 - \lambda^2}} \tag{2.77}$$

where l_0 is the height of the moving transceiver above ground. The smallest size of an object that can be seen on the ground with the help of an array with aperture given by (2.77) will be approximately equal to $\sqrt{L^2 - \lambda^2}$. The underlying assumption is that $p_1p_2 \gg \lambda$ which would allow us to simplify the Rayleigh resolution criterion (see p. 222) and then the result follows. The requirement that $p_1p_2 \gg \lambda$ is easily met by selecting a small antenna as a transmitter. The two way (one way in a passive system) travel time from the scatterer to the transceiver is given by

$$\frac{2l}{c} = \frac{2\sqrt{l_0^2 + (x - x')^2}}{c} = \frac{2\sqrt{l_0^2 + (c_s t)^2}}{c} \tag{2.78}$$
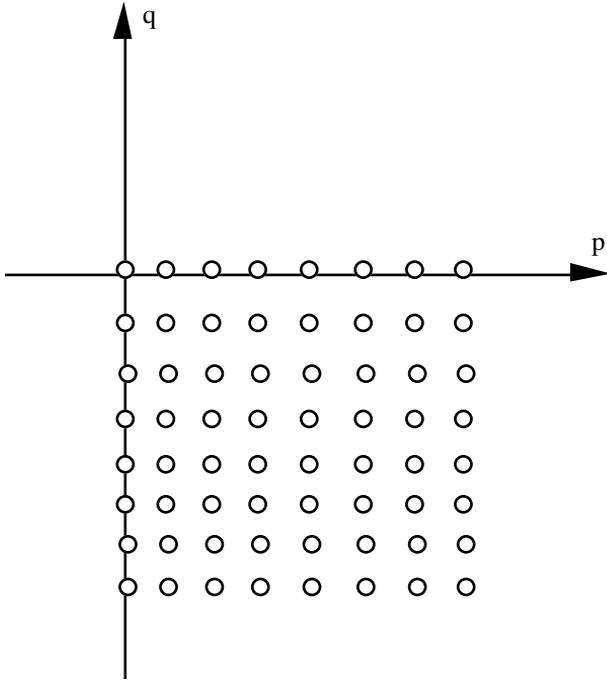
Figure 2.29 Difference coarray obtained from L-shaped receive array shown in fig. 2.24(a).

where we have expressed the horizontal distance as a product of time and transceiver speed, $(x - x') = c_s t$. Let the transmitted waveform be a sinusoid (real), $\cos(\omega t + \varphi)$. The received waveform which arrives after a delay of $\dfrac{2l}{c}$ is

$$f(t) = r_0 \cos(\omega(t - \frac{2\sqrt{l_0^2 + (c_s t)^2}}{c}) + \varphi) \tag{2.79}$$

Since the phase of the received waveform is time dependent, the instantaneous frequency [1] will be different from the frequency of the transmitted waveform. This difference is the Doppler shift. The instantaneous frequency is given by

$$\omega(t) = \omega - \frac{2\omega c_s}{c} \frac{c_s t}{\sqrt{l_0^2 + (c_s t)^2}}$$
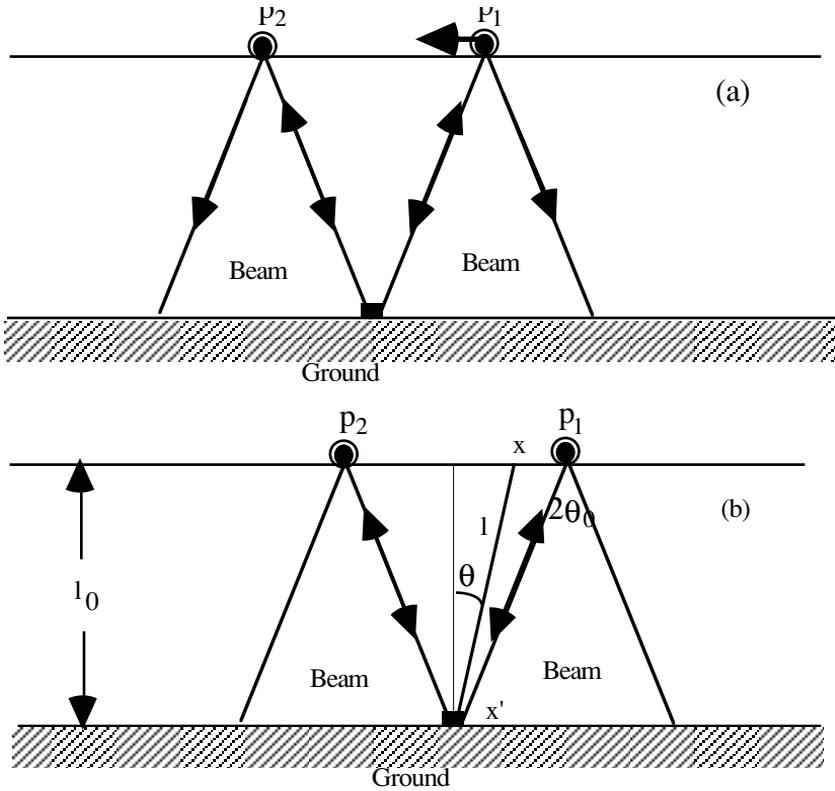
Figure 2.30: (a) A simple illustration of the principle of synthetic aperture. The energy scattered by a particle on the ground is received by a transceiver at p. (b) The two way travel time from the scatterer at x´ to the transceiver at x.

Hence, the Doppler shift is equal to

$$\Delta\omega = -\frac{2\omega c_s}{c}\frac{c_s t}{\sqrt{l_0^2 + (c_s t)^2}} = -\frac{2\omega c_s}{c}\sin\theta \qquad (2.80)$$

In the case of passive synthetic aperture sonar, the Doppler shift can be used for estimation of the direction of arrival of an unknown source [21].

At a fixed place and time the signal received consists of a sum of the scattered wavefields from a patch on the ground which is coherently illuminated (see fig. 2.31). The size of the patch is equal to the synthesized aperture. The wavefield from different scatterers reaches the sensor with delays as given by (2.78). The receiver output is given by a convolutional integral [22],

$$f(x,t) = \int\limits_{-\infty}^{\infty} r_0(x')w(x-x')\cos(\omega(t - \frac{2\sqrt{l_0^2 + (x-x')^2}}{c}) + \varphi)dx'$$

$$(2.81)$$

where $w(x - x') = 1$ for $|x - x'| \leq \dfrac{L}{2}$ and L=p1p2, the length of the synthetic aperture. For $|x - x'| << l_0$ we can approximate (2.81) as

$$f(x,t) = \int\limits_{p_1}^{p_2} r_0(x')w(x-x')\cos(\omega(t - \tau_0 - \frac{(x-x')^2}{cl_0}) + \varphi)dx'$$

$$(2.82a)$$

which simplifies to

$$f(x,t) = \cos(\omega(t - \tau_0))\int\limits_{-\infty}^{\infty} r_0(x')w(x-x')\cos(\omega(\frac{(x-x')^2}{cl_0}) + \frac{\varphi}{2})dx'$$

$$+ \sin(\omega(t - \tau_0))\int\limits_{-\infty}^{\infty} r_0(x')w(x-x')\sin(\omega(\frac{(x-x')^2}{cl_0}) + \frac{\varphi}{2})dx'$$

$$(2.82b)$$

We can recover $r_0(x)$ from $f(x,t)$, using the Fourier transform method, that is,

$$\tilde{r}_0(x) = FT^{-1}\left\{ \frac{FT\left\{\hat{f}(x)\right\}}{FT\left\{w(x)\cos(\omega(\frac{x^2}{cl_0}))\right\}} \right\}$$

$$(2.83)$$

where

$$\hat{f}(x) = \int\limits_{one\ period} f(x,t)\cos(\omega(t - \tau_0))dt$$

and FT stands for the Fourier transform. For a large aperture, we get $\tilde{r}_0(x) \xrightarrow[L \to \infty]{} r_0(x)$, that is, an exact reconstruction. It is important to remember that the increased aperture and therefore the increased resolution is the
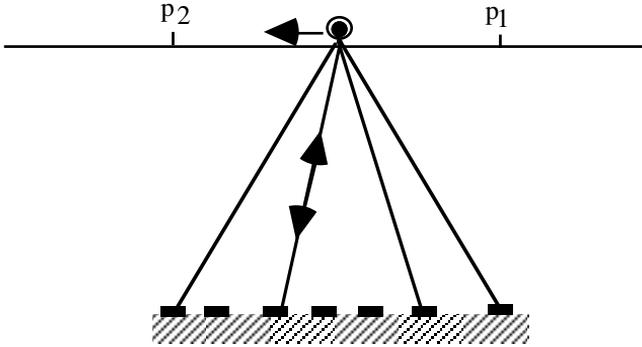
Figure 2.31: At a fixed place and time the signal received consists of a sum of scattered wavefields from a patch of scatterers which is coherently illuminated. The width of the patch is equal to the synthesized aperture.

result of geometry of the data collection as shown in fig. 2.30. Indeed, the transceiver need not move at all during pulse transmission and reception.

§2.5 Exercises:

1. The angle of arrival of a broadband signal at ULA may be estimated from the spectrum of the sum of the output of all sensors (without delays) and the spectrum of the output of anyone sensor. Find the bandwidth required to estimate the angle of arrival equal to 45° given that c=1500 m/s and d=15m. What is the role of the sensor spacing?

2. It is desired to measure the speed of wave propagation in a medium. A ULA with sensor spacing d meters is employed for this purpose. A broadband signal from a known direction is sweeping across the array. How do you go about estimating the speed of propagation?

3. Consider a circular transmit and receive array as shown in fig. 2.32. (a) Show the sum coarray along with weight coefficients assuming that the physical array has unit coefficient. (b) Do the same for the difference coarray.

4. Compare the outputs of crossed electric dipole arrays given by (2.58 & 2.59) with that of displaced identical sensors pairs (also known as dipoles) given by equations (2.55 a & b). Obtain an expression for the spectral matrix, similar to the one given in (2.57), of the output of the crossed dipole arrays.
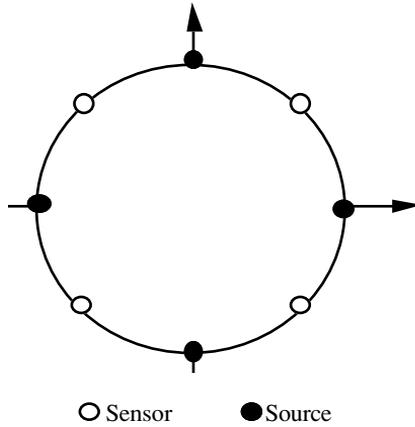
Figure 2.32: A circular sensor and source array. Compute sum and difference coarray
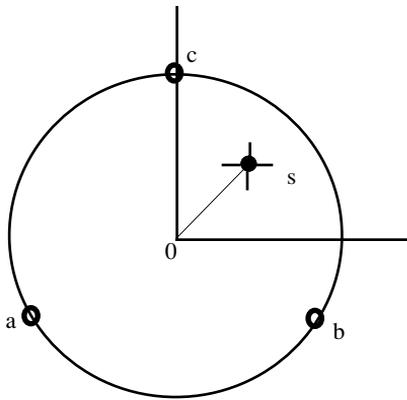
Figure 2.33: Triangular (equilateral) array and a source.

5. Consider an equilateral triangular array inscribed in a circle of radius=64λ. A point source is located at r=32λ and φ=45° (see fig. 2.33). Draw all phase trajectories passing through the point where the source is located.

6. A ULA is placed on a 100m deep sea bed and a broadband acoustic source is placed 1000m away from the array and 50m below the sea surface. Sketch the frequency wavenumber spectrum of the array signal. [Hint: Consider a direct and two reflected wavefronts.]

7. A broadband wavefront is incident on a ULA. Show that there is no aliasing for any angle of incidence if the sensor spacing is $\leq \lambda_{min}/2$.

References

1. P. S. Naidu, Modern spectrum analysis of time series, CRC Press, Boca Raton, Fl, 1996.

2. J. R. Williams, Fast beam-forming algorithm, J. Acoust. Soc. of Am., vol. 44, pp. 1454-1455, 1968.

3. S. H. Nawab, F. U. Dowla, R. T. Lacoss, Direction determination for wide band signals, IEEE Trans, ASSP-33, pp. 1114-1122, 1985.

4. M. Allam and A. Moghaddamjoo, Two dimensional DFT projection for wideband direction-of-arrival estimation, IEEE Trans. on Signal Proc., vol. 43, pp. 1728-1732, 1995

5. R. H. Tatham, Multidimensional filtering of seismic data, IEEE Proc., vol. 72, pp. 1357-1369, 1984.

6. H. Brysk and D. W. McCowan, Geophysics, vol. 51, pp. 1370-1386, 1986.

7. C. W. Therrien, Discrete random signals and statistical signal processing, Prentice Hall, Englewood Cliffs, NJ , 1992.

8 A. T. Moffet, Minimum redundancy linear arrays, IEEE Trans. Antenna Propagation, vol. AP-16, pp. 172-175, 1968.

9 Lee and S. U.Pillai, An algorithm for optimal placement of sensor elements, ICASSP-88, pp. 2674-2677, 1988.

10N. M. J. Earwicker, Signal to noise ratio gain of sparse array processors, J. Acoust. Soc. of Am., vol. 88, pp. 1129-1134, 1980.

11. B. G. Quinn, R. F. Barrett, P. J. Kootsookos and S. J. Searle, The estimation of the shape of an array using a hidden markov model, IEEE Jour. Ocean Engineering, vol. 18, pp. 557-564, 1993.

12. M. Hinich, Estimating signal and noise using a random array, J. Acoust. Soc. of Am., vol. 71, pp. 97-99, 1982.

13. M. P. Moody, Resolution of coherent sources incident on a circular antenna array, Proc. IEEE, vol.68, pp. 276-277, 1980.

14. C. P. Mathews and M. D. Zoltowaski, Eigen structure technique for 2D angle estimation with uniform circular arrays, IEEE Trans. SP-42, pp. 2395-2407, 1994.

15. M. Abramowitz and I. A. Stegun (Editors), Handbook of Mathematical Functions, National Bureau of Standards, Applied Mathematics Series 55, 1964.

16. A. H. Tewfik and H. Wong, On the application of uniform linear array bearing estimation techniques to uniform circular arrays, IEEE Trans. SP-40, pp. 1008-1011, 1992.

17. A. G. Peirsol, Time delay estimation using phase data, IEEE Trans., ASSP-29, pp.471-477, 1981.

18. C. U. Padmini and P. S. Naidu, Circular array and estimation of direction of arrival of a broadband source, Signal Processing, vol. 37, pp. 243-254, 1994.

19. R. J. Kozick, Linear imaging with source array on convex polygonal boundary, IEEE Trans on System, Man and Cybernetics, vol.21, pp. 1155- , 1991.

20. R. J. Kozick and S. A. Kassam, Coarray synthesis with circular and elliptical boundary arrays, IEEE, Trans on Image Processing, vol. 1, pp. 391-404, 1992.

21. R. Williams and Bernard Harris, Passive acoustic sythetic aperture processing techniques, IEEE Jour. of Ocean Eng., vol 17, pp. 8-14, 1992.

22. D. C. Munson, Jr., An introduction to strip-mapping sythetic aperture radar, IEEE, Proc of ICASSP, pp. 2245-2248, 1987.

Chapter Three

Frequency Wavenumber Processing

In signal processing, extraction of a signal buried in noise has been a primary goal of lasting interest. A digital filter is often employed to modify the spectrum of a signal in some prescribed manner. Notably the proposed filter is designed to possess unit response in the spectral region where the desired signal is residing and low response where the undesired signal and noise are residing. This strategy will work only when the spectrum of the desired signal does not overlap (or only partially overlapping) with the spectrum of the undesired signal and noise. This is also true in wavefield processing. Spectrum shaping is necessary whenever the aim is to enhance certain types of wavefields and suppress other types of unwanted wavefields, often termed as noise or interference. Such a selective enhancement is possible on the basis of spectral differences in the frequency wavenumber (ω-k) domain. For example, it is possible, using a digital filter, to enhance the wavefield traveling at a speed different from that of the noise or the interference as their spectra lie on different radial lines in the (ω-k) plane. Other situations where spectral shaping is required are (a) Prediction of wavefields, (b) Interpolation of wavefields, and (c) Extrapolation of wavefield into space where the field could be measured. In this chapter we shall consider the design of pass filters, specially useful in wavefield processing such as a fan and a quadrant filter. When signal and noise are overlapping in the frequency wavenumber domain simple pass filters are inadequate. Optimum filters such as Wiener filters are required. We cover this topic in some depth in view of its importance. Next, we introduce the concept of noise cancellation through prediction. Later, in chapter 6 we shall evaluate the effectiveness of some of the techniques described here.

§3.1 Digital Filters in the ω-k Domain:
In wavefield processing digital filters are often used to remove the interference corrupting the signal of interest (SOI) which arrives from one or more known directions while the interference arrives from different directions, which are often unknown. The purpose of a digital filter is to accept the wave energy arriving from a given direction or an angular sector and to reject all wave energy (interference) lying outside the assumed passband which consists of a vertical cone symmetric about the temporal frequency axis. Such a filter is known as a fan filter in seismic exploration (fig. 3.1) or as a quadrant filter (fig. 3.6) when the axis is tilted. In this section we shall study in detail the fan and the quadrant filters. The issues connected with the design and implementation will be discussed. Finally, we shall also examine the role of sampling and the associated question of aliasing.

3.1.1 <u>Two Dimensional Filters</u>: Two dimensional filters are extensively used in image processing [1], and geophysical map processing [2]. However, because of the fundamental differences between the wavefield and the image signal, the 2D filters used in the respective applications are of different types. For example, in wavefield processing the filter must satisfy the causality condition in time but there is no such requirement in image processing. The causality requirement is waived whenever delay in the output can be tolerated or the data is prerecorded as in seismic exploration or nondestructive testing. Another important difference is that the spectrum of a propagating wavefield is necessarily confined to a conical domain in the frequency wavenumber space (see chapter 1). Only locally generated disturbances (evanescent waves) and instrument generated noise are not subject to the above restriction. Hence, in wavefield applications a typical lowpass filter has a shape of a hand-held fan. Such a filter is known as a fan filter, widely used in seismic data processing. The design and implementation of the fan filter is naturally of great interest in wavefield processing. As for other types of filters, such as lowpass filters with circular or elliptical passbands, a ring type bandpass filter, Laplacian filter for edge detection, deblurring filter, etc., which are widely used in image processing, they are outside the scope of this book.

3.1.2 <u>Fan Filters</u>: The goal is to extract a broadband signal traveling with a horizontal apparent speed between c_{Hor} and ∞ (note $c_{Hor} = \dfrac{c}{\sin\theta}$). The desired fan filter will have the following transfer function:

$$H(s,\omega) = 1 \quad \begin{cases} |\omega| < \omega_{max} \\ s < \dfrac{|\omega|}{c_{Hor}} \quad (s = \sqrt{u^2 + v^2}) \end{cases} \qquad (3.1)$$
$$ = 0 \qquad otherwise$$

which is illustrated in fig. 3.1. The filter will pass all propagating waves whose angle of elevation lies between 0° and $\pm\sin^{-1}(\dfrac{c}{c_{Hor}})$ (the elevation angle is measured with respect to the vertical axis). Fail and Grau [3] were the first to introduce the fan filter in 1963. Independently, Embree [4] came up with a similar filter which he called a "Pie Slice Filter". Since the filter has a circular symmetry in a spatial frequency plane we need a series of concentric uniform circular arrays (in effect, a planar array) to measure the direction of arrival in the horizontal plane and horizontal apparent speed, when the wave speed is known. For the purpose of filter design it is enough if we consider any one radial line.
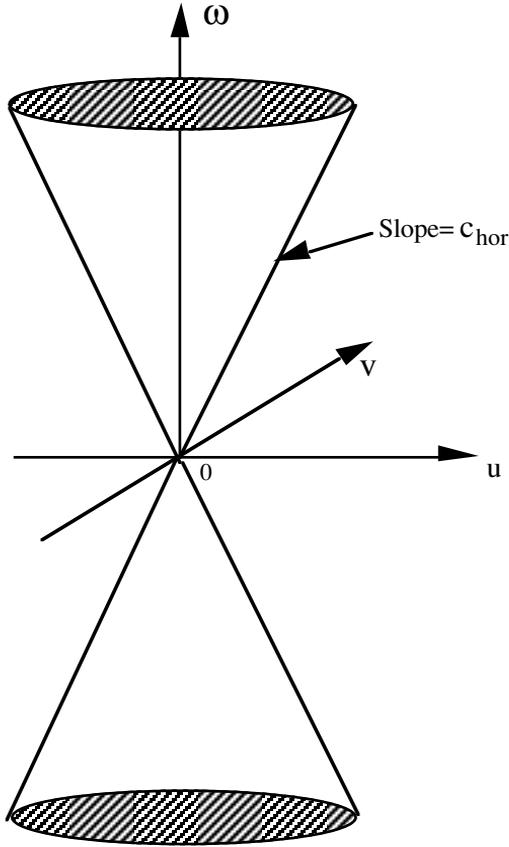
Figure 3.1: Fan filter. The passband is a cone with an axis along the ω-axis and the apex is at the center. All waves whose horizontal apparent speed lies inside the cone are allowed and others are attenuated.

The filter may be rotated about the vertical axis to yield a circularly symmetric filter. For simplicity we consider a single radial line of sensors coinciding with the x-axis and assume that all waves are propagating along the x-axis with different horizontal apparent speeds.

In order to compute the fan filter coefficients we take the inverse Fourier transform of the transfer function given in (3.1),

$$h_{m,n} = \frac{1}{4\pi^2} \int\limits_{-\infty}^{+\infty}\int H(\omega,u)e^{j(\omega n\Delta t + um\Delta x)}d\omega du$$

$$= \frac{1}{4\pi^2} \int\limits_{-\omega_{max}}^{+\omega_{max}} \int\limits_{-\frac{|\omega|}{c_{Hor}}}^{\frac{|\omega|}{c_{Hor}}} e^{j(\omega n\Delta t + um\Delta x)} d\omega du$$

(3.2)

$$= \frac{1}{2\pi^2} \int\limits_{-\omega_{max}}^{+\omega_{max}} \frac{\sin(m\Delta x \frac{|\omega|}{c_{Hor}})}{m\Delta x} e^{j\omega n\Delta t} d\omega$$

where Δt and Δx are temporal and spatial sampling intervals, respectively. Since $\sin(m\Delta x \frac{|\omega|}{c_{Hor}})$ is a symmetric function, the last integral in (3.2) can be simplified as

$$h_{m,n} =$$
$$\frac{1}{2\pi^2 m\Delta x} \int\limits_0^{+\omega_{max}} \left[\sin(\frac{m\Delta x}{c_{Hor}} - n\Delta t)\omega + \sin(\frac{m\Delta x}{c_{Hor}} + n\Delta t)\omega \right] d\omega$$

(3.3)

We are free to choose the spatial and temporal sampling intervals but within the constraints imposed by the sampling theorem. These are discussed in §3.2. Let the temporal sampling interval Δt be so chosen that $\Delta t = \frac{\Delta x}{c_{Hor}}$. In practice Δx is held fixed; therefore, to alter c_{Hor} it is necessary to resample the whole signal with a different sampling interval, consistent with the above choice. Equation (3.3) is easily evaluated to yield

$$h_{m,n} = \frac{1}{2\pi^2 m\Delta x\Delta t} \left[\frac{1 - \cos(m-n)\pi}{(m-n)} + \frac{1 - \cos(m+n)\pi}{(m+n)} \right]$$

(3.4)

Equation (3.4) takes particularly a simple form if we shift the origin to the half-way point between two sensors

$$h_{m,n} = \frac{1}{\pi^2 \Delta x\Delta t} \frac{1}{m^2 - n^2}$$

(3.5)

where $m = \pm\frac{1}{2}, \pm1\frac{1}{2}, \pm2\frac{1}{2}, \ldots$ (see fig. 3.2). Note that since $m \neq n$, $h_{m,n}$
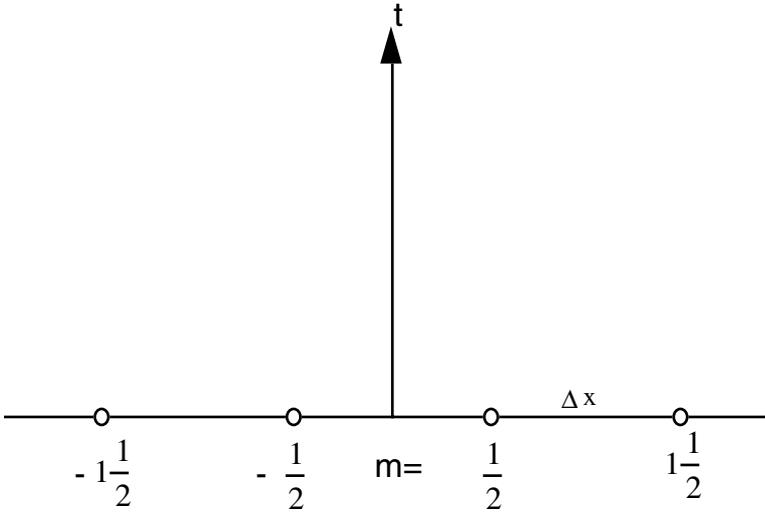
Figure 3.2: Centered array (ULA) with origin lying between two sensors. This is also the physical center of the array.

will always be finite. The frequency response (magnitude) of a fan filter given by (3.5) is shown in fig. 3.3

3.1.3 <u>Fast Algorithm</u>: In order to implement the filter the following convolution sum will have to be evaluated

$$\hat{f}(n\Delta t) = \sum_{n'}\sum_{m} h_{m,n'} f(m\Delta x, (n-n')\Delta t) \qquad (3.6)$$

The computational load for evaluating the convolutional sum can be greatly reduced by exploiting the filter structure [5]. We define a new index $\mu = (1+2m')$ where $m' = 0, \pm 1, \pm 2, \dots$. We note that the new index takes values, $\mu = \pm 1, \pm 3, \pm 5, \dots$, for different values of m'. Equation (3.5) may be reduced to

$$h_{m,n} = \frac{1}{\pi^2}\left(\frac{1}{\dfrac{\mu^2}{4} - n^2}\right)$$

which may be further split into two components
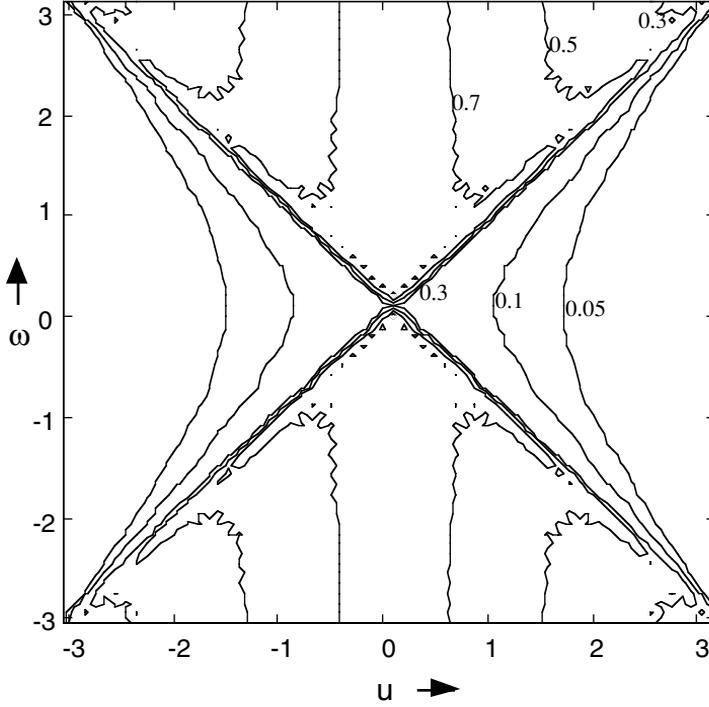
Figure 3.3: Frequency response (magnitude) of a fan filter (64 sensors and 64 time samples).

$$h_{m,n} = \frac{1}{\pi^2 \mu}\left(\frac{1}{\frac{\mu}{2} - n} + \frac{1}{\frac{\mu}{2} + n}\right)$$

$$= \frac{1}{\mu}\left[r_{\mu,n} + q_{\mu,n}\right]$$

The components $r_{\mu,n}$ and $q_{\mu,n}$ possess many symmetry properties as listed below:

• $r_{\mu,n}$ and $q_{\mu,n}$ are antisymmetric in time index (n) about a point $n = \frac{\mu}{2}$ and $n = -\frac{\mu}{2}$, respectively.

• $r_{\mu,n}$ and $q_{\mu,n}$ are of the same shape as $r_{1,n}$ except for the time shift.

$r_{\mu,n} = r_{1,n-\frac{\mu-1}{2}}$ and $q_{\mu,n} = -r_{1,n+\frac{\mu+1}{2}}$.

• $q_{\mu,n} = -r_{\mu,n+\mu}$.

Using the last property we obtain $h_{m,n} = \dfrac{1}{\mu}\left[r_{\mu,n} - r_{\mu,n+\mu}\right]$ and using the second property we obtain

$$h_{m,n} = \frac{1}{\pi^2\mu}\left(r_{1,n-\frac{\mu-1}{2}} - r_{1,n+\frac{\mu+1}{2}}\right) \tag{3.7}$$

Substituting (3.7) in (3.6) we obtain

$$\hat{f}(n\Delta t) = \sum_{n'}\sum_{\mu}\frac{1}{\pi^2\mu}\left(r_{1,n'-\frac{\mu-1}{2}} - r_{1,n'+\frac{\mu+1}{2}}\right)f(\mu\Delta x,(n-n')\Delta t)$$

$$= \frac{1}{\pi^2}\sum_{n'}r_{1,n'}\sum_{\mu=-l}^{l}\frac{1}{\mu}\left[\begin{array}{l}f(\mu\Delta x,(n-n'-\frac{\mu-1}{2})\Delta t)\\[2mm]-f(\mu\Delta x,(n-n'+\frac{\mu+1}{2})\Delta t)\end{array}\right] \tag{3.8}$$

where l=M-1 and M (an even number) stands for the number of sensors. In (3.8) there is only one convolution to be carried out in place of M/2 convolutions in (3.6). The inner summation in (3.8) stands for spatial filtering and the outer summation stands for temporal filtering. Let the output of the spatial filter be $f_1(n\Delta t)$, where

$$f_1(n\Delta t) = \sum_{\mu=-l}^{l}\frac{1}{\mu}\left[\begin{array}{l}f(\mu\Delta x,(n-n'-\frac{\mu-1}{2})\Delta t)\\[2mm]-f(\mu\Delta x,(n-n'+\frac{\mu-1}{2})\Delta t)\end{array}\right]$$

Equation (3.8) takes the form

$$\hat{f}(n\Delta t) = \sum_{n'}r_{1,n'}f_1((n-n')\Delta t) \tag{3.9}$$

Taking z transform on both sides of (3.9) we shall obtain

$$\hat{F}(z) = R_1(z)F_1(z)$$

where $\hat{F}(z)$, $F_1(z)$, $R_1(z)$ are z transforms of $\hat{f}(n), f(n), r_{1,n}$, respectively. In particular,

$$R_1(z) = \sum_{n=-N_1+1}^{N_1} \frac{1}{\left(\dfrac{1}{2} - n\right)} z^n$$

$$= -z \sum_{n=0}^{N_1-1} \frac{1}{\left(\dfrac{1}{2} + n\right)} z^n + \sum_{n=0}^{N_1-1} \frac{1}{\left(\dfrac{1}{2} + n\right)} z^{-n} \qquad (3.10)$$

$$= -z R_{N_1}(z) + R_{N_1}(z^{-1})$$

where

$$R_{N_1}(z) = \sum_{n=0}^{N_1-1} \frac{1}{\left(\dfrac{1}{2} + n\right)} z^n$$

and the number of time samples are assumed to be $2N_1$. Note that $R_{N_1}(z)$ acts on future time and $R_{N_1}(z^{-1})$ acts on past time. We assume that the data are pre-recorded; hence the question of causality is irrelevant. It is shown in [5] that as $N_1 \to \infty$, $R_{N_1}(z)$ may be approximated by a stable pole-zero filter, in particular,

$$R_{N_1}(z) \to \frac{2(1 - 0.65465z)}{1 - 0.98612z + 0.13091z^2} \qquad (3.11)$$

$$N_1 \to \infty$$

A comparison of the filter response N₁=64 and $N_1 \to \infty$ is shown in fig. 3.4. A block diagram for fast realization of the fan is shown in fig. 3.5. The realization consists of two parts, the temporal processor (upper part) and the spatial processor (lower part). The temporal part has two pole-zero filters, and one acting on past time and the other on future time. A pole-zero filter can be implemented recursively; for example, the left arm of the temporal processor may be implemented via a difference equation given below
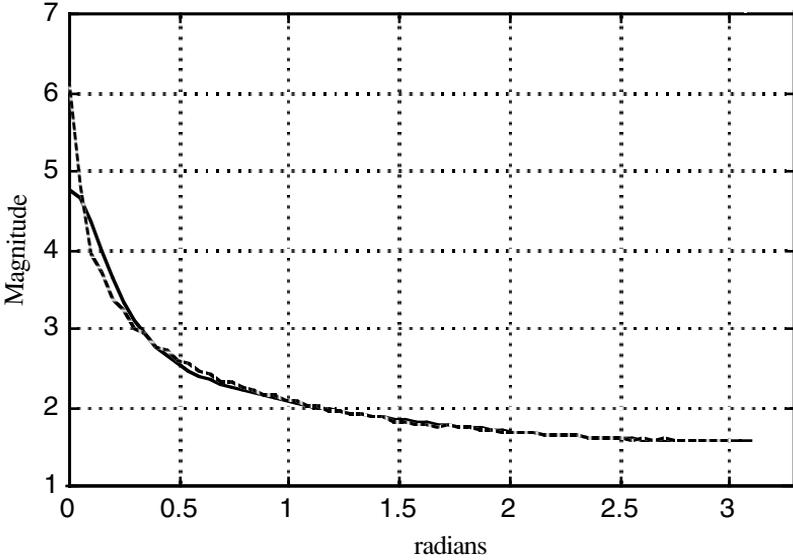
Figure 3.4: A comparison of the filter response N$_1$=64 (dashed line) and $N_1 \rightarrow \infty$ (solid line).

$$\hat{f}_l(n\Delta t) - 0.98612\hat{f}_l((n+1)\Delta t) + 0.13091\hat{f}_l((n+2)\Delta t)$$
$$= 2f_1((n+1)\Delta t) - 1.3093f_1((n+2)\Delta t)$$

and the right arm via

$$\hat{f}_r(n\Delta t) - 0.98612\hat{f}_r((n-1)\Delta t) + 0.13091\hat{f}_r((n-2)\Delta t)$$
$$= 2f_1(n\Delta t) - 1.3093f_1((n-1)\Delta t)$$

The output of the temporal processor is finally given by $\hat{f}(n\Delta t) = \hat{f}_l(n\Delta t) + \hat{f}_r(n\Delta t)$.

3.1.4 <u>Quadrant Filter</u>: The fan filter is a half-plane filter; hence it is insensitive to the direction of propagation as long as the horizontal apparent speed lies within specified limits. A wave traveling with apparent speed c_{Hor} cannot be distinguished from a wave traveling with apparent speed of $-c_{Hor}$. For such a discriminatory property the desired filter must possess a passband in only one quadrant, either the 1st or 2nd quadrant. Note that for the filter to be real the passband must be reflected diagonally into the opposite quadrant. An example of such a passband is shown in fig. 3.6. We will call such a filter a quadrant filter.
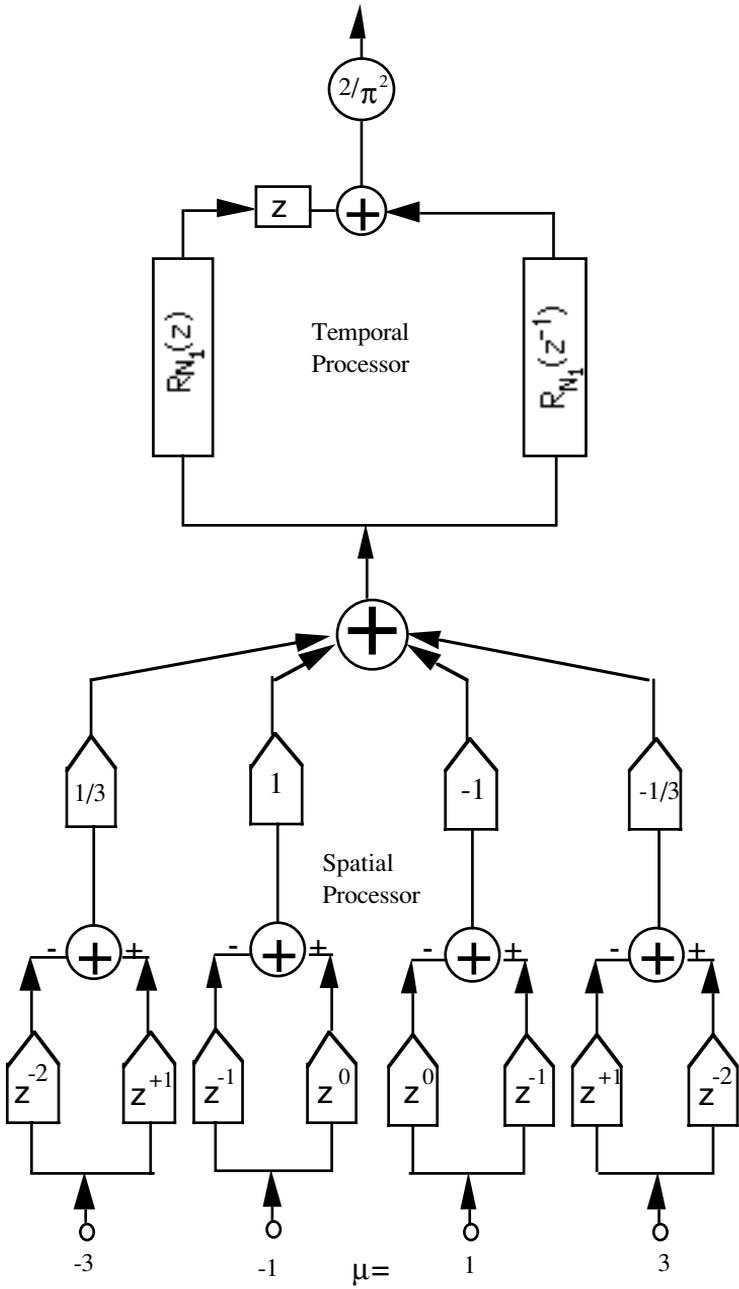
Figure 3.5: A block diagram of fan filter implementation. The filter structure has two parts, namely, spatial and temporal parts, which are independent of each other.
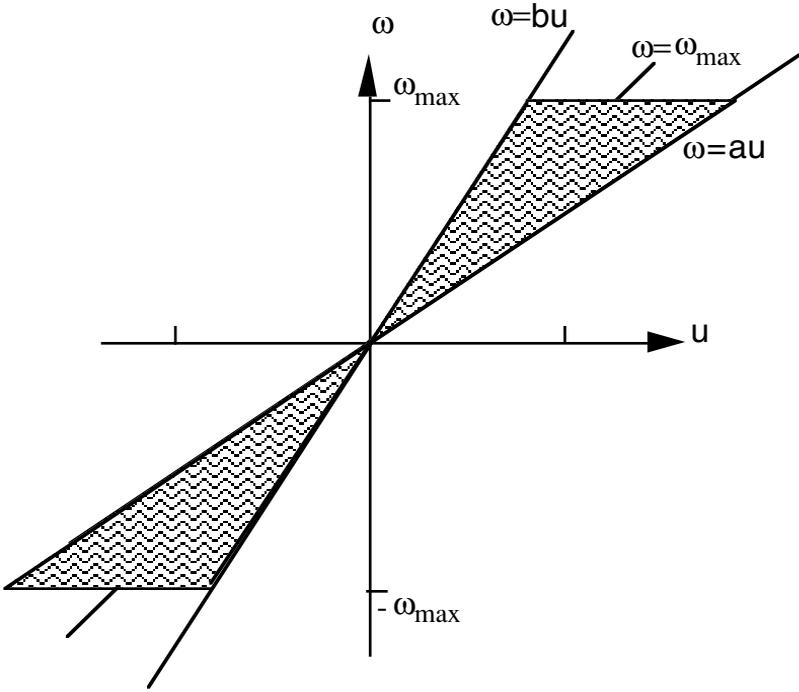
Figure 3.6: Quadrant filter in ω-u plane.

The passband is bounded with two radial lines with slopes a and b respectively. Further the passband is terminated by two horizontal lines $\omega = \pm\omega_{max}$ where ω_{max} is cut-off temporal frequency. When $|b| > 1$ and $|a| < 1$ the passband will be terminated by two cut-off lines as shown in fig. 3.6. Draw a diagonal line and divide the passband into two triangles. Straight forward integration yields the filter impulse response function for $m \neq 0$, $(\dfrac{m}{a}\dfrac{\Delta x}{\Delta t} + n) \neq 0$ and $(\dfrac{m}{b}\dfrac{\Delta x}{\Delta t} + n) \neq 0$ [6].

$$h_{m,n} = \frac{1}{4\pi^2} \int_0^{\omega_{max}} \left[\int_{\frac{\omega}{b}}^{\frac{\omega}{a}} 2\cos(m\Delta x u + n\Delta t \omega)du \right] d\omega$$

$$= \frac{1}{2\pi^2 m\Delta x\Delta t}\left[\frac{\dfrac{(1-\cos(\dfrac{m}{a}\dfrac{\Delta x}{\Delta t}+n)\Delta t\omega_{max})}{\dfrac{m}{a}\dfrac{\Delta x}{\Delta t}+n}}{\;} - \frac{(1-\cos(\dfrac{m}{b}\dfrac{\Delta x}{\Delta t}+n)\Delta t\omega_{max})}{\dfrac{m}{b}\dfrac{\Delta x}{\Delta t}+n}\right] \qquad (3.12a)$$

The special cases are given by

$$h_{0,0} = \frac{(b-a)\omega_{max}^2}{4\pi^2 ab} \qquad\qquad m = n = 0$$

$$h_{m,0} = \frac{(b-a)}{2\pi^2 n^2 \Delta t^2 ab}\left[\begin{array}{l} n\omega_{max}\Delta t\sin(n\omega_{max}\Delta t) \\ +\cos(n\omega_{max}\Delta t)-1 \end{array}\right] \qquad m = 0, n \neq 0$$

$$h_{m,n} = \frac{-1}{2\pi^2}\frac{(1-\cos(\dfrac{m}{b}\dfrac{\Delta x}{\Delta t}+n)\omega_{max}\Delta t)}{n\Delta x\Delta t(\dfrac{m}{b}\dfrac{\Delta x}{\Delta t}+n)} \qquad (\dfrac{m}{a}\dfrac{\Delta x}{\Delta t}+n)=0, n \neq 0,$$

$$(\dfrac{m}{b}\dfrac{\Delta x}{\Delta t}+n)\neq 0$$

$$h_{m,n} = \frac{1}{2\pi^2}\frac{(1-\cos(\dfrac{m}{a}\dfrac{\Delta x}{\Delta t}+n)\omega_{max}\Delta t)}{n\Delta x\Delta t(\dfrac{m}{a}\dfrac{\Delta x}{\Delta t}+n)} \qquad (\dfrac{m}{b}\dfrac{\Delta x}{\Delta t}+n)=0, n \neq 0,$$

$$(\dfrac{m}{a}\dfrac{\Delta x}{\Delta t}+n)\neq 0$$

$$(3.12b)$$

We now set $\Delta x=1$ and $\Delta t=1$, hence $\dfrac{\Delta x}{\Delta t}=1$. Using these settings (3.12a) reduces to equation (6) in [6]. The highest temporal frequency is $f_{hi} = \dfrac{1}{2}$ Hz (or $\omega_{hi} = \pi$ in angular frequency) and the corresponding lowest wavelength,
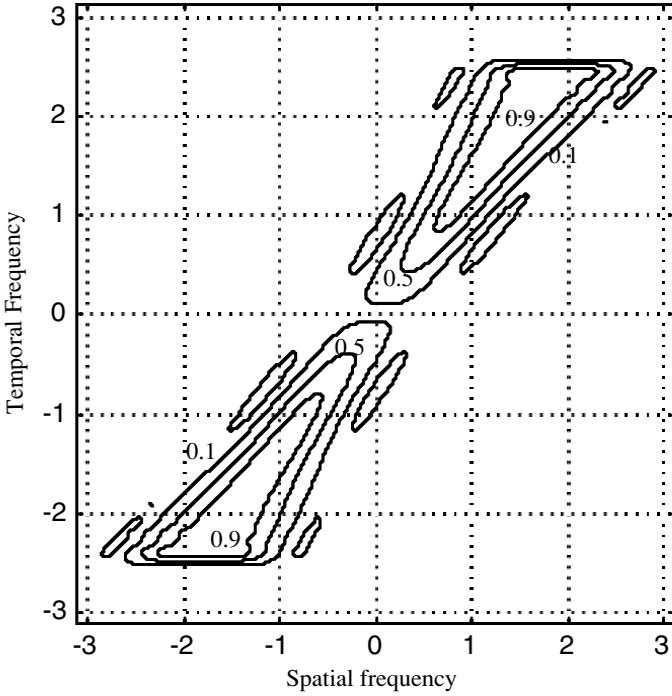
Figure 3.7: Frequency response of a quadrant filter (a=1, b=2). 16 sensors and 64 time samples are assumed. The maximum temporal frequency is 0.8π. The contour values are as shown in the figure.

$\lambda_{lowest} = 2$. The sensor spacing will be one. In real terms, consider a signal with $f_{hi} = 1500\,Hz$ propagating underwater where the wave speed is 1500 m/s. Since the wavelength is 1 meter, the sensor spacing is $\Delta x = 0.5$ meter. The sampling interval is $\Delta t = \dfrac{1}{3}$ ms. Let us redefine units of time and distance. Let one unit of time be $\dfrac{1}{3}$ ms and one unit of distance be 0.5 meter. Here onwards we shall measure the time and the distance in terms of these newly defined units. The response function of a quadrant filter for parameters listed below is shown in fig. 3.7 Parameters: 64 sensors, 64 time samples, $\omega_{max} = 0.8\pi$, a=1 and b=4. In (3.12a) if we let a=-b=1, $c_{Hor} = 1$, $\Delta t = 1$, and

No. of Sensors	rms error	maximum error
4	0.1183	0.9959
8	0.0775	0.9958
16	0.0418	0.9912
32	0.0240	0.9743
64	0.0153	0.9303

Table 3.1: The number of temporal samples is fixed at 64. While the rms error decreases with increasing number of sensors the maximum error shows only marginal decrease.

$\omega_{max} = \pi$ the resulting filter becomes a fan filter with its axis turned 90° and the filter coefficients are

$$h_{m,n} = \frac{1}{2m\pi^2}\left[\frac{(1-\cos(m+n)\pi)}{m+n} + \frac{(1-\cos(m-n)\pi)}{m-n}\right]$$

which indeed are the same as in (3.4).

3.1.5 <u>Weighted Least Squares Filter</u>: The filter size controls the root mean square (rms) difference between the ideal infinite length filter and finite length filter. In general, the rms error decreases with increasing filter size but only slowly as revealed in table 3.1 where we considered a discrete fan filter whose ideal response is one in the region bounded by two rays with slopes 2 and 8 and zero outside. The maximum temporal frequency is set at 78% of the Nyquist frequency. The impulse response function of the quadrant filter is computed by taking the inverse discrete Fourier transform of the ideal quadrant filter embedded in a large matrix (256x256) of zeros. The impulse response function is then truncated to 64 temporal samples and a variable number of spatial samples: 4, 8, 16, 32, and 64. Both root mean square (rms) error and maximum difference between the ideal filter (frequency) response and truncated filter (frequency) response, which is now interpolated to the same grid as the ideal impulse response, are shown in table 3.1.

The response in the reject region, where the ideal response is zero, also decreases with increasing filter size. From a practical point of view, for a given filter size, the response in the pass region must be as close to unity as possible while the response in the reject region or at least in some identified parts of the reject region must be close to zero. This goal can be achieved by using a weighted least squares approach [7]. Let $W(\omega,u)$ be a weighting function which takes a large value in the reject region, where it is desired to have a response close to zero. Elsewhere, $W(\omega,u)=1$. The filter coefficients are obtained by minimizing the following quantity:

$$\frac{1}{4\pi^2} \int\limits_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} \int\limits_{-\frac{\pi}{\Delta t}}^{\frac{\pi}{\Delta t}} \left| W(\omega,u) \left| H(\omega,u) - \sum_m \sum_n h_{mn} e^{-j(\omega n\Delta t - um\Delta x)} \right| \right|^2 d\omega du = \min$$

$$(3.13)$$

Minimization with respect to the filter coefficients leads to the following normal equations:

$$\sum_{m'} \sum_{n'} h_{mn} r_{m-m',n-n'} = g_{mn} \qquad\qquad (3.14a)$$

where

$$r_{m,n} = \frac{1}{4\pi^2} \int\limits_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} \int\limits_{-\frac{\pi}{\Delta t}}^{\frac{\pi}{\Delta t}} W(\omega,u) e^{+j(\omega n\Delta t - um\Delta x)} d\omega du$$

and

$$g_{m,n} = \frac{1}{4\pi^2} \int\limits_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} \int\limits_{-\frac{\pi}{\Delta t}}^{\frac{\pi}{\Delta t}} W(\omega,u) H(\omega,u) e^{+j(\omega n\Delta t - um\Delta x)} d\omega du$$

Note that, when $W(\omega,u)=1$ over the entire plane, (3.14a) reduces to the expected result, namely,

$$h_{m,n} = \frac{1}{4\pi^2} \int\limits_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} \int\limits_{-\frac{\pi}{\Delta t}}^{\frac{\pi}{\Delta t}} H(\omega,u) e^{+j(\omega n\Delta t - um\Delta x)} d\omega du$$

In order to understand the role of the weighting function, consider (3.14a) with finite limits on the summation signs

$$\sum_{m'=0}^{M-1} \sum_{n'=0}^{N-1} h_{mn} r_{m-m',n-n'} = g_{mn} \qquad\qquad (3.14b)$$

and take a finite Fourier transform on both sides of (3.14b). We obtain

$$[H]_{finite}[W]_{finite} = [WH]_{finite} \qquad (3.15)$$

where $[F]_{finite} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_{mn} e^{-j(\omega m + un)}$ stands for the finite discrete Fourier transform (DFT) for $\omega = \dfrac{2\pi}{M}k$ and $u = \dfrac{2\pi}{N}l$. Note that $[F]_{finite} \to F$ as $M \to \infty$ and $N \to \infty$. From (3.15) we obtain

$$[H]_{finite} = \frac{[WH]_{finite}}{[W]_{finite}} \qquad (3.16)$$

Although WH is a bandpass filter (because H is a bandpass filter) $[WH]_{finite}$ will not be zero outside the passband. The presence of W in the denominator will however greatly reduce (if W>>1) the out of band magnitude of $[WH]_{finite}$. Hence, $[H]_{finite}$ on the left hand side of (3.16) will have low magnitude outside the passband. In fig. 3.8 we illustrate this phenomenon. A quadrant filter with unit response in a region bounded by two radial lines with slopes 3 and 8 and with maximum temporal frequency 0.78 times the Nyquist frequency is considered. We selected a weighting function whose value is six in a fan-shaped region bounded by two radial lines with slopes 0.5 and 2 and with the maximum temporal frequency 0.78 times the Nyquist frequency and equal to one elsewhere. The impulse response functions of WH and W are now limited to 64 x16 (64 temporal samples and 16 spatial samples). The finite impulse functions are next used to compute the weighted least squares filter as given in (3.16). The plain (unweighted) finite filter response function along with the region where the weight function is six is shown in fig. 3.8(a) and the weighted least squares filter response function is shown in fig. 3.8(b). Notice how a side lobe falling within the region of the high weight function has been attenuated.

3.1.6 <u>Aliasing Effect in Fan Filter and Quadrant Filter</u>: In array signal processing there is a basic limitation imposed by discrete sensors. Consequently, the wavefield is sampled in the spatial dimension but not in the temporal dimension. It is possible to avoid temporal aliasing through a lowpass filtering and by sampling according to the sampling theorem, but it is a different story with the spatial aliasing which is intimately related to the propagation speed. This phenomenon was briefly discussed in chapter 1 (see p. 22). Here we shall examine the effect of spatial sampling on digital filters for wavefields, in particular, a quadrant filter whose pass region is defined by two radial lines with slopes c_{Lo} and c_{Hi} as shown in fig. 3.9. We further assume that sensor outputs have been prefiltered to confine the spectra to $\pm \omega_{max}$.
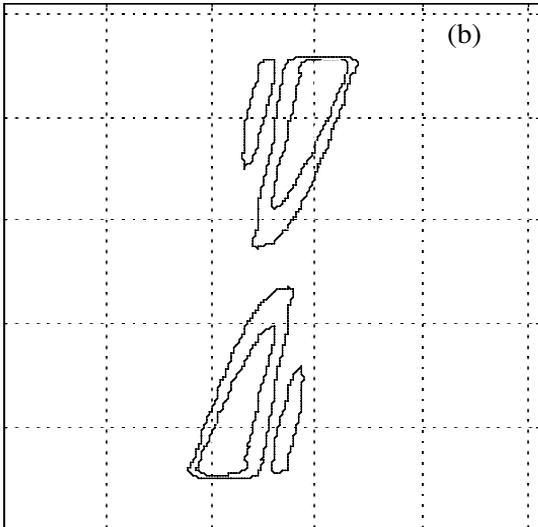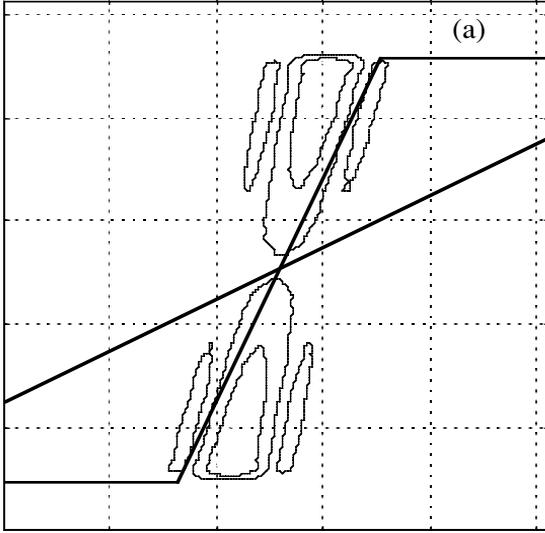
Figure 3.8: The role of a weighting function is shown here. (a) unweighted case. (b) weighted case (equal to six inside a fan shaped region bounded by two radial lines). Notice how a side lobe falling within the region of high weight function has been attenuated.

The sampled version of the quadrant filter will be periodic both in spatial and temporal frequencies but as the sensor outputs have been prefiltered replication in the temporal frequency would not cause any aliasing. However, aliasing can take place due to the replication in spatial frequency. This is shown in fig. 3.9 where we have drawn three replications, that is, three rectangles including the principal rectangle. Notice the intrusion of a radial line with a slope c_{Lo} from the neighboring rectangles into the principal rectangle. Such an intrusion is the cause of aliasing. Clearly, the aliasing can be prevented if

$$c_{Lo} \geq \frac{\omega_{max}}{B_0} = \frac{\Delta x}{\Delta t} \qquad (3.17)$$

A similar requirement was shown (in chapter 1, p. 23) to be necessary to avoid aliasing of a broadband signal traveling with a horizontal speed c_{Hor}.

§3.2 **Mapping of 1D into 2D Filters:**
As in a one dimensional filter, a sharp transition from passband to rejectband in a two dimensional filter (in frequency wavenumber space) results in large ripples in passband and rejectband. To overcome this drawback, a sharp transition is replaced by a smooth transition, but this will degrade the quality of the filter. Thus, in the design of a pass filter there is a contradictory requirement, calling for an optimum design strategy of minimum transition width while maintaining the ripple height below a prescribed limit. This optimization problem has been largely solved in one dimension [8] but in two or higher dimensions the procedure of optimization becomes highly cumbersome and computationally intensive. It has been found that optimally designed 1D filters may be mapped into 2D filters having similar transition and ripple height properties as those of 1D filters [9]. The 2D filters of interest in the wavefield analysis are fan filter and quadrant filter. The fan filter may be obtained by transforming an optimally designed 1D filter.
3.2.1 McClellan's Transformation: Consider a real linear phase FIR (finite impulse response) filter $h_0, h_{\pm 1}, h_{\pm 2} \ldots h_{\pm N}$ where for linear phase we must have $h_k = h_{-k}$. The transfer function of the filter is given by

$$H(\omega') = h_0 + 2\sum_{k=1}^{N} h_k \cos(\omega' k) \qquad (3.18)$$

which may be expressed as a polynomial in $\cos(\omega')$,
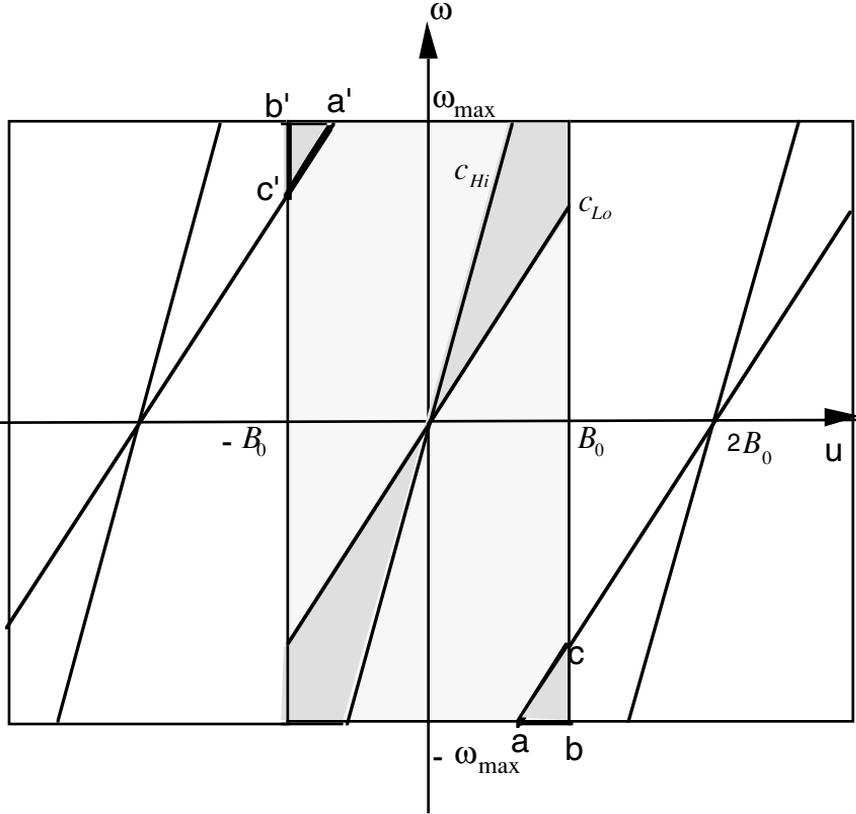
Figure 3.9: The aliasing is caused by intrusion of a radial line with slope c_{Lo} from the neighboring rectangles into the principal rectangle. Triangles abc and a'b'c' are the aliased pass regions. Aliasing can be prevented if we were to sample an analog quadrant filter according to (3.17).

$$H(\omega') = \sum_{k=0}^{N} b_k \cos^k(\omega') \qquad (3.19)$$

where the coefficients b_k, k=0, 1, ... N are expressed in terms of the FIR filter coefficients. A point on the frequency axis is mapped onto a closed contour in the frequency wavenumber space using a linear transformation

$$\cos(\omega') = \sum_{p=0}^{P} \sum_{q=0}^{Q} t_{pq} \cos(p\omega) \cos(qu) \qquad (3.20a)$$
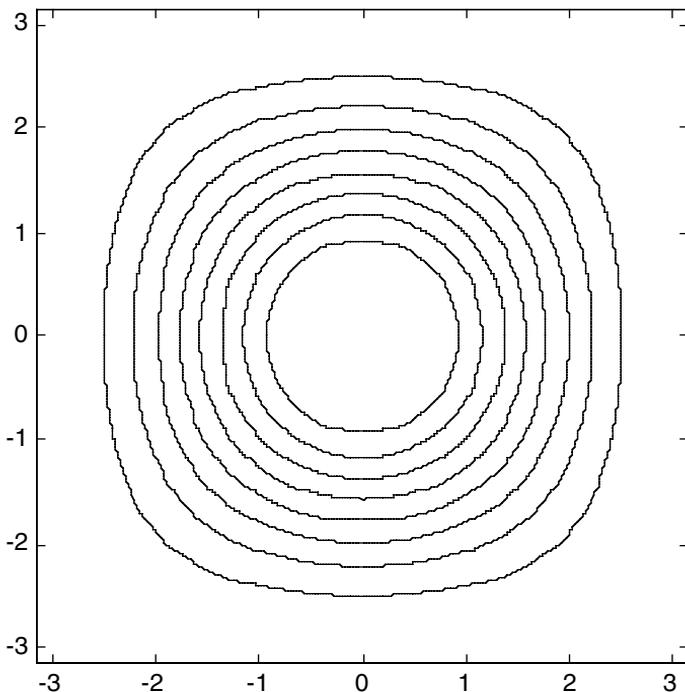
Figure 3.10: Mapping produced by (3.21). The inner contours corresponding to low frequency are close to circular shape but the outer contours corresponding to high frequency are only approximately circular.

or

$$\sin^2(\frac{\omega'}{2}) = \sum_{p=0}^{P}\sum_{q=0}^{Q} t'_{pq} \sin^{2p}(\frac{\omega}{2})\sin^{2q}(\frac{u}{2}) \qquad (3.20b)$$

where t_{pq} and t'_{pq} are yet to be determined coefficients. The shape of the contour depends upon the coefficients t_{pq} or t'_{pq}. For example, we obtain an approximate circular contour (see fig. 3.10) for

$$t_{00} = -\frac{1}{2}$$

$$t_{01} = t_{10} = t_{11} = \frac{1}{2}$$

The mapping relation (3.20a) takes a form

$$\cos(\omega') = \frac{1}{2}\{-1 + \cos(u) + \cos(\omega) + \cos(u)\cos(\omega)\} \qquad (3.21)$$

Contours corresponding to a set of fixed values of $\cos(\omega')$ are shown in fig. 3.10. Circular approximation is better in the low frequency range. Using (3.20a) in (3.19) we obtain a 2D filter transfer function,

$$H(\omega,u) = \sum_{k=0}^{N} b_k \left[\sum_{p=0}^{P} \sum_{q=0}^{Q} t_{pq} \cos(p\omega)\cos(qu) \right]^k \qquad (3.22)$$

The mapping function (3.20a) must satisfy the following conditions:

a) When $\omega'=0$, $\omega = u = 0$. This requires $\sum_{p=0}^{P} \sum_{q=0}^{Q} t_{pq} = 1$. The condition ensures that a lowpass filter remains a lowpass filter even after the transformation, that is, H(0,0)=H(0).

b) When $\omega'=\pi$, $\omega = u = \pi$. This requires $\sum_{p=0}^{P} \sum_{q=0}^{Q} t_{pq}(-1)^{p+q} = -1$. The condition ensures that a highpass filter remains a highpass filter after transformation, that is, H(π,π)=H(π).

c) When $\omega=0$, $H(0,u) = H(u)$. This requires $\sum_{p=0}^{P} t_{p1} = 1$ and $\sum_{p=0}^{P} t_{pq} = 0$ for $q \neq 1$.

d) When $u=0$, $H(\omega,0) = H(\omega)$. This requires $\sum_{q=0}^{Q} t_{1q} = 1$ and $\sum_{q=0}^{Q} t_{pq} = 0$ for $p \neq 1$.

For circularly symmetric mapping, conditions (b) and (c) must hold good. For P=Q=1, a general solution for the transformation coefficients satisfying the above conditions is given by $t_{00} = -\frac{1}{2}$, $t_{01} = t_{10} = t_{11} = \frac{1}{2}$. By relaxing (b), the solution is given by

$$t_{00} = -a, \; t_{01} = t_{10} = a, \; t_{11} = 1 - a \qquad (3.23)$$

having one free constant which may be optimally chosen for the best fit.

The next important question is to determine the coefficients b_k, $k=0,1,2,...,$ N in (3.19) given the coefficients of a digital filter. Let $h_0, h_1, h_2, ... h_N$ be the given 1D filter coefficients.

$$
\begin{aligned}
H(\omega') &= \sum_{k=0}^{N} \tilde{h}_k \cos(k\omega') \\
&= \sum_{k=0}^{N} \tilde{h}_k T_k(\cos(\omega'))
\end{aligned}
\qquad (3.24)
$$

where $\tilde{h}_0 = h_0$, $\tilde{h}_k = 2h_k, k = 1,...,N$, and $T_n(\cos(\omega'))$ is a Chebyshev polynomial which may be expressed as a polynomial in its argument,

$$T_k(\cos(\omega')) = \sum_{m=0}^{k} c_m^k \cos^m(\omega') \qquad (3.25)$$

Using (3.25) in (3.24) and comparing with (3.19) we obtain the following system of equations:

$$b_0 = \sum_{n=0}^{N} c_0^n \tilde{h}_n, \; b_1 = \sum_{n=1}^{N} c_1^n \tilde{h}_n, \; ... \; b_k = \sum_{n=k}^{N} c_k^n \tilde{h}_n, \; ... \; b_N = c_N^N \tilde{h}_N$$

The coefficients c_k^n are listed in [10, p. 795]. Finally, the 2D filter transfer function for the case P=Q=1 can be written as

$$H(\omega,u) = \sum_{k=0}^{N} b_k \left[\begin{array}{l} \{t_{00} + t_{01}\cos(u) + \\ t_{10}\cos(\omega) + t_{11}\cos(u)\cos(\omega)\} \end{array} \right]^k \qquad (3.26a)$$

Let $F(\omega,u)$ be the input and $G(\omega,u)$ be the output of a 2D filter,

$$
\begin{aligned}
G(\omega,u) &= \sum_{k=0}^{N} b_k \left[\begin{array}{l} t_{00} + t_{01}\cos(u) + t_{10}\cos(\omega) \\ +t_{11}\cos(u)\cos(\omega) \end{array} \right]^k F(\omega,u) \\
&= \sum_{k=0}^{N} b_k [H_0(\omega,u)]^k F(\omega,u)
\end{aligned}
\qquad (3.26b)
$$

which may be written in a recursive fashion. Let

$$F_0(\omega,u) = F(\omega,u)$$
$$F_1(\omega,u) = \left[H_0(\omega,u)\right]F_0(\omega,u)$$
$$F_2(\omega,u) = \left[H_0(\omega,u)\right]F_1(\omega,u)$$
$$\dots$$
$$\dots \qquad\qquad\qquad (3.26c)$$
$$\dots$$
$$F_N(\omega,u) = \left[H_0(\omega,u)\right]F_{N-1}(\omega,u)$$

where $H_0(\omega,u) = t_{00} + t_{01}\cos(u) + t_{10}\cos(\omega) + t_{11}\cos(\omega)\cos(u)$. Note that $F_0(\omega,u), F_1(\omega,u), \dots, F_N(\omega,u)$ do not depend upon the filter coefficients but only on mapping coefficients and the input. Using (3.26c), the filter output may be written as

$$G(\omega,u) = \sum_{k=0}^{N} b_k F_k(\omega,u) \qquad\qquad (3.27)$$

In filter implementation it is possible to generate $F_0(\omega,u), F_1(\omega,u), \dots,$ $F_N(\omega,u)$ in a recursive manner and then combine them after weighting each with b_k. The filter structure is illustrated in fig. 3.11.

3.2.2 <u>Fan Filter</u>: To get fan-shaped contours the required mapping coefficients are

$$t_{00} = t_{11} = 0 \ and \ t_{01} = -t_{10} = \frac{1}{2} \qquad\qquad (3.28a)$$

and the mapping function is given by [11]

$$\cos(\omega') = \frac{1}{2}[\cos(\omega) - \cos(u)] \qquad\qquad (3.28b)$$

The contours generated by (3.28b) are shown in fig. 3.12. The zero valued diagonal contour corresponds to $\omega' = \pm\dfrac{\pi}{2}$. When $\omega' = 0, \ \omega = 0$ and $u = \pm\pi$; and when $\omega' = \pm\pi, \ \omega = \pm\pi$ and $u = 0$. The positive contours
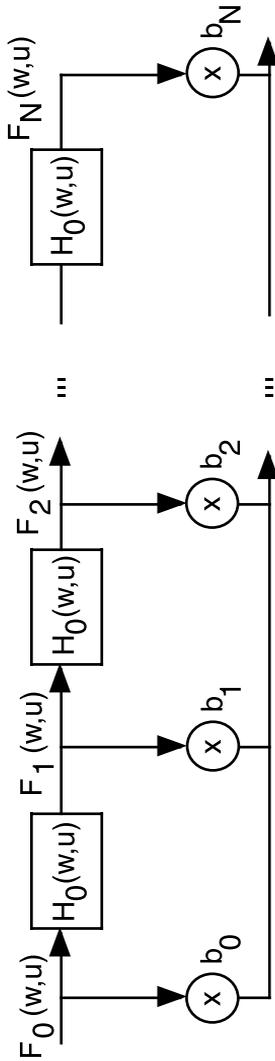
Figure 3.11: A recursive implementation of a 2D filter obtained by transforming a 1D filter. $H_0(\omega, u)$ solely depends upon the mapping coefficients while the coefficients and b_k's depend on 1D filter coefficients.

correspond to $0 \leq |\omega'| \leq \dfrac{\pi}{2}$ and the negative contours to $\dfrac{\pi}{2} < |\omega'| \leq \pi$.

Ideally, a lowpass filter having a unit response in the range $\pm\dfrac{\pi}{2}$ and zero outside
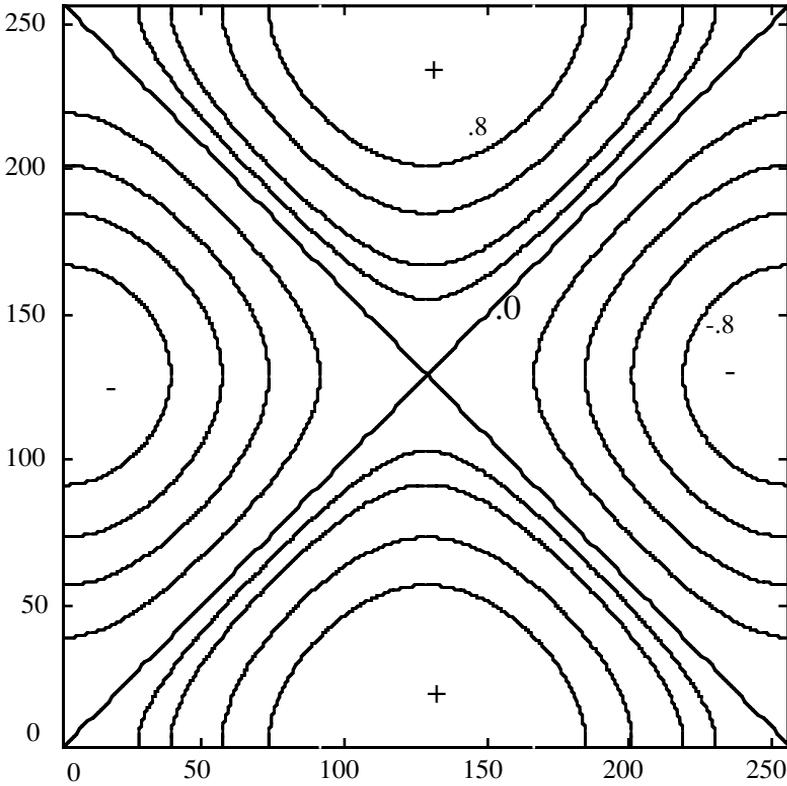
Figure 3.12: Contours generated by a mapping function given by (3.28a). The two diagonals make up the zero contour which divides the display area into four triangles. The positive contours correspond to $0 \leq |\omega'| \leq \dfrac{\pi}{2}$ and the negative contours to $\dfrac{\pi}{2} < |\omega'| \leq \pi$.

when mapped into two dimensions using the mapping function (3.28b) will result in a fan filter with a unit response in the top and bottom triangles and a zero response in the left and right triangles. The value of the 2D filter transfer function may be found by first computing the frequency corresponding to the index of each contour, that is, $\omega' = \cos^{-1}$(contour index) and then evaluating the 1D filter response at the desired frequency.

It is interesting to note that the mapping coefficients given by (3.28a) do not satisfy any of the conditions listed on page 167. This is expected as the

mapping function does not satisfy circular symmetry nor is it necessary to map a lowpass filter into a lowpass filter or a highpass filter into a highpass filter. A fan filter in 3D can also be obtained by transformation of a 1D filter. The transformation function is given by [12]

$$\cos(\omega') = \frac{1}{2}\cos(\omega) - \frac{1}{2}\cos(u) - \frac{1}{2}\cos(v) + \frac{1}{2} \qquad (3.29)$$

§3.3 Multichannel Wiener Filters:

In Wiener filtering the goal is to make the filter output as close as possible, in the least square sense, to a desired signal. The sensor output and the desired signal are assumed to be stationary stochastic signals which are characterized through their covariance functions. The Wiener filters are known after Norbert Wiener who did pioneering work on the prediction of a trajectory of a moving object from its past observations [13]. A recursive algorithm for the solution of a discrete version of the Wiener-Hopf equation was developed by Levinson [14] and by Durbin [15] in the context of time series model fitting. Multichannel extension was made by Wiggins and Robinson [16]. In this section we briefly describe the Wiener filter as applied to array signals where we like to extract a signal traveling in some *known* direction and to optimally suppress all other propagating waves and noise. Here, a straightforward solution of the Wiener-Hopf equation requires invertion of a large block toeplitz covariance matrix, leading to a dramatic increase in the computational load over the single time series version. Therefore, it is worthwhile to spend some effort to understand the principles of the Levinson-Durbin recursive algorithm for the multichannel Wiener filter [16].

3.3.1 Planar Array: We consider a planar array of sensors, not necessarily uniformly distributed (see fig. 3.13). Let $f_p(t)$ be the output of the pth sensor located at (x_p, y_p). We shall model the output as a sum of two random processes, namely, a desired signal, $\xi_p(t)$, and unwanted noise, $\eta_p(t)$,

$$f_p(t) = \xi_p(t) + \eta_p(t), \qquad p = 0,1,...P-1 \qquad (3.30)$$

A multichannel filter $h_p(t), \quad p = 0,1,...P-1$ is sought such that the output, as given by (3.31), is closest to the signal at one of the sensors, for example, at p=0. The filter output, given by,

$$\hat{f}(t) = \sum_{p=0}^{P-1} \int_0^\infty h_p(\tau) f_p(t-\tau) d\tau \qquad (3.31)$$

$x_{P-1}y_{P-1}$

x_1y_1

x_2y_2

x_0y_0
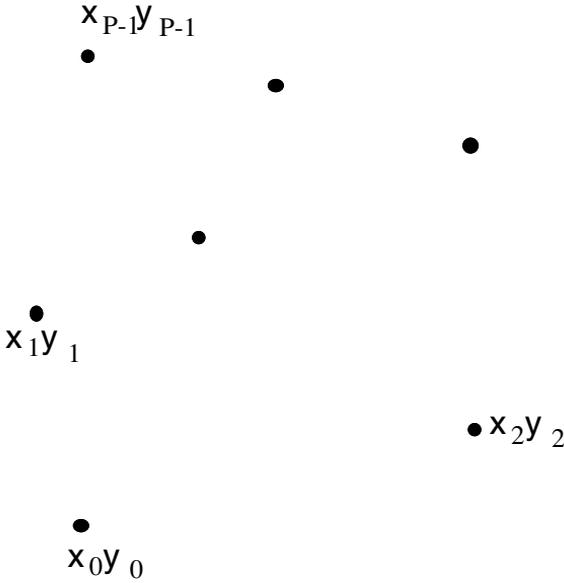
Figure 3.13: A distribution of sensors in a planar array. All sensors are identical but for their position.

must be closest to $\xi_0(t)$ in the sense

$$E\left\{\left|\hat{f}(t) - \xi_0(t)\right|^2\right\} = \min$$

that is, the mean square error (mse) is minimum. This requires minimization of an expression for mse,

$$mse = c_{\xi_0}(0) - 2\sum_{p=0}^{P-1} \int_0^\infty h_p(\tau)c_{\xi_0 f_p}(\tau)d\tau$$

$$+ \sum_{p=0}^{P-1}\sum_{q=0}^{P-1}\int_0^\infty \int_0^\infty h_p(\tau)h_q(\tau')\left[c_{\xi_p \xi_q}(\tau - \tau') + c_{\eta_p \eta_q}(\tau - \tau')\right]d\tau d\tau'$$

$$(3.32)$$

We shall minimize (3.32) with respect to $h_p(\tau)$. So we need to differentiate (3.32) with respect to $h_p(\tau)$ which occurs inside an integral. To see how such

a differentiation may be carried out let us express the integral as a limiting sum. For example,

$$\int_0^\infty h_p(\tau)c_{\xi_0\xi_p}(t-\tau)d\tau \to \Delta\tau\sum_{n=0}^\infty h_p(n\Delta\tau)c_{\xi_0\xi_p}(t-n\Delta\tau) \tag{3.33}$$

$$\Delta\tau \to 0$$

Now, differentiate the sum on the right hand side in (3.33) with respect to $h_p(n\Delta\tau)$ for fixed n. The result is $\Delta\tau c_{\xi_0\xi_p}(t-n\Delta\tau)$. A similar approach is used to differentiate an expression involving a double integral. The derivative is first set to zero and then a limiting operation ($\Delta\tau \to 0$) is carried out. We finally obtain a set of equations known as normal equations:

$$\sum_{q=0}^{P-1}\int_0^\infty h_q(\tau')c_{f_pf_q}(\tau-\tau')d\tau' = c_{\xi_0f_p}(\tau), \qquad p=0,1,...P-1 \tag{3.34a}$$

The minimum mean square error (mmse) may be derived by using (3.34a) in (3.32). We obtain

$$mse\big|_{\min} = c_{\xi_0}(0) - \sum_{p=0}^{P-1}\int_0^\infty h_p(\tau)c_{\xi_0f_p}(\tau)d\tau \tag{3.34b}$$

Let us express (3.34) in discrete form. In order to do this we introduce the following vectors and matrix:

$$\mathbf{c}_{f_pf_q} = \begin{bmatrix} c_{f_pf_q}(0) & c_{f_pf_q}(\Delta\tau) & ... & c_{f_pf_q}((N-1)\Delta\tau) \\ c_{f_pf_q}(-\Delta\tau) & c_{f_pf_q}(0) & ... & c_{f_pf_q}((N-1)\Delta\tau) \\ & & ... & \\ & & ... & \\ & & ... & \\ c_{f_pf_q}((1-N)\Delta\tau) & c_{f_pf_q}((2-N)\Delta\tau), & ... & c_{f_pf_q}(0) \end{bmatrix}$$

$$\mathbf{h}_p = \left[h_p(0),\, h_p(\Delta\tau),\, h_p(2\Delta\tau),\, ... \, h_p((N-1)\Delta\tau)\right]^T$$

and

$$\mathbf{c}_{\xi_0 f_p} = \left[c_{\xi_0 f_p}(0), c_{\xi_0 f_p}(\Delta\tau), c_{\xi_0 f_p}(2\Delta\tau), \ldots c_{\xi_0 f_p}((N-1)\Delta\tau) \right]^T$$

where $\Delta\tau$ is the sampling interval and $\left[\ \right]^T$ stands for matrix or vector transpose. Equation (3.34) may now be expressed using the above vectors and matrices

$$\sum_{q=0}^{P-1} \mathbf{c}_{f_p f_q} \, \mathbf{h}_q = \mathbf{c}_{\xi_0 f_p}, \qquad p = 0,1,2,\ldots P-1 \qquad (3.35a)$$

$$mse\Big|_{\min} = c_{\xi_0}(0) - \sum_{p=0}^{P-1} \mathbf{c}_{\xi_0 f_p}^T \, \mathbf{h}_p \qquad (3.35b)$$

An alternate representation of (3.34) is through block matrices defined as

$$\mathbf{h} = \left[\mathbf{h}(0), \mathbf{h}(\Delta\tau), \ldots \mathbf{h}((N-1)\Delta\tau) \ \right]^T$$

where

$$\mathbf{h}(n\Delta\tau) = \left[h_0(n\Delta\tau), h_1(n\Delta\tau), \ldots h_{P-1}(n\Delta\tau) \right]^T$$

$$\mathbf{C}_0 = \left[\mathbf{c}_0(0), \mathbf{c}_0(\Delta\tau), \ldots \mathbf{c}_0((N-1)\Delta\tau) \right]^T$$

where

$$\mathbf{c}_0(n\Delta\tau) = \left[c_{\xi_0 f_0}(n\Delta\tau), c_{\xi_0 f_1}(n\Delta\tau), c_{\xi_0 f_2}(n\Delta\tau), \ldots c_{\xi_0 f_{P-1}}(n\Delta\tau) \right]^T$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{c}(0), & \mathbf{c}(\Delta\tau), & \ldots & \mathbf{c}((N-1)\Delta\tau) \\ \mathbf{c}(-\Delta\tau), & \mathbf{c}(0), & \ldots & \mathbf{c}((N-2)\Delta\tau) \\ & & \ldots & \\ & & \ldots & \\ \mathbf{c}((1-N)\Delta\tau), & & \ldots & \mathbf{c}(0) \end{bmatrix}$$
$$(PN \text{x} PN)$$

where each element is a block matrix of the type shown below

$$\mathbf{c}(n\Delta\tau) = \begin{bmatrix} c_{f_0f_0}(n\Delta\tau) & c_{f_0f_1}(n\Delta\tau) & \dots & c_{f_0f_{P-1}}(n\Delta\tau) \\ c_{f_1f_0}(n\Delta\tau) & c_{f_1f_1}(n\Delta\tau) & \dots & c_{f_1f_{P-1}}(n\Delta\tau) \\ & & \dots & \\ & & \dots & \\ & & \dots & \\ c_{f_{P-1}f_0}(n\Delta\tau) & c_{f_{P-1}f_0}(n\Delta\tau),\dots & c_{f_{P-1}f_{P-1}}(n\Delta\tau) \end{bmatrix}$$
$$(\text{PxP})$$

As an example consider a three sensor array with two time samples (N=2). The above quantities, $\mathbf{C}$, $\mathbf{h}$, and $\mathbf{C}_0$, become

$$\mathbf{C} = \begin{bmatrix} c_{f_0f_0}(0) & c_{f_0f_1}(0) & c_{f_0f_2}(0) & c_{f_0f_0}(1) & c_{f_0f_1}(1) & c_{f_0f_2}(1) \\ c_{f_1f_0}(0) & c_{f_1f_1}(0) & c_{f_1f_2}(0) & c_{f_1f_0}(1) & c_{f_1f_1}(1) & c_{f_1f_2}(1) \\ c_{f_2f_0}(0) & c_{f_2f_1}(0) & c_{f_2f_2}(0) & c_{f_2f_0}(1) & c_{f_2f_1}(1) & c_{f_2f_2}(1) \\ c_{f_0f_0}(-1) & c_{f_0f_1}(-1) & c_{f_0f_2}(-1) & c_{f_0f_0}(0) & c_{f_0f_1}(0) & c_{f_0f_2}(0) \\ c_{f_1f_0}(-1) & c_{f_1f_1}(-1) & c_{f_1f_2}(-1) & c_{f_1f_0}(0) & c_{f_1f_1}(0) & c_{f_1f_2}(0) \\ c_{f_2f_0}(-1) & c_{f_2f_1}(-1) & c_{f_2f_2}(-1) & c_{f_2f_0}(0) & c_{f_2f_1}(0) & c_{f_2f_2}(0) \end{bmatrix}$$

$$\mathbf{h} = \left[h_0(0), h_1(0), h_2(0), h_0(1), h_1(1), h_2(1) \right]^T$$

and

$$\mathbf{C}_0 = \left[c_{\xi_0f_0}(0), c_{\xi_0f_1}(0), c_{\xi_0f_2}(0), c_{\xi_0f_0}(1), c_{\xi_0f_1}(1), c_{\xi_0f_2}(1) \right]^T$$

Equation (3.35) may be expressed in a compact form as

$$\mathbf{C}\,\mathbf{h} = \mathbf{C}_0 \tag{3.36a}$$

$$mse\big|_{\min} = c_{\xi_0}(0) - \mathbf{C}_0^T\mathbf{h} \tag{3.36b}$$

3.3.2 <u>Frequency Domain</u>: Sometimes it is advantageous to express the normal equations (3.34) in the frequency domain. Taking the Fourier transform on both sides of (3.34) we obtain the following result:

$$\sum_{q=1}^{P-1} H_q(\omega) S_{f_p f_q}(\omega) = S_{\xi_0 f_p}(\omega), \qquad p = 0,1,...P-1 \qquad (3.37a)$$

where $H_q(\omega)$ is the transfer function of the qth filter given by

$$H_q(\omega) = \int_0^\infty h_q(\tau) e^{j\omega\tau} d\tau$$

$S_{f_p f_q}(\omega)$ is the cross-spectrum between $f_q(t)$ and $f_p(t)$ and similarly $S_{\xi_0 f_p}(\omega)$ is the cross-spectrum between $\xi_0(t)$ and $f_p(t)$. Similarly, minimum mean square error (mmse) in the frequency domain may be obtained from (3.34b)

$$mse\big|_{\min} = \int_{-\infty}^{\infty} \left\{ S_{\xi_0}(\omega) - \sum_{p=0}^{P-1} H_p(\omega) S_{\xi_0 f_p}(-\omega) \right\} d\omega \qquad (3.37b)$$

To write (3.37) in a matrix form define the following vectors and matrix:

$$\mathbf{S}_f(\omega) = \begin{bmatrix} S_{0,0}(\omega) & S_{0,1}(\omega) & ... & S_{0,P-1}(\omega) \\ S_{1,0}(\omega) & S_{1,1}(\omega) & ... & S_{1,P-1}(\omega) \\ & & ... & \\ & & ... & \\ & & ... & \\ S_{P-1,0}(\omega) & S_{P-1,1}(\omega) & ... & S_{P-1,P-1}(\omega) \end{bmatrix}$$

$$\mathbf{H}(\omega) = \left[H_0(\omega), H_1(\omega) ... H_{P-1}(\omega) \right]^T$$

and

$$\mathbf{S}_0(\omega) = \left[S_{00}(\omega), S_{01}(\omega) ... S_{0P-1}(\omega) \right]^T$$

The normal equations in the frequency domain (3.37) may be expressed in a compact form

$$\mathbf{S}_f(\omega)\mathbf{H}(\omega) = \mathbf{S}_0(\omega) \tag{3.38a}$$

for all ω in the range $\pm\infty$. Formally, the solution of (3.38) may be expressed as

$$\mathbf{H}(\omega) = \mathbf{S}_f^{-1}(\omega)\mathbf{S}_0(\omega) \tag{3.38b}$$

Now consider a plane wave sweeping across an array of sensors. Let the background noise be spatially white. The spectral matrix for this model is given by

$$\mathbf{S}(\omega) = \mathbf{A}(\omega)S_\eta(\omega) \tag{3.39}$$

where

$$\mathbf{A}(\omega) = \begin{bmatrix} 1+T(\omega) & T(\omega)e^{j(u_0 x_{1,0}+v_0 y_{1,0})} & \cdots & T(\omega)e^{j(u_0 x_{P-1,0}+v_0 y_{P-1,0})} \\ T(\omega)e^{-j(u_0 x_{1,0}+v_0 y_{1,0})} & 1+T(\omega) & \cdots & T(\omega)e^{j(u_0 x_{P-1,1}+v_0 y_{P-1,1})} \\ & & \cdots & \\ & & \cdots & \\ & & \cdots & \\ T(\omega)e^{-j(u_0 x_{P-1,0}+v_0 y_{P-1,0})} & T(\omega)e^{-j(u_0 x_{P-1,1}+v_0 y_{P-1,1})} & \cdots & 1+T(\omega) \end{bmatrix}$$

$T(\omega) = \dfrac{S_0(\omega)}{S_\eta(\omega)}$, $x_{pq} = x_p - x_q$, $y_{pq} = y_p - y_q$, $S_0(\omega)$ is the signal

spectrum, and $S_\eta(\omega)$ is the noise spectrum common to all sensors. Similarly, the vector on the right hand side of (3.38) may be expressed as

$$\mathbf{S}_0(\omega) = \left[1,\, e^{-j(u_0 x_{1,0}+v_0 y_{1,0})},\, e^{-j(u_0 x_{2,0}+v_0 y_{2,0})},\, \ldots,\, e^{-j(u_0 x_{P-1,0}+v_0 y_{P-1,0})}\right] S_0(\omega)$$

$$= \mathbf{B}(\omega)S_0(\omega) \tag{3.40}$$

Using (3.39) and (3.40) in (3.38) we obtain special normal equations

$$\mathbf{A}(\omega)\mathbf{H}(\omega) = \mathbf{B}(\omega)\frac{S_0(\omega)}{S_\eta(\omega)} \tag{3.41a}$$

$$= \mathbf{B}(\omega)T(\omega)$$

or

$$H(\omega) = A^{-1}(\omega)B(\omega)T(\omega) \qquad (3.41b)$$

We note that $A(\omega)$ has an useful structure in that it can be expressed as

$$A(\omega) = I + T(\omega)a(\omega)a^H(\omega) \qquad (3.42a)$$

where

$$a(\omega) = \left[e^{j(u_0 x_0 + v_0 y_0)}, e^{j(u_0 x_1 + v_0 y_1)}, e^{j(u_0 x_2 + v_0 y_2)}, \ldots e^{j(u_0 x_{P-1} + v_0 y_{P-1})} \right]^T$$

$$(3.42b)$$

3.3.3 <u>Constrained Minimization</u>: Let the filtered signal output of an array of sensors be given by

$$\hat{f}(t) = \sum_{p=0}^{P-1} \int_0^\infty h_p(\tau)\xi_p(t-\tau)d\tau \qquad (3.43)$$

Assume a signal model where a plane wave is sweeping across the array maintaining its waveform unchanged. The outputs of any two sensors differ only in propagation delays. The signal at the pth sensor may be given by $\xi_p(t) = \xi_0(t - \tau_p)$ where τ_p is the propagation delay at the pth sensor ($\tau_0 = 0$ when delays are measured with respect to the 0th sensor). Since the geometry of the array, speed of propagation and direction of arrival are known or can be estimated independently the propagation delays are presumed to be known. The output of each sensor is advanced to make it in phase with the output of the reference sensor. Equation (3.43) reduces to

$$\hat{f}(t) = \sum_{p=0}^{P-1} \int_0^\infty h_p(\tau)\xi_0(t-\tau)d\tau$$

$$= \int_0^\infty \left[\sum_{p=0}^{P-1} h_p(\tau) \right] \xi_0(t-\tau)d\tau \qquad (3.44)$$

The filters are to be chosen to satisfy the condition that $\hat{f}(t) = \xi_0(t)$. From (3.44) it is clear that this constraint can be satisfied if

$$\sum_{p=0}^{P-1} h_p(t) = \delta(t) \qquad (3.45a)$$

or in the frequency domain

$$\sum_{p=0}^{P-1} H_p(\omega) = 1 \tag{3.45b}$$

Thus for distortion free extraction of waveforms the filters must satisfy (3.45) [17]. While an individual filter transfer function is allowed to be of any form, the sum must be a constant.

Another type of constraint arises when it is required to minimize the background noise power. The noise in the output of an array processor is given by

$$\hat{\eta}(t) = \sum_{p=0}^{P-1} \int_0^\infty h_p(\tau)\eta_p(t-\tau)d\tau$$

and the noise power is given by

$$\sigma_{\hat{\eta}}^2 = \sum_p \sum_q \int_0^\infty \int_0^\infty h_p(\tau)h_q(\tau')E\{\eta_p(t-\tau)\eta_q(t-\tau')\}d\tau d\tau'$$

$$= \sum_p \sum_q \int_0^\infty \int_0^\infty h_p(\tau)h_q(\tau')c_{pq}(\tau'-\tau)d\tau d\tau'$$

$$= \sum_p \sum_q \frac{1}{2\pi} \int_0^\infty H_p(\omega)H_q^*(\omega)S_{\eta_{pq}}(\omega)d\omega \tag{3.46}$$

For spatially and temporally uncorrelated noise (3.46) reduces to

$$\sigma_{\hat{\eta}}^2 = \sum_{p=0}^{P-1} \frac{1}{2\pi} \int_0^\infty |H_p(\omega)|^2 d\omega \, \sigma_\eta^2$$

The noise power in the array output shall be minimum whenever the filter transfer functions satisfy the condition

$$\sum_{p=0}^{P-1} \frac{1}{2\pi} \int_0^\infty |H_p(\omega)|^2 d\omega = \min \tag{3.47}$$

The trivial solution, namely, $H_p(\omega) = 0$ for all p, is not acceptable as it will set the output signal power also to zero. The constraint (3.45) or (3.47) is usually imposed along with other constraints. Sometimes the output noise power is set to a given fraction (noise reduction factor, NRF) of the input noise power.

$$\sum_{p=0}^{P-1} \frac{1}{2\pi} \int_0^\infty |H_p(\omega)|^2 \, d\omega = NRF \ <1 \qquad (3.48)$$

§3.4 **Wiener Filters for ULA and UCA:**
We shall now turn to some specific array geometries; in particular, we consider uniform linear array (ULA) and uniform circular array (UCA).
3.4.1 <u>Uniform Linear Array (ULA)</u>: As in the last section, the signal model assumed here is a plane wavefront sweeping across a linear array of sensors spaced at an interval d on x-axis. The noise is assumed to be spatially white. In (3.41a) the matrix $\mathbf{A}(\omega)$ and the column $\mathbf{B}(\omega)$ take the following form:

$$\mathbf{A}_{ULA}(\omega) = \begin{bmatrix} 1+T(\omega) & T(\omega)e^{ju_0 d} & \dots & T(\omega)e^{j(u_0(P-1)d} \\ T(\omega)e^{-ju_0 d} & 1+T(\omega) & \dots & T(\omega)e^{j(u_0(p-2)d} \\ & & \dots & \\ & & \dots & \\ & & \dots & \\ T(\omega)e^{-ju_0(P-1)d} & T(\omega)e^{-ju_0(P-2)d} & \dots 1+T(\omega) \end{bmatrix}$$

and

$$\mathbf{B}_{ULA} = col[1, e^{-ju_0 d}, e^{-ju_0 2d}, \dots e^{-ju_0(P-1)d}]$$

$\mathbf{A}_{ULA}(\omega)$ can be expressed as $(\mathbf{I} + T(\omega)\mathbf{a}(\omega)\mathbf{a}^H(\omega))$ (see (3.42)) where the vector $\mathbf{a}(\omega)$ for a ULA is given by

$$\mathbf{a}(\omega) = col[1, e^{-ju_0 d}, e^{-ju_0 2d}, \dots e^{-ju_0(P-1)d}]$$

Remember that in a ULA the right most sensor is conventionally taken as the reference sensor. Using the Woodbury's identity [18] we obtain its inverse in a closed form,

$$\mathbf{A}_{ULA}^{-1}(\omega) = \frac{1}{1 + PT(\omega)} \times$$

$$\begin{bmatrix} 1 + (P-1)T(\omega) & -T(\omega)e^{ju_0 d} & \dots & -T(\omega)e^{ju_0(P-1)d} \\ -T(\omega)e^{-ju_0 d} & 1 + (P-1)T(\omega) & \dots & -T(\omega)e^{ju_0(p-2)d} \\ & & \dots & \\ & & \dots & \\ & & \dots & \\ -T(\omega)e^{-ju_0(P-1)d} & -T(\omega)e^{-ju_0(P-2)d} & \dots & 1 + (P-1)T(\omega) \end{bmatrix}$$

$$(3.49)$$

Using (3.49) in (3.41b) we obtain transfer functions of Wiener filters

$$H_0(\omega) = \frac{T(\omega)}{1 + PT(\omega)}$$

$$H_1(\omega) = \frac{T(\omega)}{1 + PT(\omega)} e^{-ju_0 d}$$

$$H_2(\omega) = \frac{T(\omega)}{1 + PT(\omega)} e^{-j2u_0 d}$$

$$\begin{matrix} . \\ . \\ . \end{matrix}$$

$$(3.50)$$

$$H_{P-1}(\omega) = \frac{T(\omega)}{1 + PT(\omega)} e^{-ju_0(P-1)d}$$

The frequency wavenumber response of the Wiener filter for T=4 and sixteen sensor array (ULA) is shown in fig. 3.14. Notice that whenever $PT(\omega) \gg 1$, that is, either $T(\omega) = \dfrac{S_0(\omega)}{S_\eta(\omega)} \gg 1$ or P>>1 or both or the spectra of the signal and noise are nonoverlapping, the Wiener filter reduces to a simple delay filter

$$H_k(\omega) = \frac{1}{P} e^{-jk u_0 d} \tag{3.51}$$

3.4.2 <u>Uniform Circular Array</u>: The sensors are uniformly spaced on the circumference of a circular array of radius R units (see chapter 2 for more on circular arrays). For circular array the matrix $\mathbf{A}(\omega)$ takes the form,

$$\mathbf{A}_{UCA}(\omega) = \mathbf{I} + T(\omega)\mathbf{a}(\omega)\mathbf{a}^H(\omega) \tag{3.52a}$$

where

$$\mathbf{a}(\omega) = col\begin{bmatrix} e^{-ju_0 R} & e^{-j[u_0 R\cos\frac{2\pi}{P}+v_0 R\sin\frac{2\pi}{P}]} & \cdots \\ e^{-j[u_0 R\cos(P-1)\frac{2\pi}{P}+v_0 R\sin(P-1)\frac{2\pi}{P}]} \end{bmatrix} \tag{3.52b}$$

and

$$\mathbf{B}_{UCA}(\omega) = \mathbf{a}(\omega) \tag{3.53}$$

The reference point is at the center of the circle; however, there is no sensor physically present there. The Wiener filter is designed to predict the waveform as seen by an hypothetical sensor kept at the center of the circle. To solve for the Wiener filter (3.41b) we need to invert the $\mathbf{A}(\omega)$ matrix. We shall once again use Woodbury's identity and obtain

$$\mathbf{A}_{UCA}^{-1}(\omega) = \mathbf{I} - \frac{T(\omega)}{1+PT(\omega)}\left[\mathbf{a}(\omega)\mathbf{a}^H(\omega)\right] \tag{3.54}$$

Using (3.53) and (3.54) in (3.41b) we get

$$\mathbf{H}(\omega) = \frac{T(\omega)}{1+PT(\omega)}\mathbf{a}(\omega) \tag{3.55}$$

The Wiener filter for a circular array is similar to that for a linear array except for the difference arising out of the definition of $\mathbf{a}(\omega)$. The frequency wavenumber response of 16 sensor circular array of radius 10λ is shown in fig. 3.15. Although the mainlobe width is narrow, the sidelobe level is quite high. This is clearly brought out in a cross-sectional plot passing through the maximum (see fig. 3.16). When we increase the number of sensors to 64 the sidelobe level is brought down considerably, but the main lobe width remains practically unchanged. It may be emphasized that the sensors need not be spaced at $\leq 0.5\lambda$ as in a ULA [19]. What is gained by increasing the number of sensors (keeping the radius fixed) is the reduction of the sidelobe level. In contrast, in case of a ULA, by increasing the number of sensors the array aperture is increased which in turn sharpens the mainlobe but does not reduce the sidelobe level.
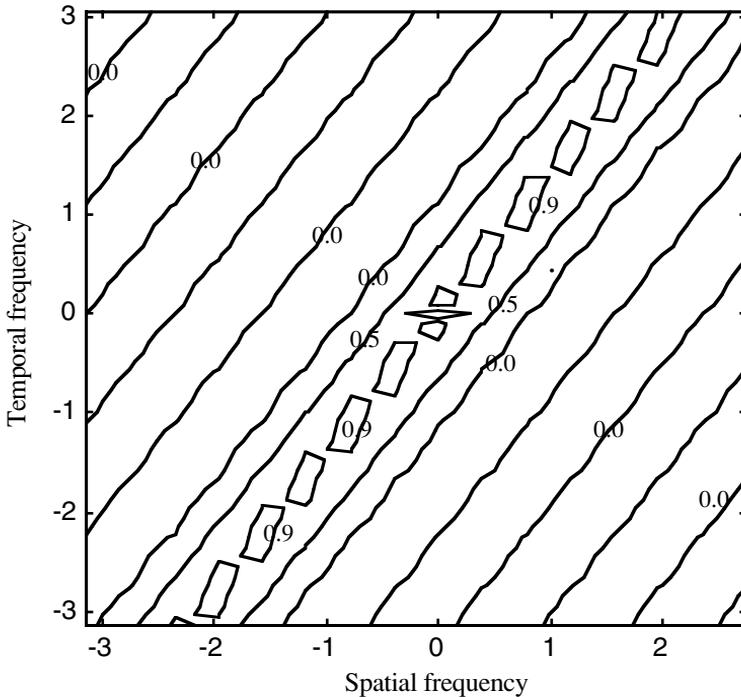
Figure 3.14: Frequency wavenumber response of the Wiener filter for ULA. The direction of arrival (DOA) is assumed to be known. In the above example it is equal to 30°.

3.4.3 <u>Robustification</u>: The Wiener filters given by (3.51) and (3.55) require a knowledge of $\mathbf{a}(\omega)$ for which we need to know the apparent speed of the wavefront sweeping across the array. Prior to waveform estimation it is a common practice to estimate the direction of arrival (DOA) of a wavefront. (This topic is covered in some detail in chapter 5). The DOA estimation is not without error. Hence, it would be nice if the Wiener filters are made robust so that the degradation in its performance is minimum. We shall confine to a linear array (ULA). An error in DOA estimation will introduce an error in the wavenumber. For a ULA, the erroneous wavenumber may be expressed as $(u_0 + \varepsilon)$ where ε is a uniformly distributed random variable. Naturally, there will be an error in each element of $\mathbf{A}_{ULA}(\omega)$ and $\mathbf{B}_{ULA}(\omega)$. We shall use the stochastically averaged $\mathbf{A}_{ULA}(\omega)$ and $\mathbf{B}_{ULA}(\omega)$ matrices in (3.41b). This approach was suggested in [20] in the context of optimal velocity filters in
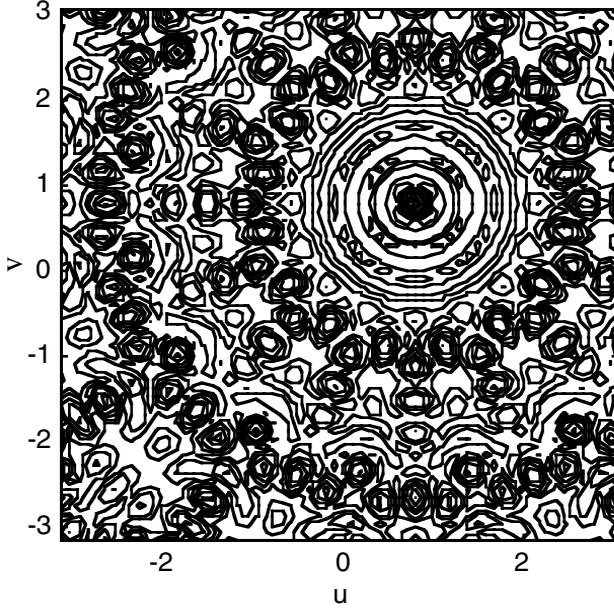
Figure 3.15: Frequency wavenumber response of Wiener filter for UCA. The DOAs are known (azimuth=elevation=45° and $\omega=\pi/2$). 16 sensors and constant snr=4 are assumed.

seismic exploration. Observe that in each element of the matrices there is an extra term which does not permit the matrix to be written as an outer product of two vectors as in (3.52a).

$$\overline{\mathbf{A}}_{ULA}(\omega) =$$

(3.56a)

$$
\begin{bmatrix}
1+T(\omega) & T(\omega)\sin c(\varepsilon_0 d)e^{ju_0 d} & \dots & T(\omega)\sin c(\varepsilon_0 (P-1)d)e^{ju_0 (P-1)d} \\
T(\omega)\sin c(\varepsilon_0 d)e^{-ju_0 d} & 1+T(\omega) & \dots & T(\omega)\sin c(\varepsilon_0 (P-2)d)e^{ju_0 (p-2)d} \\
& & \dots & \\
& & \dots & \\
& & \dots & \\
T(\omega)\sin c(\varepsilon_0 (P-1)d)e^{-ju_0 (P-1)d} & T(\omega)\sin c(\varepsilon_0 (P-2)d)e^{-ju_0 (P-2)d} & \dots & 1+T(\omega)
\end{bmatrix}
$$

and

$$\overline{\mathbf{B}}_{ULA} = col\begin{bmatrix} 1, \sin c(\varepsilon_0 d)e^{-ju_0 d}, \sin c(2\varepsilon_0 d)e^{-ju_0 2d}, \dots, \\ \sin c((P-1)\varepsilon_0 d)e^{-ju_0 (P-1)d} \end{bmatrix}$$

(3.56b)
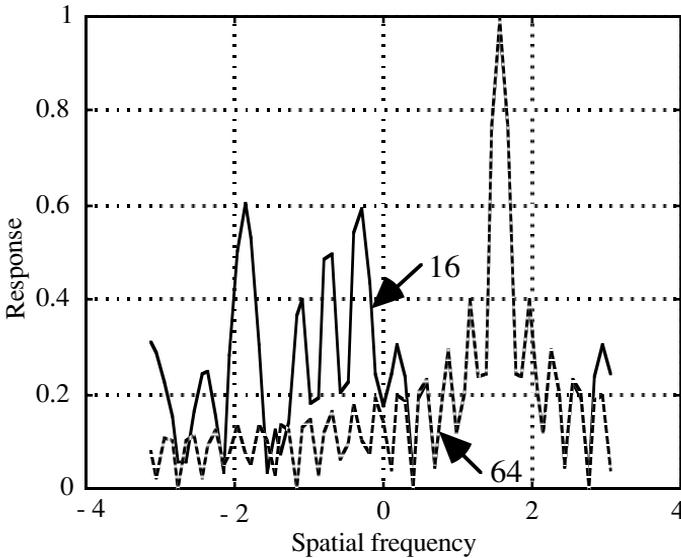
Figure 3.16: A cross-section of the Wiener filter response taken through the peak. Number of sensors are 16 and 64. The sidelobe level relative to the peak has been reduced when the number of sensors is increased from 16 to 64. Radius of circular aperture is 10 units.

The effectiveness of the proposed approach is demonstrated through frequency wavenumber response of the Wiener filters before and after robustification. We compute the frequency wavenumber response given that the DOA is known to an accuracy of $\pm 2.0^o$. The plot is shown in fig. 3.17a for a ULA with 16 sensors and constant snr (=4). Notice the splitting of the peak particularly in the higher temporal frequency range. Next, we compute the frequency wavenumber response of a UCA designed to tolerate an error of 2°. The frequency wavenumber response is shown in fig. 3.17b. The main lobe shape remains practically unchanged but there is an increase in the sidelobe level. Cross-sectional plots passing through the maximum of the response function of the UCA with and without DOA error are shown in fig. 3.18. While the shape of the main lobe remains practically unchanged the side lobe level seems to have slightly increased. This is the price one has to pay for the lack of exact knowledge of the DOA.

3.4.4 <u>Levinson-Durbin Algorithm</u>: In (3.36), in order to solve for **h**, we have to invert a large covariance matrix (e.g., with the array size, M=24, and filter length, N=16, the size of the covariance matrix will be 384x384). The
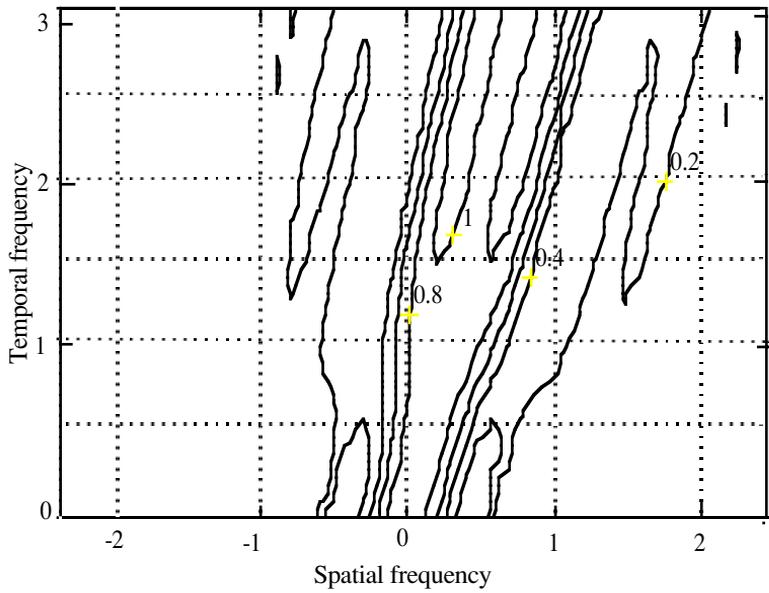
(a)
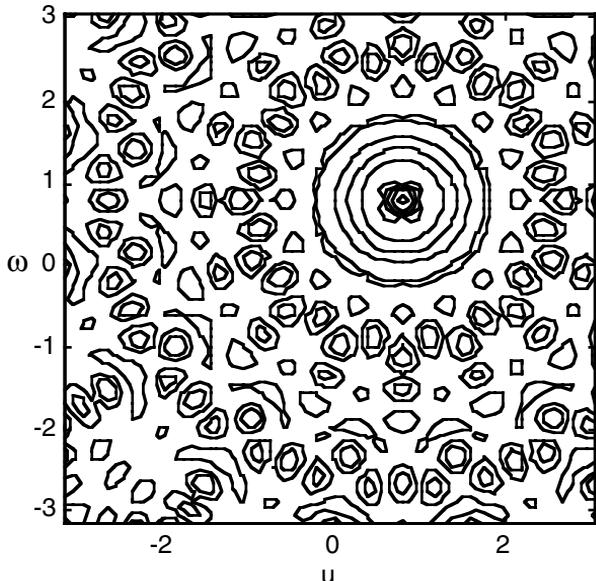
(b)

Figure 3.17: Frequency wavenumber response of Wiener filter when the error in DOA estimate is $\pm 2.0^{o}$. (a) ULA and (b) UCA. 16 sensors and 64 time samples are assumed.
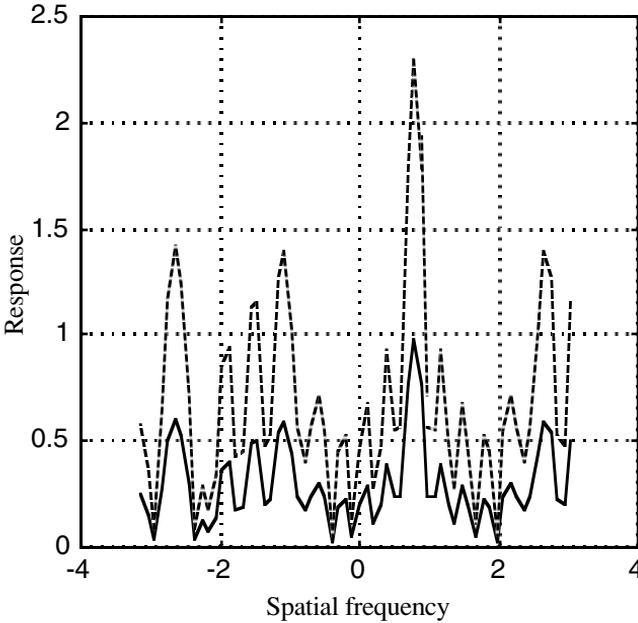
Figure 3.18: A cross-section through the maximum (v0=0.7854) for UCA. Solid line shows response when there is no error and the dashed line shows one with DOA error ($\pm 2.0^o$). 16 sensors and 64 time samples are assumed.

computational load for inverting such a large matrix will be very high. We shall outline a recursive method applicable to a ULA. For any other geometry of an array the covariance matrix becomes a block symmetric matrix, but only for a ULA does the covariance matrix become toeplitz. This important property enables us to devise a recursive algorithm, known as Levinson-Durbin algorithm, which requires inversion of a matrix of size $M \times M$ in place of a matrix of size $MN \times MN$. Briefly, the algorithm is as follows [16]: Let $\mathbf{h}_N$ ($\mathbf{h}_N = col[\mathbf{h}(0), \mathbf{h}(\Delta\tau), \dots \mathbf{h}((N-1)\Delta\tau)]$) be the solution of the Nth order normal equations, (3.36). Let us now increase the size of the covariance matrix by padding one row and one column of covariance matrices as below,

$$\begin{bmatrix} \mathbf{C}_N & \mathbf{c}(N) \\ & \vdots \\ \mathbf{c}(N) \dots & \mathbf{c}(0) \end{bmatrix} \begin{bmatrix} \mathbf{h}_N \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_0 \\ \gamma_N \end{bmatrix} \qquad (3.57)$$

where

$$\gamma_N \quad = \quad \begin{bmatrix} \mathbf{c}(N) \dots & \mathbf{c}(1) \end{bmatrix} \mathbf{h}_N$$
$$(M \times 1) \qquad (M \times MN) \quad (MN \times 1)$$

Note that the square matrix on the left hand side of (3.57) is a block covariance matrix of size $M(N+1) \times M(N+1)$; therefore, (3.57) is similar to (3.36) but of order N+1. Indeed, if we subtract (3.57) from (3.36) of the same order, we shall obtain

$$\mathbf{C}_{N+1}\left[\mathbf{h}_{N+1} - \begin{bmatrix} \mathbf{h}_N \\ \mathbf{0} \end{bmatrix} \right] = \begin{bmatrix} \mathbf{0} \\ \mathbf{c}_0(N) - \gamma_N \end{bmatrix} \qquad (3.58)$$

Define a set of auxiliary coefficient matrices

$$\mathbf{b}_N = col\{\mathbf{b}_{NN}, \mathbf{b}_{NN-1}, \dots \mathbf{b}_{N1}, \mathbf{I}\}$$
$$(N+1)M \times M$$

where $\mathbf{b}_{NN-i}, i = 0, \dots, N-1$ are $M \times M$ matrices yet to be defined, but they satisfy the following recurrence relation:

$$\mathbf{h}_{N+1}(i) = \mathbf{b}_{NN-i}\mathbf{h}_{N+1}(N) + \mathbf{h}_N(i), \quad i = 0,1,\dots,N-1 \quad (3.59)$$

$$\mathbf{C}_{N+1}\begin{bmatrix} \mathbf{b}_{NN}\mathbf{h}_{N+1}(N) \\ \mathbf{b}_{NN-1}\mathbf{h}_{N+1}(N) \\ \mathrm{M} \\ \mathbf{h}_{N+1}(N) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \beta_N\mathbf{h}_{N+1}(N) \end{bmatrix} \qquad (3.60a)$$

where $\beta_N\mathbf{h}_{N+1}(N) = \mathbf{c}_0(N) - \gamma_N$. Eliminating $\mathbf{h}_{N+1}(N)$ from both sides of (3.60a) we obtain

$$\mathbf{C}_{N+1}\mathbf{b}_N = \begin{bmatrix} \mathbf{0} \\ \beta_N \end{bmatrix} \qquad (3.60b)$$

We shall once again increase the order of $\mathbf{C}_{N+1}$ in (3.60b) by padding one more row and one more column. We get

$$\mathbf{C}_{N+2}\begin{bmatrix}\mathbf{0}\\\mathbf{b}_N\end{bmatrix}=\begin{bmatrix}\beta'_N\\\mathbf{0}\\\beta_N\end{bmatrix} \qquad (3.60c)$$

where

$$\beta'_N=\begin{bmatrix}\mathbf{c}(1),\ \mathbf{c}(2),...\mathbf{c}(N+1),\end{bmatrix}\quad \mathbf{b}_N$$
$$(M\times M(N+1))\quad M(N+1)\times M$$

We shall now introduce another set of auxiliary coefficients, $\mathbf{a}_N = col\{\mathbf{I}, \mathbf{a}_{N1}, \mathbf{a}_{N2}, ..., \mathbf{a}_{NN}\}$ as a solution of the following system of equations:

$$\mathbf{C}_{N+1}\mathbf{a}_N = \begin{bmatrix}\alpha_N\\\mathbf{0}\end{bmatrix} \qquad (3.61a)$$

where

$$\alpha_N = \begin{bmatrix}\mathbf{c}(0)\mathbf{I} + \mathbf{c}(1)\mathbf{a}_{N1} + ... + \mathbf{c}(N)\mathbf{a}_{NN}\end{bmatrix}$$

$M \times M$

Let us now increase the size of $\mathbf{C}_{N+1}$ in (3.61a) by padding one more row and one more column of covariance matrices. We obtain

$$\mathbf{C}_{N+2}\begin{bmatrix}\mathbf{a}_N\\\mathbf{0}\end{bmatrix}=\begin{bmatrix}\alpha_N\\\mathbf{0}\\\alpha'_N\end{bmatrix} \qquad (3.61b)$$

where

$$\alpha'_N = \begin{bmatrix}\mathbf{c}(N+1)\mathbf{I} + \mathbf{c}(N)\mathbf{a}_{N1} + ... + \mathbf{c}(1)\mathbf{a}_{NN}\end{bmatrix}$$

$M \times M$

We linearly combine (3.60c) and (3.61b) such that the resulting equation is the $(N+2)^{th}$ order equivalent of (3.60b). Let the linear combination be given by

$$C_{N+2}\left\{\begin{bmatrix} \mathbf{0} \\ \mathbf{b}_N \end{bmatrix} + \begin{bmatrix} \mathbf{a}_N \\ \mathbf{0} \end{bmatrix} \delta_N \right\} = \left\{\begin{bmatrix} \beta'_N \\ \mathbf{0} \\ \beta_N \end{bmatrix} + \begin{bmatrix} \alpha_N \\ \mathbf{0} \\ \alpha'_N \end{bmatrix} \delta_N \right\} \qquad (3.62a)$$

where δ_N is a $M \times M$ matrix of constants which we shall select in such a manner that the right hand side of (3.62a) resembles the right hand side of (3.60b). This may be achieved by requiring

$$\beta'_N + \alpha_N \delta_N = \mathbf{0}$$

or

$$\delta_N = -\alpha_N^{-1}\beta'_N \qquad (3.62b)$$

The resulting equation is equivalent to (3.60) but of order N+2. Then, we have

$$\mathbf{b}_{N+1} = \begin{bmatrix} \mathbf{0} \\ \mathbf{b}_N \end{bmatrix} + \begin{bmatrix} \mathbf{a}_N \\ \mathbf{0} \end{bmatrix} \delta_N \qquad (3.63a)$$

$$\beta_{N+1} = \beta_N + \alpha'_N \delta_N \qquad (3.63b)$$

Further we take a linear combination of (3.60c) and (3.61b) such that the resulting equation resembles (3.61a) but it is of order N+2. This may be achieved by requiring

$$\Delta_N \beta_N + \alpha'_N = 0$$

or

$$\Delta_N = -\beta_N^{-1}\alpha'_N \qquad (3.64a)$$

where Δ_N is also a $M \times M$ matrix of constants for linear combination. Now we have

$$\mathbf{a}_{N+1} = \begin{bmatrix} \mathbf{a}_N \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{b}_N \end{bmatrix} \Delta_N \qquad (3.64b)$$

$$\alpha_{N+1} = \alpha_N + \beta'_N \Delta_N \qquad (3.64c)$$

Eqs. (3.63a) and (3.64b) form a set of recursive relations to compute $\mathbf{a}_{N+1}$ and $\mathbf{b}_{N+1}$ given $\mathbf{a}_N$ and $\mathbf{b}_N$. Similarly, equations (3.63b) and (3.64c) form another

set of recursive relations to compute α_{N+1} and β_{N+1} given α_N and β_N. The initial conditions are:

1) $\mathbf{b}_{0,0} = \mathbf{a}_{0,0} = \mathbf{I}$,   2) $\alpha_0 = \beta_0 = \mathbf{c}(0)$ and   3) $\mathbf{h}_0 = 0$

Finally, to compute the filter coefficients we need $\mathbf{h}_{N+1}(N)$ which we get from

$$\mathbf{h}_{N+1}(N) = \left[\beta_N\right]^{-1}\left(\mathbf{c}_0(N) - \gamma_N\right)$$

It must be emphasized that, since the Levinson-Durbin algorithm exploits the toeplitz property of the covariance matrix, the array signal must be stationary and sufficiently large data must be available for the estimation of statistically averaged covariance matrix. Such a filter will naturally be optimum to an ensemble of time series having the same second order structure. In practice, this is often difficult to realize; consequently, the toeplitz character is often lost and we cannot use the Levinson-Durbin algorithm altogether. Later in chapter 6 we shall describe a deterministic least squares approach which uses the actual measured signal in place of covariance function.

§3.5 **Predictive Noise Cancellation:**
Noise suppression is based on the principle of differential spectral properties of signal and noise. A filter is often used to maximally attenuate the noise power but at the same time to minimally attenuate the signal power. There is an alternate approach to noise attenuation known as noise cancellation. This involves prediction of the noise that is actually corrupting the signal. For prediction we need a sample of noise which is not contaminated with the signal but which is correlated with the noise present in the actual observed signal plus noise sample. In principle, it is possible to devise a Wiener filter for the prediction of the unknown noise from the sample of noise which is correlated with the unknown. A signal free noise sample can be obtained as follows:
(a) Place an extra sensor within the correlation distance of the noise but away from the signal source. This is possible when the signal source is in the near field region but the noise sources are in the far field region, for example, when a speaker is close to a microphone and the noise sources are far away from the microphone.
(b) Both signal and noise are in the far field region but reach the microphone array from different directions. An array of sensors may be simultaneously used to receive the signal and the noise coming from different directions. However, since the array response is finite in a direction other than the direction to which it is tuned, some amount of noise will leak into the array output. The noise, which has leaked into the array output, will be strongly correlated with the

incident noise and hence it may be predicted using the noise output when the array is steered in the direction of the noise.

In the above approach the noise is canceled electronically. It is also possible to achieve the noise cancellation acoustically. This would, however, require use of many coherently generated noise sources whose combined effect is to produce a noise wavefront with a phase equal but opposite to that of the noise present in the observed waveform.

3.5.1 <u>Signal Source in Near Field</u>: Consider a speaker close to a microphone (M_1) and another microphone (M_2) away from the speaker but well within the correlation distance of the noise (see fig. 3.19). It is assumed that all noise sources are in the far field region. Let $f_1(t)$ be the output of M_1, $f_2(t)$ be the output of M_2, $\xi_0(t)$ be the signal emitted by the speaker and $\eta_1(t)$ be the noise in M_1. The signal and noise in M_2 are related to those in M_1 through impulse response functions, $h_1(t)$ and $h_2(t)$.

$$f_1(t) = \xi_0(t) + \eta_1(t)$$

$$f_2(t) = \int_0^\infty \xi_0(t-t')h_1(t')dt' + \int_0^\infty \eta_1(t-t')h_2(t')dt' \qquad (3.65)$$

The output of M_2 is passed through a prediction filter, $h_{pred}(t)$, which is found by minimizing a quantity,

$$E\left\{ \left| f_1(t) - \int_0^\infty f_2(t'-t)h_{pred}(t')dt' \right|^2 \right\} = \min$$

or in the frequency domain by minimizing the following

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \begin{bmatrix} S_1(\omega) + S_2(\omega)\left|H_{pred}(\omega)\right|^2 - \\ S_{12}(\omega)H_{pred}^*(\omega) - S_{12}^*(\omega)H_{pred}(\omega) \end{bmatrix} d\omega$$

We obtain

$$H_{pred}(\omega) = \frac{S_{12}(\omega)}{S_2(\omega)} = \frac{S_0(\omega)H_1^*(\omega) + S_{\eta_1}(\omega)H_2^*(\omega)}{S_0(\omega)\left|H_1(\omega)\right|^2 + S_{\eta_1}(\omega)\left|H_2(\omega)\right|^2} \qquad (3.66a)$$

and
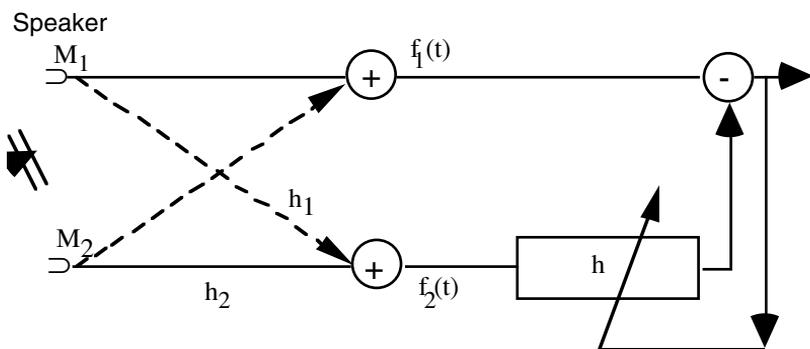
Figure 3.19: Source is in near field and noise is in far field. Microphone M_2 (reference microphone) receives very little of the signal.

$$Error\big|_{min} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[S_1(\omega) - S_2(\omega) \big| H_{pred}(\omega) \big|^2 \right] d\omega \qquad (3.66b)$$

<u>Some Special Cases</u>:

(a) $h_1(t) = 0$. The signal from the speaker does not reach M_2. $h_2(t) = \delta(t - \tau_0)$. The noise reaching the microphone M_2 is simply a delayed version of the noise reaching microphone M_1. For this case, the prediction filter is simply a delay filter, $H_{pred}(\omega) = e^{j\omega\tau_0}$ and the minimum error is equal to the signal power,

$$Error\big|_{min} = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_0(\omega) d\omega$$

In this special case complete noise cancellation takes place.

(b) $h_1(t) = h_2(t) \neq 0$. The output of the reference microphone is a filtered version of the output of M_1. For this case, the prediction filter is given by

$$H_{pred}(\omega) = \frac{1}{H_1(\omega)}$$

and the minimum error is equal to zero. Apparently, both signal and noise are canceled and the output power is zero. For noise cancellation to take place we must have $h_2(t) > h_1(t)$.

Define gain G as a ratio of snr at the output to snr at the input. The output power is given by

$$Output\ power = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} [S_1(\omega) - S_2(\omega)|H_{pred}(\omega)|^2]d\omega$$

$$= \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \left\{ \begin{array}{c} S_0(\omega) + S_{\eta_1}(\omega) - (S_0(\omega)|H_1(\omega)|^2 + \\ S\eta_1(\omega)|H_2(\omega)|^2)|H_{pred}(\omega)|^2 \end{array} \right\} d\omega$$

$$= \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \left\{ S_0(\omega)(1-|H_1(\omega)|^2|H_{pred}(\omega)|^2) \right\} d\omega \quad Signal\ power$$

$$+ \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \left\{ S_{\eta_1}(\omega)(1-|H_2(\omega)|^2|H_{pred}(\omega)|^2) \right\} d\omega \quad Noise\ power$$

The snr at the output as a function of frequency is given by

$$SNR_{output} = \frac{\left\{ S_0(\omega)(1-|H_1(\omega)|^2|H_{pred}(\omega)|^2) \right\}}{\left\{ S_{\eta_1}(\omega)(1-|H_2(\omega)|^2|H_{pred}(\omega)|^2) \right\}}$$

from which the gain as defined here turns out to be

$$G = \frac{[SNR]_{output}}{[SNR]_{input}} = \frac{1-|H_1(\omega)|^2|H_{pred}(\omega)|^2}{1-|H_2(\omega)|^2|H_{pred}(\omega)|^2} \qquad (3.67)$$

For G>1 we must have $|H_1(\omega)|^2 < |H_2(\omega)|^2$.

3.5.2 <u>Source in Far Field</u>: Both signal and noise sources are in the far field region but the DOAs of their wavefronts are different. Let τ_0 and τ_1 be the incremental delays produced by the signal wavefront and noise wavefront respectively. The array can be steered to receive the signal or noise at the same time (see fig. 3.20)[21]. Let $f_1(t)$ be the output of an array when it is steered to the signal wavefront and $f_2(t)$ be the output of an array when steered to the noise wavefront. Since the array response function has finite side lobes, some

amount of wave energy will leak through the sidelobes. Hence, we model the array output as

$$f_1(t) = \xi_0(t) + \frac{1}{2\pi} \int_{-\infty}^{\infty} N_1(\omega) H(\omega(\tau_0 - \tau_1)) e^{j\omega t} d\omega \qquad (3.68a)$$

and

$$f_2(t) = \eta_1(t) + \frac{1}{2\pi} \int_{-\infty}^{\infty} \Xi_0(\omega) H(\omega(\tau_1 - \tau_0)) e^{j\omega t} d\omega \qquad (3.68b)$$

where $\Xi_0(\omega)$ is the Fourier transform of the signal and $N_1(\omega)$ is that of the noise. By comparing (3.65) with (3.68) it is possible to write

$$H_1(\omega) = H(\omega(\tau_1 - \tau_0)) \qquad (3.69a)$$

and

$$H_2(\omega) = \frac{1}{H(\omega(\tau_0 - \tau_1))} \qquad (3.69b)$$

Using (3.69) in (3.66a) we obtain a filter to predict the noise in $f_1(t)$,

$$
\begin{aligned}
H_{pred}(\omega) &= \frac{S_{\xi_0}(\omega) H^*(\omega(\tau_1 - \tau_0)) + \dfrac{S_{\eta_1}(\omega)}{H^*(\omega(\tau_0 - \tau_1))}}{S_{\xi_0}(\omega)\left|H(\omega(\tau_1 - \tau_0))\right|^2 + \dfrac{S_{\eta_1}(\omega)}{\left|H(\omega(\tau_0 - \tau_1))\right|^2}} \\[2mm]
&= \frac{1 + SNR_{input}}{1 + SNR_{input}\left|H(\omega(\tau_1 - \tau_0))\right|^2} H(\omega(\tau_0 - \tau_1))
\end{aligned}
$$

Let us now consider a few special cases:

(a) When $SNR_{input} \gg 1$ and $SNR_{input}\left|H(\omega(\tau_1 - \tau_0))\right|^2 \gg 1$

$$H_{pred}(\omega) \approx \frac{1}{H(\omega(\tau_1 - \tau_0))}$$

If this filter is used on $f_2(t)$ (see eq.(3.68b)) for predicting the noise in $f_1(t)$, the signal component will be restored causing the cancellation of the signal.
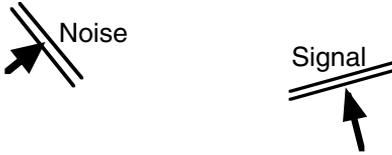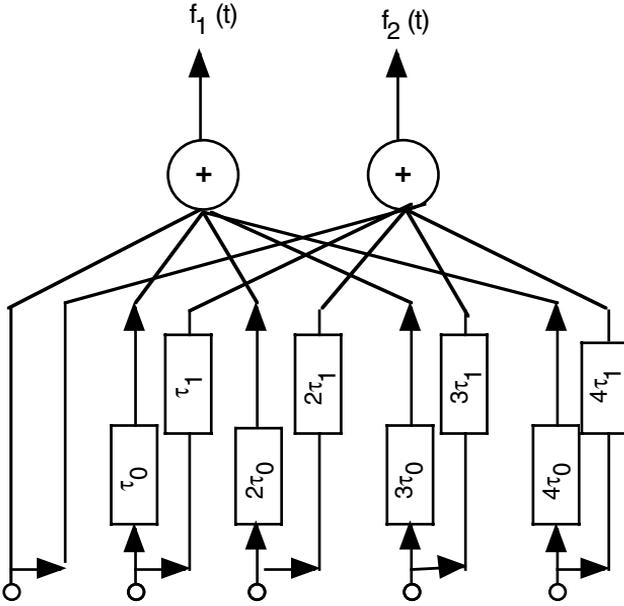
Figure 3.20: An array of sensors can be steered simultaneously in the direction of the signal and in the direction of noise. When the array is steered in the direction of signal the output $f_1(t)$ is mostly signal and when it is steered in the direction of noise the output $f_2(t)$ is mostly noise.

(b) $SNR_{input} \left| H(\omega(\tau_1 - \tau_0)) \right|^2 << 1$

$$H_{pred}(\omega) \approx (1 + SNR_{input}) H(\omega(\tau_0 - \tau_1))$$

If this filter is used on $f_2(t)$ (see eq.(3.68b)) for predicting the noise component in $f_1(t)$ the noise component will be largely canceled without canceling the signal. As an illustration, we consider two pure sinusoidal signals (of same frequency) arriving with different DOAs (0° and 5.7^o) at a ULA of 16 sensors spaced at $\lambda/2$ spacing. The second sinusoid arrives 50 time units later with an amplitude of 0.8. Figure 3.21(a) shows a sum of the two tones as received by the first sensor. The array is steered in the direction of the first sinusoid and at the same time in the direction of the second sinusoid. The array outputs are described by (3.68), which is now considerably simplified for pure sinusoidal inputs.

$$f_1(t) = s_1(t) + s_2(t)H(\omega_0(\tau_0 - \tau_1))$$
$$f_2(t) = s_2(t) + s_1(t)H(\omega_0(\tau_0 - \tau_1))$$

(3.70a)

where $s_1(t)$ and $s_2(t)$ are the first and the second sinusoid, respectively and ω_0 is the frequency of the sinusoids. Solving (3.70) we obtain, for $|H(\omega_0(\tau_0 - \tau_1))|^2 < 1$,

$$s_1(t) = \frac{f_1(t) - f_2(t)H(\omega_0(\tau_0 - \tau_1))}{1 - H^2(\omega_0(\tau_0 - \tau_1))}$$

(3.70b)

$$s_2(t) = \frac{f_2(t) - f_1(t)H(\omega_0(\tau_0 - \tau_1))}{1 - H^2(\omega_0(\tau_0 - \tau_1))}$$

The results are shown in figs. 3.21 (b & c).

We explore another possibility, where, instead of steering the array in the direction of interference, we steer a null in the direction of a signal with finite response in the direction of interference [22]. Equation (3.68b) may be written as

$$f_2(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{H}_{null}^H(\omega)\mathbf{N}_1(\omega)e^{j\omega t}d\omega$$

(3.71a)

where

$$\mathbf{H}_{null}(\omega) = \left[\mathbf{I} - \frac{\mathbf{a}_0(\omega)\mathbf{a}_0^H(\omega)}{M}\right]\mathbf{a}_1(\omega)$$
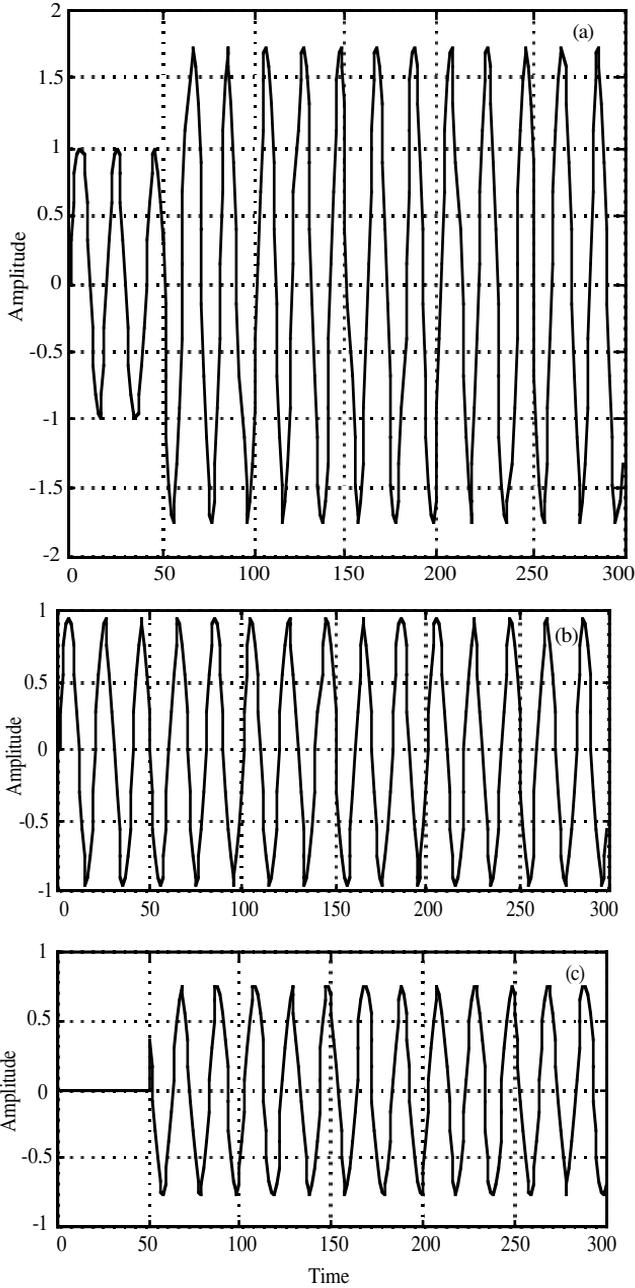
Figure 3.21: (a) sum of two sinusoids, (b) first sinusoid after subtraction and (c) second sinusoid after subtraction.

is a filter with a null in the direction of the signal. Noting that $N_1(\omega) = \mathbf{a}_1(\omega)N_1(\omega)$, (3.71a) may be written as

$$f_2(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} M\left[1 - \frac{\left|\mathbf{a}_1^H(\omega)\mathbf{a}_0(\omega)\right|^2}{M^2}\right]N_1(\omega)e^{j\omega t}d\omega \quad (3.71b)$$

Define a filter

$$w = \frac{\dfrac{\mathbf{a}_0^H(\omega)\mathbf{a}_1(\omega)}{M}}{M\left[1 - \dfrac{\left|\mathbf{a}_1^H(\omega)\mathbf{a}_0(\omega)\right|^2}{M^2}\right]} \quad (3.71c)$$

If we now pass $f_2(t)$ through above filter (3.71c) the output will be exactly the same as the noise term in (3.68a); therefore it may be removed by simple subtraction.

3.5.3 <u>Adaptive Filter</u>: We have derived the multichannel Wiener filter in §3.3. A single channel version may be derived along the same lines (see [22] for derivation). Here we shall state the final result. The Wiener filter which predicts $f_1(t)$ from $f_2(t)$ is given by $\mathbf{h} = \mathbf{C}_{f_2}^{-1}\mathbf{C}_{f_1 f_2}$ where $\mathbf{C}_{f_2}$ is the covariance matrix of $f_2(t)$ and $\mathbf{C}_{f_2 f_1}$ is the cross-covariance matrix between $f_2(t)$ and $f_1(t)$. For real time estimation of the filter and also to account for temporal variations in the covariance functions it is appropriate to devise an adaptive approach which in the limiting case reduces to the Wiener solution. Let $\mathbf{h} = \begin{bmatrix} h_0, h_1, h_2 & \dots & h_{N-1} \end{bmatrix}^T$ be the prediction filter vector and

$$\mathbf{f}_2 = \begin{bmatrix} f_2(t), f_2(t-\Delta t), f_2(t-2\Delta t), \dots f_2(t-(N-1)\Delta t) \end{bmatrix}^T$$

be the data vector. The filter output is given by $\mathbf{h}^T\mathbf{f}_2$, which is required to be as close as possible to $f_1(t)$. This filter is known as a transversal filter acting on the delayed outputs (see fig. 3.22). For this we need to minimize the mean square error

$$E\left\{\left|\varepsilon(t)\right|^2\right\} = E\left\{\left|f_1(t) - \mathbf{h}^T\mathbf{f}_2\right|^2\right\}$$
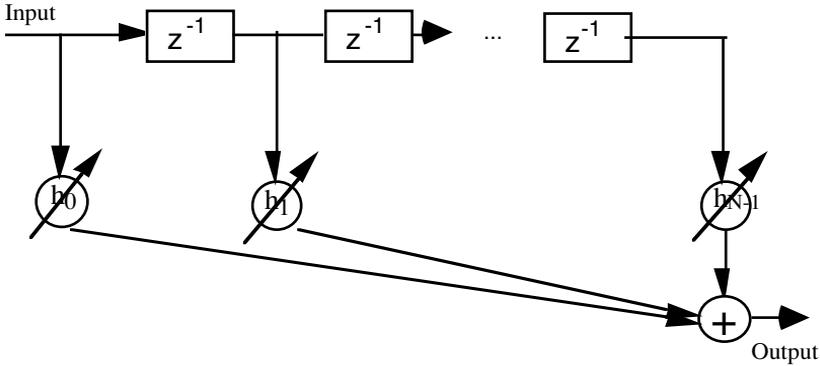
Figure 3.22: Structure of transversal filter. The filter coefficients are made adaptable to the changing input.

with respect to the prediction filter coefficients. To minimize the error power we need to compute a gradient of $E\{|\varepsilon(t)|^2\}$ with respect to $h_0, h_1, h_2 \ldots h_{N-1}$ and go down the path of the steepest descent until a minimum (possibly a local minimum) is encountered. The difficulty, however, is in estimating $E\{|\varepsilon(t)|^2\}$, which requires averaging over a finite interval (ideally infinite) of time. Instead, in the least mean squares (LMS) algorithm, it is proposed to use $|\varepsilon(t)|^2$ in place of $E\{|\varepsilon(t)|^2\}$. The gradient of $|\varepsilon(t)|^2$ is now easily computed,

$$\nabla|\varepsilon(t)|^2 = 2\varepsilon(t)\left\{\frac{\partial\varepsilon(t)}{\partial h_0}, \frac{\partial\varepsilon(t)}{\partial h_1}, \ldots \frac{\partial\varepsilon(t)}{\partial h_{N-1}},\right\} \quad (3.72)$$
$$= -2\varepsilon(t)\mathbf{f}_2(t)$$

In the steepest descent search method the current filter vector is adjusted by an amount proportional to negative of the gradient of the error function, $\nabla|\varepsilon(t)|^2$ [24]. The idea of adaptation is illustrated in fig. 3.23. The current filter vector is updated by an amount proportional to the product of prediction error and current input,

$$\mathbf{h}_{i+1} = \mathbf{h}_i + 2\mu\varepsilon(t)\mathbf{f}_2(t) \quad (3.73)$$

where μ is a gain constant which regulates the speed of adaptation.

It is interesting to note that the filter vector converges to the Wiener filter, that is, as $i \to \infty$ $\mathbf{h}_i \to \mathbf{C}_{f_2}^{-1}\mathbf{C}_{f_1 f_2}$. To show this consider the expected value of $\mathbf{h}_{i+1}$

$$
\begin{aligned}
E\{\mathbf{h}_{i+1}\} &= E\{\mathbf{h}_i\} + E\{2\mu\varepsilon(t)\mathbf{f}_2(t)\} \\
&= E\{\mathbf{h}_i\} + 2\mu E\{\mathbf{f}_2(t)(f_1(t) - \mathbf{f}_2^T(t)\mathbf{h}_i)\} \\
&= E\{\mathbf{h}_i\} + 2\mu\mathbf{C}_{f_2 f_1} - 2\mu\mathbf{C}_{f_2} E\{\mathbf{h}_i\}
\end{aligned}
\tag{3.74}
$$

where we have assumed that $\mathbf{h}_i$ and $\mathbf{f}_2(t)$ are independent. Let $\tilde{\mathbf{h}}$ represent filter coefficients obtained by solving the normal equation (3.36), that is, $\tilde{\mathbf{h}} = \mathbf{C}_{f_2}^{-1}\mathbf{C}_{f_2 f_1}$. Equation (3.74) reduces to

$$
E\{\mathbf{h}_{i+1} - \tilde{\mathbf{h}}\} = E\{\mathbf{h}_i - \tilde{\mathbf{h}}\} + 2\mu\tilde{\mathbf{h}}\mathbf{C}_{f_2} - 2\mu\mathbf{C}_{f_2} E\{\mathbf{h}_i\} \tag{3.75a}
$$

$$
E\{\Delta\mathbf{h}_{i+1}\} = E\{\Delta\mathbf{h}_i\} - 2\mu\mathbf{C}_{f_2} E\{\Delta\mathbf{h}_i\} \tag{3.75b}
$$

$$
E\{\Delta\mathbf{h}_{i+1}\} = (\mathbf{I} - 2\mu\mathbf{C}_{f_2})E\{\Delta\mathbf{h}_i\} \tag{3.75c}
$$

Let us use the eigendecomposition of the covariance matrix, $\mathbf{C}_{f_2} = \mathbf{V}\Lambda\mathbf{V}^H$, in (3.75c) and obtain

$$
E\{\mathbf{V}^H\Delta\mathbf{h}_{i+1}\} = (\mathbf{I} - 2\mu\Lambda)E\{\mathbf{V}^H\Delta\mathbf{h}_i\} \tag{3.76}
$$

The solution of the above difference equation (3.76) is given by $E\{\mathbf{V}^H\Delta\mathbf{h}_i\} = (\mathbf{I} - 2\mu\Lambda)^i \chi_0$ where χ_0 is the initial condition. As $i \to \infty$ $E\{\mathbf{V}^H\Delta\mathbf{h}_i\} \to \mathbf{0}$ provided $(\mathbf{I} - 2\mu\Lambda)^i \to \mathbf{0}$ for $i \to \infty$. This is possible iff $|(\mathbf{I} - 2\mu\lambda_i)| < 1$ for all i. This can be easily achieved if we were to select μ such that $0 < \mu < \dfrac{1}{\lambda_{max}}$ where λ_{max} stands for the maximum eigenvalue of the covariance matrix. Note that $\lambda_{max} \le tr\{\mathbf{C}_{f_2}\} =$ the sum of the eigenvalues of the covariance matrix. Hence, $E\{\Delta\mathbf{h}_i\} \to \mathbf{0}$ as $i \to \infty$. From this result it follows that $E\{\mathbf{h}_i\} \to \tilde{\mathbf{h}}$ as $i \to \infty$.
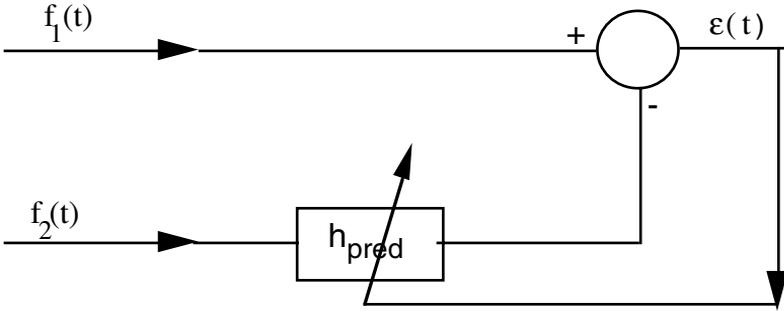
Figure 3.23: The idea of adaptive prediction filter is illustrated in the above figure. The prediction error modifies the filter so as to reduce the prediction error.

§3.6 **Exercises:**

1. A tapered fan filter is defined as [3],

$$H(u,\omega) = \frac{1}{|\omega|} rect(\frac{u}{\omega}) \bullet rect(\frac{u}{2\omega})$$

where $\bullet$ stands for the convolution in u. Sketch the filter frequency response function. Compute the impulse response function.

2. In the weighted least squares filter design the minimum mean square error, which is given by

$$\varepsilon_{min}^2 = \frac{1}{4\pi^2} \int\limits_{-\pi}^{+\pi}\int \left|\frac{WH - [WH]_{finite}}{W}\right|^2 dudv,$$

becomes independent of the weight function as the filter size increases, ideally at infinity.

3. The output of a ULA with its sensors spaced at one meter apart is sampled at the rate of 5kHz (Nyquist rate). A quadrant filter is desired with upper and lower cut off speeds 7 km/sec and 3 km/sec, respectively. Sketch the pass regions including the aliased part, if any.

4. The following are the low pass 1D filter coefficients:

h(0)=0.52
h(1)=0.3133176
h(2)=-0.01808986
h(3)=-0.09138802
h(4)=0.01223454
h(5)=0.04000004

h(6)=-0.001945309
h(7)=-0.014112893

Compute the coefficients b_k, $k = 0, 1, \ldots, N$ appearing in (3.26b) on page 168 and then, using these coefficients, evaluate a circular and a fan shaped 2D filters.
5. A UCA is split into two UCAs. The first UCA has all even sensors and the second UCA has all odd sensors. Show that the response of the first UCA may be obtained by rotating the response of the second UCA through an angle equal to angular separation between the sensors. Using this property give a heuristic explanation on the behaviour of the sidelobe as a function of the number of sensors (see p. 180).

References

1. H. C. Andrews and B. R. Hunt, *Digital Image Restoration*, Prentice-Hall, Englewood Cliffs, NJ, 1977.
2. P. S. Naidu and M. P. Mathew, *Geophysical Potential Field Analysis- A Digital Signal Processing Approach*, Advances in Geophysical Exploration vol. 5, Elsevier Science Publishers, Amsterdam, 1998.
3. J. P. Fail and G. Grau, Les filtere en eventail, Geophy. Prosp., vol. 11, pp. 131-163, 1963.
4. P. Embree, J. P. Burg, and M. M. Backus, Wide-band velocity filtering- The pie slice process, Geophysics, vol. 28, pp. 948-974, 1963.
5. S. Treitel, J. L. Shanks and C. W. Frasier, Some aspects of fan filtering, Geophysics, vol. 32, pp. 789-800, 1967.
6. Soo-Chang Pei and Sy-Been Jaw, Two dimensional general fan-type FIR digital filter design, Signal Processing (Eurosip), vol. 37, pp. 265-274, 1994
7. R. A. Wiggins, ω-k filter design, Geophysical Prospecting, vol. 14, pp. 427-440, 1966.
8. J. G. Proakis and D. G. Manolakis, *Digital Signal Processing* (chapter eight), Second Edition, Prentice-Hall of India Private Ltd, New Delhi, 1995.
9. J. H. McClellan, The design of two dimensional digital filter by transformation, Proc. of 7th Annual Princeton Conf on Information Sciences and Systems, pp. 247-251, 1973.
10. M. Abramowitz and I. A. Stegun (Editors), *Handbook of Mathematical Functions*, National Bureau of Standards, Applied Mathematics Series 55, 1964.
11. J. H. McClellan, T. W. Parks, Equiripple approximation of fan filter, Geophysics, vol. 37, pp. 573-583, 1972.
12. D. E. Dudgeon, Fundamentals of digital array processing, Proc. IEEE, vol. 65, pp. 898-905, 1975.
13. N Wiener, Extrapolation, interpolation and smoothing of stationary time series, The MIT Press, Cambridge, MA, 1949.

14. N. Levinson, The Wiener RMS (root mean square) error criterion in filter design and prediction, J. of Mathematics and Physics, vol. 25, pp. 261-278, 1947 (see also Appendix B in [13]).

15. J. Durbin, The fitting of time series models, Revue de l'Instutute International de Statistique, vol. 28, pp. 233-243, 1960.

16. R. A. Wiggins and E. A. Robinson, Recursive solution to the multichannel filtering problem, Journal of Geophysical Research, vol. 70, pp. 1885-1891, 1965.

17. P. E. Green, E. J. Kelly, Jr., and M. J. Levin, A comparison of seismic array processing methods, Geophy. J. Roy Soc, vol. 11, pp. 67-84, 1966.

18. S. M. Kay, Modern Spectrum Analysis, Prentice-Hall, Englewood Cliffs, NJ, p. 24, 1989.

19. C. Usha Padmini and Prabhakar S. Naidu, Circular array and estimation of direction of arrival of a broadband source, Signal Processing, vol. 37, pp. 243-254, 1994.

20. R. C. Sengbush and M. R. Foster, Design and application of optimal velocity filters in seismic exploration, IEEE Trans G21, pp. 648-654, 1972.

21. A. Cantoni and L. C. Godara, Performance of a postbeamformer interference canceller in the presence of broadband directional signals, J. Acoust. Soc. Am., vol. 76, pp. 128-138, 1984.

22. L. C. Godara, A robust adaptive array processor, IEEE Trans., vol. CAS-34, pp.721-730, 1987.

23. C. W. Therrien, *Discrete random signals and statistical signal processing*, Prentice-Hall, Englewood Cliffs, NJ, 1992. (1D Wiener filter, p. 422).

24. B. Widrow and J. D. Stearms, Adaptive Signal Processing, Prentice-Hall, Englewood Cliffs, NJ, p. 47, 1985.

Chapter Four

Source Localization:
Frequency Wavenumber Spectrum

In this chapter we consider the most important problem in sensor array signal processing which is estimation of the coordinates of a source emitting a signal (passive localization) or a point target illuminated by an external signal (active localization). A point in three dimensional space is defined by three parameters, namely, range (r), azimuth (φ) and elevation (θ). The range is often measured by means of return time of travel in active systems and by means of time delays measured at a number of sensors in passive systems. The azimuth and elevation angles are obtained from the measurements of direction of arrival (DOA) by an array of sensors. A horizontal array of sensors is required for azimuth measurement and a vertical array for elevation measurement. The basic quantity used for estimation of location parameters is the frequency wavenumber (ω, k) spectrum (see chapter 2). A source is assumed to be present where there is a concentration of power. We shall describe three different methods; namely, beamformation, Capon spectrum and maximum entropy spectrum. The last two methods fall under the nonlinear category while the first method belongs to the linear category. The important difference between the linear and nonlinear methods lies in their response to an input which consists of a sum of two or more uncorrelated signals. The output of a linear method will be a sum of the spectra of input signals but the output of a nonlinear method may contain an additional cross term. In spite of this drawback the nonlinear methods have become quite popular [1].

§4.1 **Frequency Wavenumber Spectrum**:
A wavefield produced by sources in the far field region may be expressed as a sum of plane waves with random phase (see plane wave decomposition in chapter 1, p. 13). The quantity of interest is power (or energy when transient waves are involved) as a function of azimuth and elevation. This is the frequency wavenumber (ω, k) spectrum which we had introduced in chapter 2. As the array is of limited size, the task of estimating the frequency wavenumber spectrum becomes too ambitious. What can be estimated with a reasonable certainty is the spectral matrix by treating the array output as a multichannel time series. The spectral matrix is indeed related to the frequency wavenumber spectrum. Fortunately, for the purpose of source localization it is enough if we can accurately estimate the spectral matrix.
4.1.1 <u>Spectral Representation of the Wavefield</u>: As in chapter 2 we shall model the wavefield as a stochastic process; hence the basic tool will be the spectral representation of the wavefield,

$$f(x,y,z,t) = \frac{1}{(2\pi)^3} \int_{-\pi}^{+\pi}\!\!\int\!\!\int dF(\omega,u,v)e^{j(\omega t - ux - vy - \sqrt{k^2 - s^2}z)}$$

where $k = \dfrac{\omega}{c}$ and $s = \sqrt{u^2 + v^2}$. Using the stochastic properties of the spectral representation of a wavefield the frequency wavenumber (ω,k) spectrum may be given by

$$\frac{1}{(2\pi)^3} S_f(\omega,u,v)d\omega dudv$$

$$= E\left\{ \frac{1}{(2\pi)^3} dF(\omega,u,v) \frac{1}{(2\pi)^3} dF^*(\omega,u,v) \right\} \quad (4.1)$$

Note that $S_f(\omega,u,v) \geq 0$ and it satisfies symmetry relations

$$S_f(\omega,+u,-v) = S_f(\omega,-u,+v)$$
$$S_f(\omega,u,v) = S_f(\omega,-u,-v)$$

Further, the (ω,k) spectrum, for a propagating wavefield, must satisfy a condition, $S_f(\omega,u,v) = 0$ for $\sqrt{u^2 + v^2} > k$. The spectrum is also defined as the Fourier transform of the spatio-temporal covariance function,

$$S_f(\omega,u,v) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty}\!\!\int\!\!\int c_f(\bar{x},\bar{y},\tau)e^{j(\omega\tau - u\bar{x} - v\bar{y})} d\bar{x}d\bar{y}d\tau \quad (4.2)$$

where

$$c_f(\bar{x},\bar{y},\tau) = E\left\{ f(x,y,t)f^*(x+\bar{x}, y+\bar{y}, t+\tau) \right\}$$

is the spatio-temporal covariance function of the wavefield on the z=0 surface. Consider an example of a wideband plane wavefront. Let c_x and c_y be the apparent wave speed in the x and y directions respectively. The spectrum of the wavefront is given by

$$S_f(\omega,u,v) = S_0(\omega)\delta(\omega - c_x u)\delta(\omega - c_y v) \quad (4.3)$$

where $S_0(\omega)$ is the spectrum of the temporal waveform. The spectrum of a plane wavefront given in (4.3) is a line in the (ω, k) space passing through the origin with direction cosines $\alpha = \sin\theta\cos\varphi$, $\beta = \sin\theta\sin\varphi$, and $\gamma = \cos\theta$ (see fig. 4.1). Note that $c_x = \dfrac{c}{\alpha}$ and $c_y = \dfrac{c}{\beta}$. A stochastic wavefield may be modeled as a sum of uncorrelated plane waves; therefore its spectrum will consist of a set of radial line segments.

4.1.2 <u>Aliasing</u>: In the spatial domain a wavefield is always measured at a set of discrete points (using point detectors) while in the temporal domain the wavefield is measured as a continuous function of time or over dense sample points, as required. Sampling in the spatial domain is dictated by the cost of deploying a large array of sensors but in the temporal domain the sampling rate may be as high as required, only at a marginally higher cost. Hence, the phenomenon of aliasing in the spatial domain becomes important. Consider a ULA of infinite length with interelement spacing equal to d. The highest spatial frequency beyond which there is a reflection of power resulting in aliasing is $\dfrac{\pi}{d}$. This phenomenon is illustrated in fig. 3.9 with reference to a digital filter. Let us assume that the signal has been prefiltered to limit the temporal spectrum to a band, $\pm\omega_{max}$. The effect of spatial sampling on the spectrum of a broadband plane wave is illustrated in fig. 4.2. The (ω, k) spectrum lies on a line abcd. The segments ab and cd lie outside the principal domain but reappear as c'd' and a'b' as shown. Since, in practice we have a finite array, the line spectrum will be broadened. To show this consider the Fourier transform of the output of a finite ULA.

$$
\begin{aligned}
F(\omega, k) &= \sum_{m=0}^{M-1} \int_{-\infty}^{+\infty} f(t, md) e^{-j\omega t} dt\, e^{-j\frac{2\pi}{M}mk} \\
&= \sum_{m=0}^{M-1} e^{jm(\frac{\omega}{c_x}d - \frac{2\pi}{M}k)} \int_{-\infty}^{+\infty} f_0(t) e^{-j\omega t} dt \qquad (4.4) \\
&= F_0(\omega) H(\frac{\omega}{c_x}d - \frac{2\pi}{M}k)
\end{aligned}
$$

where $H(\dfrac{\omega}{c_x}d - \dfrac{2\pi}{M}k)$ is the response function of a ULA of length M. Note that
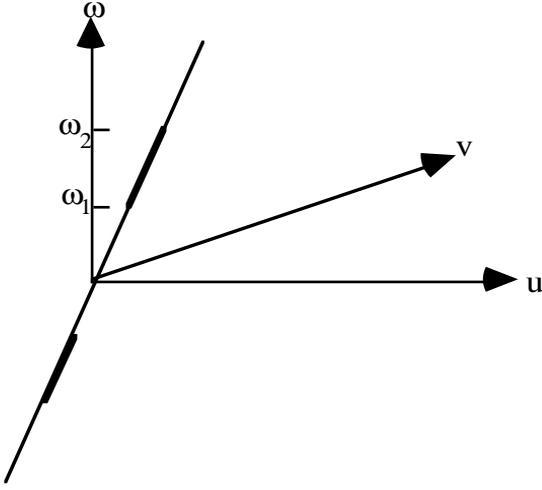
Figure 4.1: Frequency-wavenumber spectrum of a broadband plane wavefront. It lies on a line passing through the center and has direction cosines α, β and γ. The temporal frequency bandwidth is from ω_1 to ω_2.

$$H(\frac{\omega}{c_x}d - \frac{2\pi}{M}k) \to \delta(\frac{\omega}{c_x}d - u)$$

$$as \quad M \to \infty$$

where $\dfrac{2\pi}{M}k \to u$ as $M, k \to \infty$. To avoid the aliasing error in the spectrum the temporal sampling interval and sensor spacing must satisfy the relation shown in (1.34b) which, when expressed in terms of ω_{max} and d, reduces to

$$\omega_{max}\frac{d}{\pi} \le \frac{c}{\sin\theta} \tag{4.5a}$$

For fixed ω_{max}, d and c, in order to avoid aliasing error, the angle of incidence will have to satisfy the following inequality,

$$\theta \le \sin^{-1}(\frac{\pi}{\omega_{max}}\frac{c}{d}) \tag{4.5b}$$
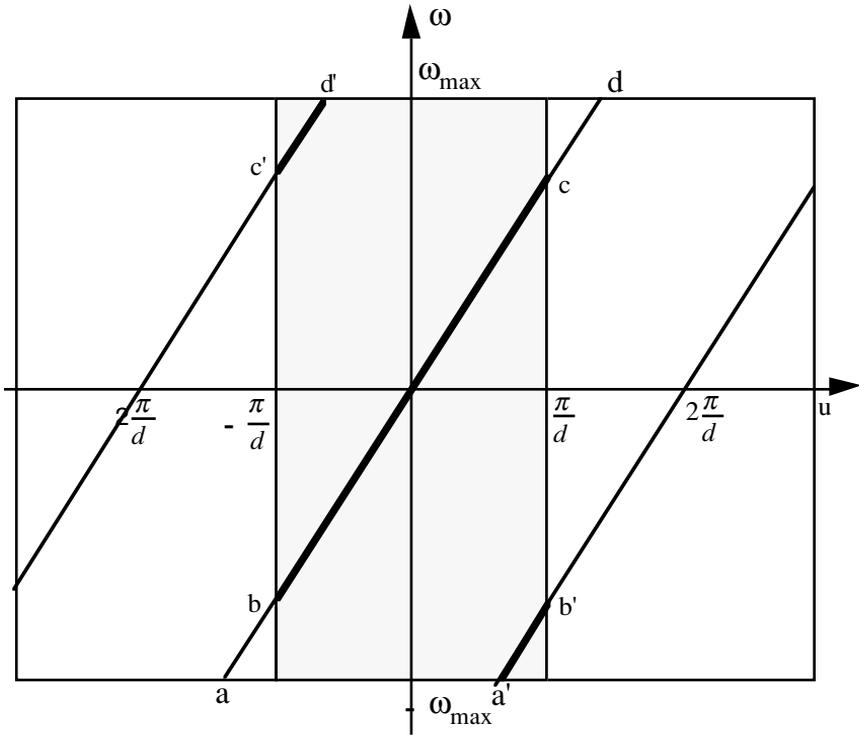
Figure 4.2: Aliasing error due to spatial sampling of a broadband plane wave.

From (4.5b) it may be seen that for $d = \dfrac{\lambda}{2}$ and $\Delta t = \dfrac{\pi}{\omega_{max}}$ there is no aliasing for any angle of incidence. Aliasing error will occur whenever the above requirements are not satisfied. As an example, consider a stochastic plane wave, incident on a ULA with 16 sensors at an angle of 45°. The sensors are 15 meters apart. The bandwidth of the waveform is $\pm 100 Hz$ and it is sampled with a sampling interval of 0.005 sec. The aliasing error is present in the top left and bottom right corners (see fig. 4.3).

4.1.3 <u>Spectral Matrix</u>: The output of an array of sensors may be treated as a collection of time series or vector time series. A spectral matrix whose elements
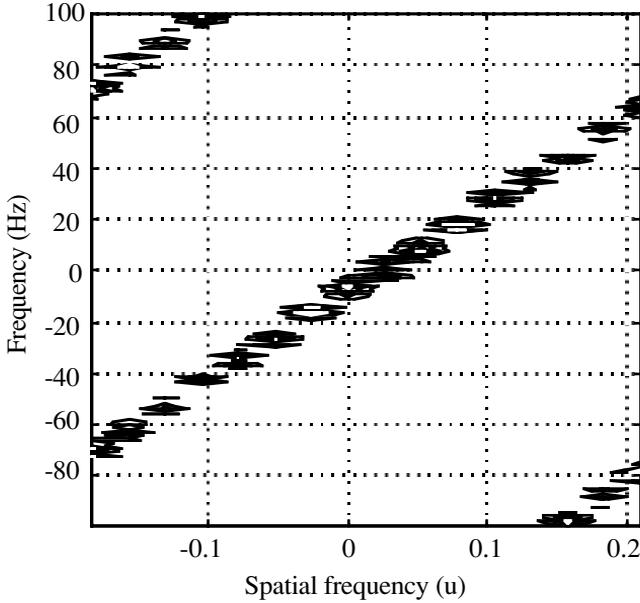
Figure 4.3: Aliasing error due to spatial undersampling. A stochastic plane wave is incident on a ULA at 45°. The ULA consists of 16 sensors spaced 15 meters apart. The wave speed is 1500m/s. The bandwidth of the waveform is $\pm 100 Hz$ and it is sampled with a sampling interval of 0.005sec.

are the spectra and cross-spectra of a pair of outputs provides a complete characterization, particularly when the outputs are Gaussian. We would like to relate a spectral matrix to the (ω, k) spectrum. Note that the output of the mth sensor, $f_m(t)$, has the following spectral representation:

$$f_m(t) = f(t, x = md) = \frac{1}{4\pi^2} \int\!\!\int_{-\infty}^{\infty} dF(\omega, u) e^{j(\omega t - umd)} \qquad (4.6)$$

Using (4.6), the cross-covariance function between two sensor outputs is given by

$$c_{mn}(\tau) = E\{f_m(t) f_n(t + \tau)\}$$

$$= \frac{1}{4\pi^2} \int\!\!\int_{-\infty}^{\infty} s_f(\omega, u) e^{j\omega\tau} e^{-j(m-n)ud} \, d\omega du \qquad (4.7)$$

Further, the spectral representation of a cross-covariance function [2] is

$$c_{mn}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{mn}(\omega)e^{j\omega\tau}d\omega \qquad (4.8)$$

Comparing (4.7) and (4.8) we obtain the following relation between elements of the spectral matrix and (ω, k) spectrum:

$$S_{mn}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_f(\omega, u)e^{jd(m-n)u}du$$

$$\qquad (4.9a)$$

$$S_{mm}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_f(\omega, u)du$$

The reverse relation, that is, (ω, k) spectrum in terms the elements of spectral matrix, is

$$S_f(\omega, u) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} S_{mn}(\omega)e^{-jd(m-n)u} \qquad (4.9b)$$

The spectral matrix has Hermitian symmetry. Additionally, for a ULA, it has toeplitz symmetry. Consider an example of a stochastic plane wave incident on a ULA. The output of the mth sensor is given by

$$f_m(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dF(\omega)e^{j\omega(t-m\frac{d}{c}\sin\theta)}$$

and the cross-covariance function between two outputs is given by

$$c_{mn}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_f(\omega)e^{j\omega(\tau-(m-n)\frac{d}{c}\sin\theta)}d\omega \qquad (4.10)$$

From (4.10) the cross-spectrum between the outputs of the mth and n^{th} sensors is given by

$$S_{mn}(\omega) = S_f(\omega)e^{-j\omega(m-n)\frac{d}{c}\sin\theta} \qquad (4.11a)$$

which may also be expressed as

$$S_{mn}(\omega) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} S_f(\omega)\delta(\omega - u)e^{-jd(m-n)u}du \qquad (4.11b)$$

where $u = \dfrac{\omega}{c}\sin\theta$. Comparing (4.11b) with (4.9) we obtain $S_f(\omega, u) = S_f(\omega)\delta(\omega - u)$.

The spectral matrix for this model has a very useful representation, that is, as an outer product of two vectors

$$\mathbf{S}_f(\omega) = S_f(\omega)\mathbf{a}(\omega)\mathbf{a}^H(\omega) \qquad (4.12a)$$

where

$$\mathbf{a}(\omega) = col\left\{1, e^{-j\omega\frac{d}{c}\sin\theta}, \ldots, e^{-j(M-1)\omega\frac{d}{c}\sin\theta}\right\}$$

is the direction vector of the incident plane wave. Eq. (4.12a) may be easily generalized for P uncorrelated sources,

$$\mathbf{S}_f(\omega) = \sum_{i=0}^{P-1} S_{f_i}(\omega)\mathbf{a}_i(\omega)\mathbf{a}_i^H(\omega)$$

which we shall express in matrix form. Define the following matrices:

$$\mathbf{S}_0(\omega) = diag\left\{S_{f_0}(\omega), S_{f_1}(\omega), \ldots, S_{f_{P-1}}(\omega)\right\}$$
$$\mathbf{A}(\omega) = \left[\mathbf{a}_0(\omega), \mathbf{a}_1(\omega), \ldots, \mathbf{a}_{P-1}(\omega)\right]$$

The spectral matrix for a case of P uncorrelated waves and uncorrelated white background noise is given by

$$\mathbf{S}_f(\omega) = \mathbf{A}(\omega)\mathbf{S}_0(\omega)\mathbf{A}^H(\omega) + \sigma_\eta^2\mathbf{I} \qquad (4.12b)$$

The columns of matrix $\mathbf{A}(\omega)$ possess an interesting structure (for a ULA only), namely, each column can be expressed as powers of a constant, $col\left\{1, \mu^1, \mu^2, \ldots, \mu^{M-1}\right\}$, where μ is a constant. Let μ_m and μ_n be the constants corresponding to m^{th} and n^{th} columns. If $\mu_m \neq \mu_n$ for all m and n, $m \neq n$ the matrix. $\mathbf{A}(\omega)$ will have full column rank, that is, equal to P. Such a matrix is also known as a Vandermonde matrix [3]. Even when the

plane waves are correlated, as in multipath propagation, the representation of a spectral matrix given by (4.12b) still hold good, with the only difference that $\mathbf{S}_0(\omega)$ is no more a diagonal matrix. The nondiagonal terms will represent cross-spectra between the outputs of sources. An important consequence of the nondiagonal character of $\mathbf{S}_0(\omega)$ is the loss of the toeplitz character (see (4.9)) of the spectral matrix. Interestingly, the toeplitz character of the spectral matrix is lost whenever the sensors of the ULA are disturbed. In a more general situation where we have a large number of random plane waves incident on a ULA the spectral matrix is a spatial covariance matrix at a fixed frequency.

<u>Propagation Matrix</u>: The matrix $\mathbf{A}(\omega)$ is partitioned as follows:

$$\mathbf{A}(\omega) = \begin{bmatrix} \mathbf{A}_1 \ _{(P \times P)} \\ \mathbf{A}_2 \ _{(M - P \times P)} \end{bmatrix}$$

Since $\mathbf{A}(\omega)$ has full column rank, there is a unique linear operator known as propagation matrix, $\Gamma (P \times M - P)$, such that $\Gamma^H \mathbf{A}_1 = \mathbf{A}_2$ which may also be written as

$$\mathbf{A}^H \begin{bmatrix} \Gamma \\ -\mathbf{I} \end{bmatrix} = \mathbf{A}^H \mathbf{Q} = 0 \qquad\qquad (4.13a)$$

It follows that $\mathbf{Q}$ spans the null space of $\mathbf{A}$. Now let us use the partitioned $\mathbf{A}(\omega)$ matrix in (4.12b)

$$\mathbf{S}_f(\omega) = \begin{bmatrix} \mathbf{G}_1 & \mathbf{H}_1 \\ \mathbf{G}_2 & \mathbf{H}_2 \end{bmatrix} = \overset{\overset{P \qquad\qquad M-P}{}}{\begin{bmatrix} \mathbf{A}_1\mathbf{S}_0\mathbf{A}_1^H & \mathbf{A}_1\mathbf{S}_0\mathbf{A}_2^H \\ \mathbf{A}_2\mathbf{S}_0\mathbf{A}_1^H & \mathbf{A}_2\mathbf{S}_0\mathbf{A}_2^H \end{bmatrix}}\begin{matrix} P \\ M-P \end{matrix} + \sigma_\eta^2 \mathbf{I}$$

Note that the spectral matrix has also been partitioned so that

$$\begin{aligned} \mathbf{G}_1 &= \mathbf{A}_1\mathbf{S}_0\mathbf{A}_1^H + \sigma_\eta^2\mathbf{I}_P, \quad \mathbf{G}_2 = \mathbf{A}_2\mathbf{S}_0\mathbf{A}_1^H \\ \mathbf{H}_1 &= \mathbf{A}_1\mathbf{S}_0\mathbf{A}_2^H, \quad \mathbf{H}_2 = \mathbf{A}_2\mathbf{S}_0\mathbf{A}_2^H + \sigma_\eta^2\mathbf{I}_{M-P} \end{aligned} \qquad (4.13b)$$

where $\mathbf{I}_P$ stands for a unit matrix of size PxP. It may be shown from (4.13b) that

$$\mathbf{G}_2 = \Gamma^H \mathbf{A}_1\mathbf{S}_0\mathbf{A}_1^H = \Gamma^H(\mathbf{G}_1 - \sigma_\eta^2\mathbf{I}_P)$$

and hence,

$$\Gamma^H = \mathbf{G}_2(\mathbf{G}_1 - \sigma_\eta^2 \mathbf{I}_P)^{-1} \qquad (4.13c)$$

Thus, the propagation matrix may be derived from the partitioned spectral matrix. The background noise variance is assumed to be known (see Exercises, 4.4). $\mathbf{Q}$ may be used to find the direction of arrival in place of eigenvectors corresponding to noise eigenvalues as in the MUSIC algorithm to be described later in chapter 5 [4, 5].

4.1.4 <u>Eigenstructure</u>: The spectral matrix possesses interesting eigenstructure. $\mathbf{S}_f(\omega)$ is a hermitian symmetric toeplitz (only for a ULA) matrix; hence its eigenvalues are real. Further, they are positive as shown below: Let $\mathbf{e}(\omega)$ be some arbitrary vector and consider a quadratic form $\mathbf{e}^H(\omega)\mathbf{S}_f(\omega)\mathbf{e}(\omega)$. It follows from (4.13a) for any $\mathbf{e}(\omega)$,

$$\mathbf{e}^H(\omega)\mathbf{S}_f(\omega)\mathbf{e}(\omega) = \sum_{i=0}^{P-1} S_{f_i}(\omega)\mathbf{e}^H(\omega)\mathbf{a}_i(\omega)\mathbf{a}_i^H(\omega)\mathbf{e}(\omega)$$

$$= \sum_{i=0}^{P-1} S_{f_i}(\omega)\left|\mathbf{e}^H(\omega)\mathbf{a}_i(\omega)\right|^2 \geq 0$$

Therefore, a spectral matrix is always positive definite or positive semidefinite (when P<M and there is no noise), that is, all its eigenvalues are positive or zero. Let $\mathbf{v}_m$, $m = 0,1,...,M-1$ be the eigenvectors of $\mathbf{S}_f(\omega)$ and the corresponding eigenvalues be λ_m, $m = 0,1,...,M-1$. From (4.13b) it follows that

$$\mathbf{v}_m^H \mathbf{S}_f(\omega)\mathbf{v}_m = \mathbf{v}_m^H \mathbf{A}(\omega)\mathbf{S}_0(\omega)\mathbf{A}^H(\omega)\mathbf{v}_m + \sigma_\eta^2 \mathbf{v}_m^H \mathbf{I} \mathbf{v}_m$$

$$\lambda_m \qquad\qquad \alpha_m \qquad\qquad\qquad \sigma_\eta^2$$

hence,

$$\lambda_m = \alpha_m + \sigma_\eta^2 \qquad (4.14a)$$

where α_m is an eigenvalue of the noise-free spectral matrix. Note that $\mathbf{A}(\omega)\mathbf{S}_0(\omega)\mathbf{A}^H(\omega)$ is a rank P matrix, as all P columns of the Vandermonde matrix $\mathbf{A}(\omega)$ are independent, and $\mathbf{S}_0(\omega)$ is a PxP diagonal matrix. Hence, the remaining M-P eigenvalues must be equal to zero, that is, $\alpha_P, \alpha_{P+1},...,\alpha_{M-1} = 0$. It therefore follows that

$$\mathbf{v}_m^H \mathbf{A}(\omega)\mathbf{S}_0(\omega)\mathbf{A}^H(\omega)\mathbf{v}_m = 0 \quad m = P, P+1,...,M-1 \qquad (4.14b)$$

Since $\mathbf{S}_0(\omega)$ is positive definite, by assumption, for (4.14b) to be valid we must have

$$\mathbf{v}_m^H \mathbf{A} = 0, \qquad \text{for} \qquad m = P, P+1, ..., M-1 \qquad (4.14c)$$

that is, $\mathbf{v}_m$, $m = P, P+1, ..., M-1$ are $\perp$ to direction vectors of all incident wavefronts. Equations (4.14a) and (4.14c) are of great importance as it will be evident in the sequel (chapter 5). Some of these properties of a spectral matrix were first noted by Nakhamkin et al. [25] in connection with the separation of seismic wavefronts.

4.1.5 <u>Frequency Wavenumber Spectrum</u>: When a large number of random plane waves are incident on an array, the (ω, k) spectrum may be computed from the spectral matrix using (4.9b). Writing (4.9b) in a matrix form for a finite array we obtain

$$S(\omega, \theta) = \mathbf{a}^H \mathbf{S}_f(\omega)\mathbf{a} \qquad (4.15a)$$

Since the above spectrum is analogous to the BT (Blackman-Tuckey) spectrum in time series literature [2], we shall call it a BT frequency wavenumber spectrum.

The (ω, k) spectrum turns into an angular spectrum when integrated over the bandwidth of the incident signal for a fixed angle of propagation

$$S(\theta) = \frac{1}{2\omega_{max}} \int_{-\omega_{max}}^{\omega_{max}} S_f(\omega, \frac{\omega}{c}\sin\theta)d\omega \qquad (4.15b)$$

In the (ω, k) domain the integration is carried out over a radial line sloping at angle θ (see fig. 4.4). A peak in the angular spectrum is an indication of wave energy arriving from a direction where the peak is found.

The angular spectrum defined in (4.15b) is also the beam power integrated over the frequency bandwidth as a function of the look angle.

4.1.6 <u>Parametric Spectrum</u>: The frequency wavenumber spectrum defined in (4.15a) assumes a plane wave model which is more appropriate in open space (see chapter 1 for wavefield representation in open and bounded space). In bounded space the wavefronts are far from planar. Such nonplanar wavefronts may be represented in terms of source location parameters measured with reference to the bounded space geometry. For example, in shallow water the
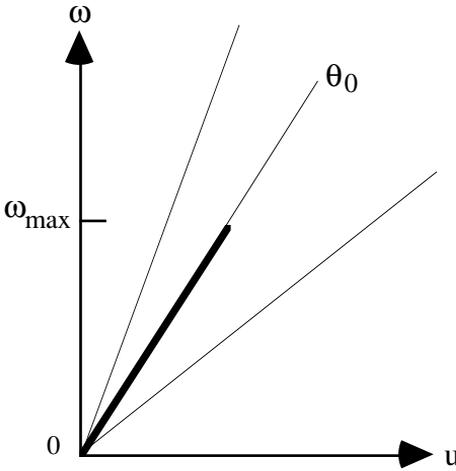
Figure 4.4: In the frequency wavenumber plane the spectrum is averaged over a series of radial lines. The spectrum of a plane wave which is incident at angle θ_0 is shown by the bold line.

source location parameters are range, depth (measured from the surface) and azimuth. In place of frequency wavenumber spectrum, where the parameters of interest are frequency and wavenumbers which depend on azimuth and elevation, we introduce a similarly defined quantity called parametric spectrum

$$S(\omega,\theta) = \mathbf{a}^H(\theta)\mathbf{S}_f(\omega)\mathbf{a}(\theta) \qquad (4.16)$$

where θ now stands for generalized location parameters and $\mathbf{a}(\theta)$ is the wavefield which the array would sense if the source were to be located at θ. $S(\omega,\theta)$ is computed over a range of values of θ spanning the entire parameter space. The actual source position is indicated by the position of the maximum in $S(\omega,\theta)$. Evidently, $\mathbf{a}(\theta)$ must be computed for the assumed geometry and the boundary conditions there on. Since the central idea is to match the computed field with the observed field, the processing is also known as matched field processing. It was first introduced by Bucker [6] in 1976 and since then a lot of research effort has been devoted toward its development as a tool in underwater detection. An account of this effort is summarized in a monograph [7]. The chief obstacle in the use of matched field processing lies in the requirement of an exact knowledge of the propagation environment for the purpose of computing the wavefield.

§4.2 **Beamformation**:

In this section we shall deal with the estimation of the (ω, k) spectrum, starting with the beamformation both in time and frequency domains including fast Fourier transform (fft) based method. Next, we describe nonlinear methods: Capon's ML method and maximum entropy method. The nonlinear methods provide higher resolution when the signal-to-noise ratio (snr) is reasonably high. The importance of spectrum estimation arises on account of the fact that a signal wavefield approaching an array of sensors in a particular direction will produce a strong peak in the frequency wavenumber spectrum. Given the peak position we can estimate the center frequency of radiation and the direction of approach, that is, the directions of arrival (azimuth and elevation).

4.2.1 <u>Beamformation</u>: A beam in a desired direction is formed by introducing delays before summation. The required delay per sensor in a ULA is equal to $\tau = \dfrac{d}{c} \sin\theta$ and in a UCA the delay for the mth sensor is equal to $\tau_m = \dfrac{a}{c} \sin\theta \cos(\dfrac{2\pi m}{M} - \varphi)$. In analog beamformation, introduction of continuously varying delays is achieved through analog delay lines, but in digital beamformation the delays can be achieved only as integral steps of sampling time units. Consider a ULA with sensor spacing d equal to $\dfrac{\lambda}{2}$ and time sampling interval, Δt, equal to $\dfrac{d}{c}$. As noted in the previous section there will be no aliasing, spatial or temporal for this choice of parameters. However, we can form just one beam, namely, for $\theta = 0$ (excluding endfire beam). Clearly, to form more beams we need more samples between two Nyquist samples. Assume that we have sampled at q times the Nyquist rate, that is, we have q equispaced samples between two Nyquist samples which will enable us to form beams at angles, θ_i, i=0,1,..., q-1, where $\theta_i = \sin^{-1}(\dfrac{i}{q})$. For example, let q=8, the beam angles are: 0°, 7.18^O, 14.48^O, 22.04^O, 30.0^O, 38.68^O, 48.59^O, $61.04^{O\cdot}$ Evidently, only a fixed number of beams can be formed for a given oversampling rate. It is not possible to form a beam in any arbitrary direction. The situation with a UCA is far more difficult as for no direction of arrival can a uniformly sampled sensor output be used for beamformation. For example, consider a UCA of 16 sensors and $\dfrac{a}{c} = 8$ time samples. The delays to be introduced in the sensor outputs, in units of the temporal sampling interval, for $\theta = 90^0$ and $\varphi = 0$ are: 8.00, 7.39, 5.66, 3.06, 0.0, -3.06, -5.65, -7.39, -8.00, -7.39, -5.66, -3.06, 0.0, 3.06, 5.66, 7.39 (rounded to second decimal place). All these delays are with respect to a

hypothetical sensor at the center of the circle. Notice that the delays are not in integral steps of the sampling interval. This leaves us with the only alternative of nonuniform sampling through interpolation of uniformly sampled sensor output. To minimize the computational load a simple linear interpolation has been suggested [8].

Narrowband: For narrowband signals the delays applied to the sensor outputs before summation may be expressed in terms of phase rotation. A narrowband signal output of the mth sensor may be represented as

$$f_m(t) \approx \frac{1}{2\pi} \int_{\omega_0 - \frac{\Delta\omega}{2}}^{\omega_0 + \frac{\Delta\omega}{2}} F_{nb}(\omega) e^{j\omega(t - m\frac{d}{c}\sin\theta_0)} \, d\omega$$

$$= \frac{1}{2\pi} \int_{-\frac{\Delta\omega}{2}}^{\frac{\Delta\omega}{2}} F(\omega_0 + \delta\omega) e^{j\left[\omega_0(t - m\frac{d}{c}\sin\theta_0) + \delta\omega(t - m\frac{d}{c}\sin\theta_0) \right]} \, d\omega \qquad (4.17a)$$

$$= e^{-j\omega_0 m\frac{d}{c}\sin\theta_0} f_{nb}(t)$$

where the subscript nb stands for narrowband. The approximation is valid when the bandwidth satisfies the condition, $\Delta\omega(m\dfrac{d}{c}\sin\theta_0) << 2\pi$ for all m, which implies that the time taken for a wave to sweep across the array must be much less than the inverse of the bandwidth, expressed in Hertz. In vector notation (4.17a) may be expressed as

$$\mathbf{f}(t) = \mathbf{a}_0 f_{nb}(t) \qquad (4.17b)$$

where $\mathbf{a}_0 = \left[1, e^{-j\omega_0 \frac{d}{c}\sin\theta_0}, \ldots, e^{-j\omega_0(M-1)\frac{d}{c}\sin\theta_0} \right]$ is the direction vector on the incident wavefront. The delays applied to sensor outputs may be expressed in terms of a vector dot product. Define a vector, $\mathbf{a} = \left[1, e^{-j\omega_0 \frac{d}{c}\sin\theta}, \ldots, e^{-j\omega_0(M-1)\frac{d}{c}\sin\theta} \right]$, known as the steering vector, which will rotate the phase of each sensor output by an amount equal to $\omega_0 m \dfrac{d}{c}\sin\theta$ for the mth sensor. Thus, a narrowband beam is formed in the direction θ as

$$\mathbf{a}^H \mathbf{f}(t) = \mathbf{a}^H \mathbf{a}_0 f_{nb}(t)$$

or in terms of beam power, that is, the (ω, k) spectrum is given by

$$S(\omega_0, \theta) = E\left\{\left|\mathbf{a}^H \mathbf{a}_0 f_{nb}(t)\right|^2\right\} = \left|\mathbf{a}^H \mathbf{a}_0\right|^2 \sigma_f^2 \qquad (4.17c)$$

<u>Window</u>: The sensor outputs are often weighted before summation, the purpose being to reduce the sidelobes of the response function just as in spectrum estimation where a window is used to reduce the sidelobes and thereby reduce the power leakage. As this topic is extensively covered under spectrum estimation, for example see [2,9], we shall not pursue any further. Instead, we like to explain the use of a weight vector to reduce the background noise variance or to increase the snr. Let us select a weight vector, $\mathbf{w}$, such that the signal amplitude is preserved but the noise power is minimized.

$$\mathbf{w}^H \mathbf{a}_0 = 1 \text{ and } \mathbf{w}^H \mathbf{c}_\eta \mathbf{w} = \min \qquad (4.118a)$$

where $\mathbf{c}_\eta$ is the noise covariance function. The solution to the constrained minimization problem in (4.18a) results in

$$\mathbf{w} = \frac{\mathbf{c}_\eta^{-1} \mathbf{a}_0}{\mathbf{a}_0^H \mathbf{c}_\eta^{-1} \mathbf{a}_0} \qquad (4.18b)$$

It may be observed that for spatially white noise $\mathbf{c}_\eta = \sigma_\eta^2 \mathbf{I}$ and therefore $\mathbf{w} = \dfrac{\mathbf{a}_0}{M}$. In other words, the weights are simply phase shifts or delays as in beamformation. The variance of the noise in the output is equal to $\sigma_{\hat{\eta}}^2 = \sigma_\eta^2/M$.

A weight vector may be chosen to maximize the snr. The output power of an array with a weight vector $\mathbf{w}$, when there is no noise, is given by $\mathbf{w}^H \mathbf{c}_s \mathbf{w}$ and when there is no signal, the output power is $\mathbf{w}^H \mathbf{c}_\eta \mathbf{w}$. We select that weight vector which will maximize the output power ratio

$$\frac{\mathbf{w}^H \mathbf{c}_s \mathbf{w}}{\mathbf{w}^H \mathbf{c}_\eta \mathbf{w}} = \max \text{ (with respect to } \mathbf{w})$$

The solution is simply the generalized eigenvector corresponding to the largest eigenvalue of the pencil matrix $\begin{bmatrix} \mathbf{c}_s, \mathbf{c}_\eta \end{bmatrix}$[10]. There is a large body of knowledge on how to obtain an optimum weight vector which meets different types of constraints. A brief review of the relevant literature is given in [10].

Rayleigh Resolution: When two wavefronts are simultaneously incident on an array, naturally we would like to measure their directions of arrival. For this to be possible the spectrum given by (4.17c) must show two distinct peaks. Let $f_{nb_1}(t)$ and $f_{nb_2}(t)$ be two uncorrelated signals incident at angles θ_1 and θ_2, with the center frequencies being the same for both signals. The beam power is given by

$$s(\omega_0, \theta) = \left| \mathbf{a}^H \mathbf{a}_1 \right|^2 \sigma_{f1}^2 + \left| \mathbf{a}^H \mathbf{a}_2 \right|^2 \sigma_{f2}^2$$

In order that each signal gives rise to a distinct peak, $\left| \mathbf{a}^H \mathbf{a}_1 \right|^2 \sigma_{f1}^2$ when plotted as a function of θ should not overlap with $\left| \mathbf{a}^H \mathbf{a}_2 \right|^2 \sigma_{f2}^2$. A condition for nonoverlap is necessarily arbitrary as the array response to an incident wavefront is strictly not limited to a fixed angular range The Rayleigh resolution criterion states that two wavefronts are resolved when the peak of the array response due to the first wavefront falls on the first zero of the array response due to the second wavefront. The first zero is located at an angle, $\sin^{-1} \dfrac{\lambda}{Md}$, away from the direction of arrival (broad side). Thus, two wavefronts are resolved, according to the Rayleigh resolution criterion when their directions of arrival differ by $\sin^{-1} \dfrac{\lambda}{Md}$. An example of resolution is shown in fig. 4.5. For a UCA we can derive a simple expression when it is fully populated. In this case, its response function is a Bessal function of 0th order (see eq. 2.43b). The first zero of the Bessal function of 0th order is at 2.45. Two wavefronts are said to be resolved according to the Rayleigh resolution criterion when the angular separation is greater than $\sin^{-1}(\dfrac{1.225\lambda}{\pi a})$. Let us compare the resolution properties of a ULA and a UCA having *equal* aperture, for example, a 16 sensor ULA with 7.5λ aperture and the corresponding UCA with a radius equal to 3.75λ but fully populated with more than 32 sensors. The relative performance
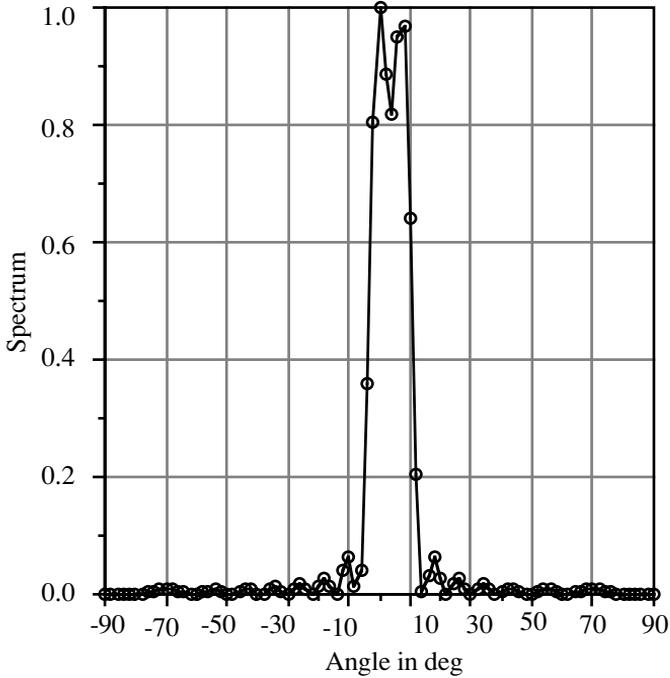
Figure 4.5: Two uncorrelated wavefronts with DOA's, 0° and 7.18^o, are incident on a 16-sensor ULA. The waves are clearly resolved. The DOAs were chosen to satisfy the Rayleigh resolution criterion.

No of Sensors	Rayleigh Resolution Angle in deg.(ULA)	Rayleigh Resolution Angle in deg.(UCA)
4	30	30.61
8	14.48	12.67
16	7.18	5.84
32	3.58	2.83
64	1.79	1.39

Table 4.1: The Rayleigh resolution angle as a function of the number of sensors (ULA with $\dfrac{\lambda}{2}$ sensor spacing).

is shown in table 4.1. The performance of the UCA is marginally better than that of the ULA. Beamformation in the frequency domain requires the 2D Fourier transform. For a fixed temporal frequency, the magnitude of the spatial Fourier transform coefficients is related to the power of a wave coming from a

direction which may be computed from the spatial frequency number, $\theta = \sin^{-1}(\frac{k}{M})$ (see chapter 2, page 80) where k is the spatial frequency number. Here too, only a finite number of fixed beams are formed. This number is equal to the number of sensors. However, the discrete Fourier transform allows interpolation between fixed beams through a simple means of padding zeros or placing dummy sensors giving no output. Consider an example of a wavefront incident at an angle of $21.06^{o\cdot}$ on a 16 sensor array. The output is subjected to the temporal Fourier transform. The spatial Fourier transform is performed before and after padding zeros. In fig. 4.6a the spatial Fourier transform before padding is shown. The peak appears at frequency number 6 corresponding to an angle of 22.02^{o} ($\sin^{-1}(6/16)$). Next, the sequence is padded with 48 zeros before Fourier transformation. The result is shown in fig. 4.6b where the peak appears at frequency number 23 corresponding to an angle 21.06^{o} ($\sin^{-1}(23/64)$) which is the correct figure. Note that the correct peak position lies between frequency numbers 5 and 6 (closer to 6). By padding zeros we are able to interpolate between the frequency numbers 5 and 6 and are thus able to capture the peak at its correct position. Further, the peak is better defined but the peak width remains unchanged. It may be emphasized that by introducing dummy sensors (zeros) we cannot achieve higher resolution.

<u>Sources of Error</u>: In practical beamformation we encounter several sources of phase errors such as those caused by sensor position errors, variable propagation conditions, sensor and associated electronics phase errors, quantization error in the phase shifter, etc. The array response is highly prone to such phase errors. Nominally, the array response may be expressed as

$$H(\omega) = \mathbf{a}^{H}(\omega\frac{d}{c}\sin\theta)\mathbf{a}(\omega\frac{d}{c}\sin\theta_{0})$$

where θ_{0} is DOA of the incident wave and θ is the steering angle. We shall model two types of phase errors, namely, those caused by position errors and phase errors caused by all other sources lumped into one. The corrupted direction vector has the following form:

$$\tilde{\mathbf{a}} = col\left\{e^{-j\phi_{0}}, e^{-j(\omega\frac{d+\Delta d_{1}}{c}\sin\theta_{0}+\phi_{1})}, \ldots, e^{-j(\omega(M-1)\frac{d+\Delta d_{M-1}}{c}\sin\theta_{0}+\phi_{M-1})}\right\} \quad (4.19)$$

where Δd_{1} is the position error of the ith sensor and ϕ_{1} is the phase error. We have assumed that the first sensor is a reference sensor and hence there is no
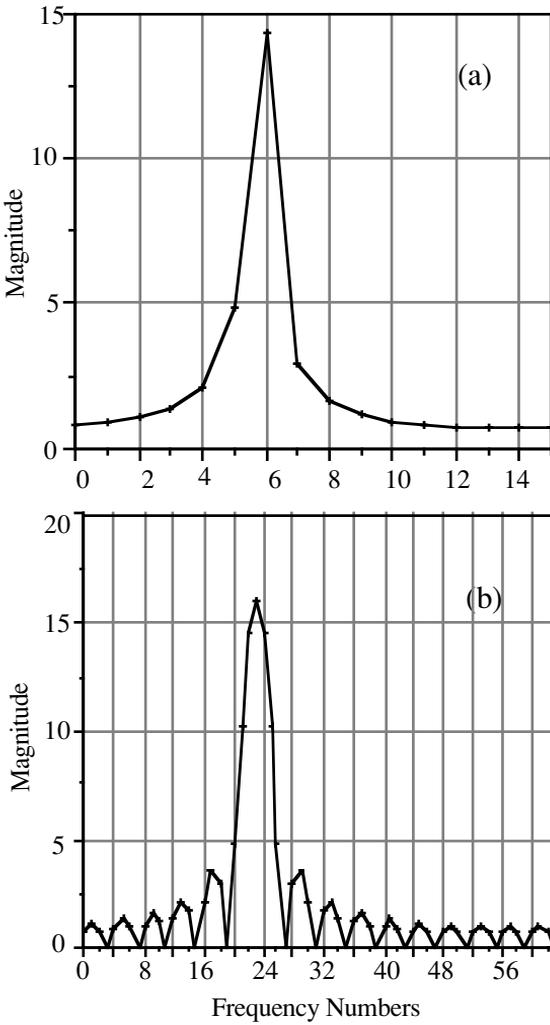
Figure 4.6: The role of padding zeros or introducing dummy sensors is to correctly position the peak (a) before padding zeros and (b) after padding zeros. The beam width remains unchanged.

position error. We shall illustrate the effect of position and phase errors on the array response function. We assume that the ULA has 16 sensors which are equispaced but with some position error. Let $d = \dfrac{\lambda}{2}$ and Δd be a uniformly
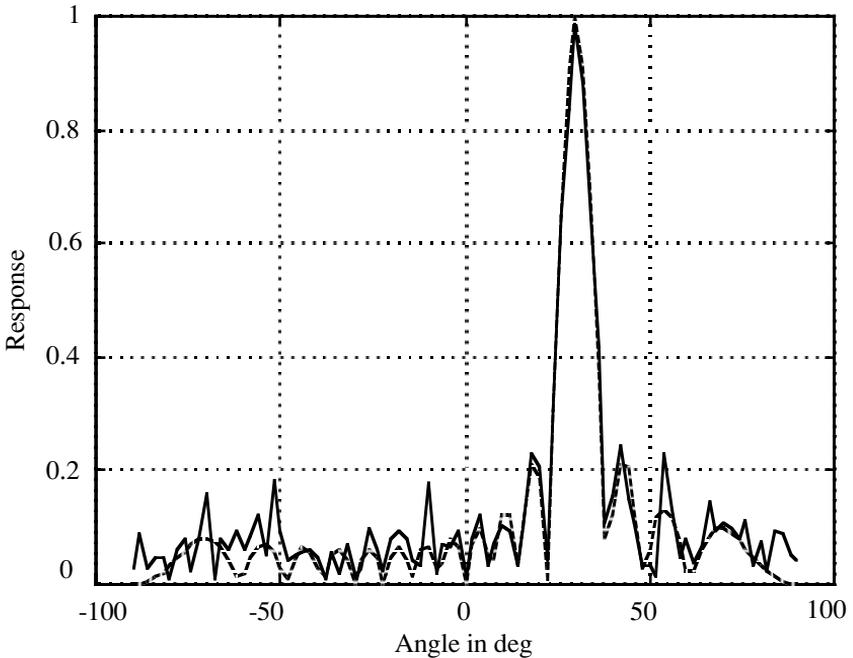
Figure 4.7: Response of a ULA with position errors which are uniformly distributed in the range $\pm \dfrac{\lambda}{4}$ (solid curve). Compare this with the response of the ULA without any position errors (dashed curve).

distributed random variable in the range $\pm \dfrac{\lambda}{16}$. The resulting response is shown in fig. 4.7. The array response due to phase errors, caused by other factors, is shown in fig. 4.8. The phase errors seem to cause less harm compared to the position errors. The sensor position and phase errors largely affect the sidelobe structure of the response function while the main lobe position and the width remain unchanged.

4.2.2 <u>Broadband Beamformation</u>: Beamformation with a broadband signal can also be written in a form similar to that for a narrowband signal. We must first Fourier transform (temporal) the broadband signal output from each sensor and treat each Fourier coefficient as a Fourier transform of a narrowband signal whose bandwidth is approximately equal to the inverse of the time duration of the signal. The frequency wavenumber spectrum in this case is given by
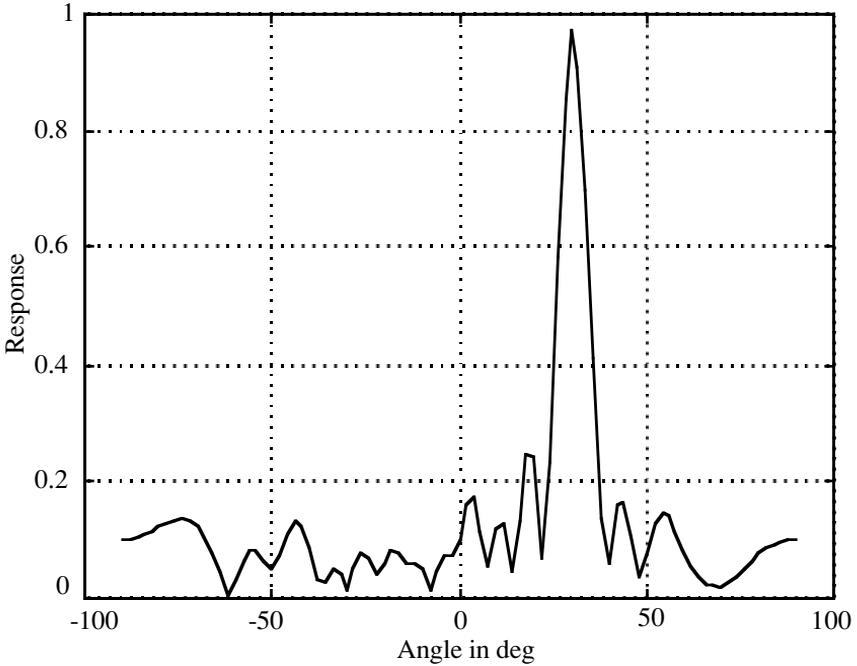
Figure 4.8: Response of a ULA with phase errors which are uniformly distributed in the range $\pm\dfrac{\pi}{4}$

$$S(\omega,\theta) = E\left\{\left|\mathbf{a}^H\mathbf{a}_0 F(\omega)\right|^2\right\} = \left|\mathbf{a}^H\mathbf{a}_0\right|^2 S_f(\omega) \qquad (4.20)$$

By integrating over the temporal frequency as shown in (4.15b) we get an estimate of the total power arriving from the direction θ (over an angular interval determined by the array). We had previously called this quantity an angular spectrum. If the objective is to estimate power received from a given direction the angular spectrum meets the requirement. On the other hand, when the aim is to estimate the waveform arriving from a given direction it is not obviously enough if we estimate the power.

Delayed Snapshots: In chapter 2 we have introduced the concept of delayed snapshots to represent a broadband wavefield. We shall make use of that representation in devising a 2D spatio-temporal filter for beamformation. A desired frequency wavenumber response is shown in fig. 3.6 where we have a passband lying between two radial lines with prescribed slopes and a horizontal

line representing the maximum temporal frequency. Let $w_{mn}, m = 0,1,...,M-1$; and $n = 0,1,...,N-1$, where M stands for the number of sensors and N for the number of delayed snapshots be the required finite 2D filter whose response is as close to the desired frequency wavenumber response as possible. A block diagram showing the filter structure is given in fig. 4.9.

We shall express the frequency wavenumber response of the filter in a matrix form. For this purpose we define the following quantities:

$$\mathbf{w} = col\left\{\begin{matrix} w_{0,0}, w_{1,0}, ..., w_{M-1,0}; w_{0,1}, w_{1,1}, ..., w_{M-1,1}; ...; \\ w_{0,N-1}, w_{1,N-1}, ..., w_{M-1,N-1} \end{matrix}\right\} \quad (4.21a)$$

$$\mathbf{A} = col\left\{\begin{matrix} 1, e^{jud}, ..., e^{ju(M-1)d}; e^{j\omega\Delta t}, e^{j(ud+\omega\Delta t)}, ..., e^{j(ud(M-1)+\omega\Delta t)}; ...; \\ e^{j\omega(N-1)\Delta t}, e^{j(ud+\omega(N-1)\Delta t)}, ..., e^{j(ud(M-1)+\omega(N-1)\Delta t)} \end{matrix}\right\}$$
$$(4.21b)$$

It is easy to show that the response function can be expressed as an inner product of two vectors defined in (4.21)

$$H(u,\omega) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} w_{mn} e^{-j(umd+\omega n\Delta t)}$$
$$= \mathbf{A}^H \mathbf{w} \quad (4.22)$$

The energy output of the filter (4.22) is given by

$$Output\ energy = \mathbf{w}^H\left[\frac{1}{4\pi^2}\int_{\omega_1}^{\omega_2}\int_{\omega/a}^{\omega/b}\mathbf{A}\mathbf{A}^H du d\omega\right]\mathbf{w} \quad (4.23a)$$

which we like to maximize with respect to $\mathbf{w}$ in relation to the total energy,

$$Total\ energy = \mathbf{w}^H\left[\frac{1}{4\pi^2}\int_{\omega_{min}}^{\omega_{max}}\int_{-\pi/d}^{\pi/d}\mathbf{A}\mathbf{A}^H du d\omega\right]\mathbf{w} \quad (4.23b)$$

where a and b are slopes of the radial lines defining the passband (see fig. 3.6), ω_2 and ω_1 are respectively the upper and the lower cut-off frequency for the beam, and ω_{max} and ω_{min} refer to the maximum and the minimum frequency present in the signal, respectively. The problem may be expressed as a problem
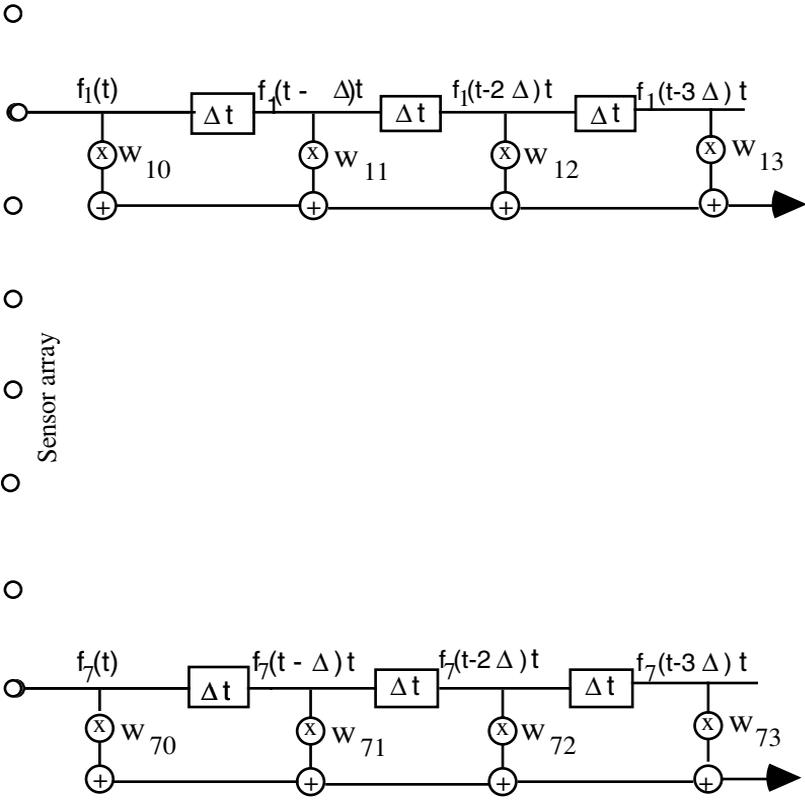
Figure 4.9: The structure of a 2D filter for broadband beamformation. M=8 and N=4.

in maximization of a ratio, $\dfrac{\mathbf{w}^H \Gamma_1 \mathbf{w}}{\mathbf{w}^H \Gamma_2 \mathbf{w}} = \lambda$, which is achieved by solving the

following generalized eigendecomposition problem:

$$\Gamma_1 \mathbf{w} = \lambda \, \Gamma_2 \mathbf{w} \qquad\qquad (4.24)$$

where $\Gamma_1 = \dfrac{1}{4\pi^2} \displaystyle\int\limits_{\omega_1}^{\omega_2} \int\limits_{\omega/a}^{\omega/b} \mathbf{A}\mathbf{A}^H du\, d\omega$ and $\Gamma_2 = \dfrac{1}{4\pi^2} \displaystyle\int\limits_{\omega_{min}}^{\omega_{max}} \int\limits_{-\pi/d}^{\pi/d} \mathbf{A}\mathbf{A}^H du\, d\omega$. The

solution is given by the eigenvector corresponding to the largest eigenvalue and the maximum relative power is equal to the largest eigenvalue.

To evaluate Γ_1 and Γ_2, we must first simplify the elements of $\mathbf{A}\mathbf{A}^H$,

$$\left[\mathbf{A}\mathbf{A}^H\right]_{pq} = e^{j(p_1 ud + p_2 \omega \Delta t)} e^{-j(q_1 ud + q_2 \omega \Delta t)}$$

$$= e^{j((p_1 - q_1)ud + (p_2 - q_2)\omega \Delta t)}$$

where $p_1 = p - Int\left[\dfrac{p}{M}\right]M$, $p_2 = Int\left[\dfrac{p}{M}\right]$ and Int[x] stands for the largest integer less than or equal to x. q_1 and q_2 are similarly defined. The elements of Γ_1 may be evaluated as in (3.12a) and those of Γ_2 are given below:

$$[\Gamma_2]_{p,q} = \frac{1}{d}\sin c((p_1 - q_1)\pi) \left[\begin{array}{c} \dfrac{\omega_{max}}{\pi}\sin c((p_2 - q_2)\omega_{max}\Delta t) \\[2mm] -\dfrac{\omega_{min}}{\pi}\sin c((p_2 - q_2)\omega_{min}\Delta t) \end{array} \right] \quad (4.25)$$

The principle of maximizing the power in the passband was first suggested in [11] in the context of optimum window for spectrum estimation of time series. Later, this principle with additional constraints on the magnitude and derivative has been applied in beamformation [12, 13]. In fig. 4.10 we show a numerical example of spatio-temporal filter for broadband beamformation. The desired response is unity in the region bounded by two radial lines and the upper and the lower frequency cutoff lines. The actual response of the maximum energy filter is contoured in the same figure. It is observed that the maximum side lobe level is about 4dB less than that in the simple quadrant filter shown in fig. 3.7.

§4.3 **Capon's ω-k Spectrum**:
We consider a stochastic plane wave incident on a ULA. Let us represent the array output in a matrix form,

$$\mathbf{f}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\mathbf{F}(\omega) e^{j\omega t} \quad (4.26a)$$

where

$$d\mathbf{F}(\omega) = d\mathbf{F}_0(\omega)\mathbf{a}(\omega \frac{d}{c}\sin\theta) + d\eta(\omega) \quad (4.26b)$$

where $d\mathbf{F}_0(\omega)$ is the generalized Fourier transform of the stochastic waveform, $\mathbf{a}(\omega \dfrac{d}{c}\sin\theta)$ is the direction vector and $d\eta(\omega)$ is the background

noise. From (4.26) and using the properties of the generalized Fourier transform we can derive an expression for the spectral matrix
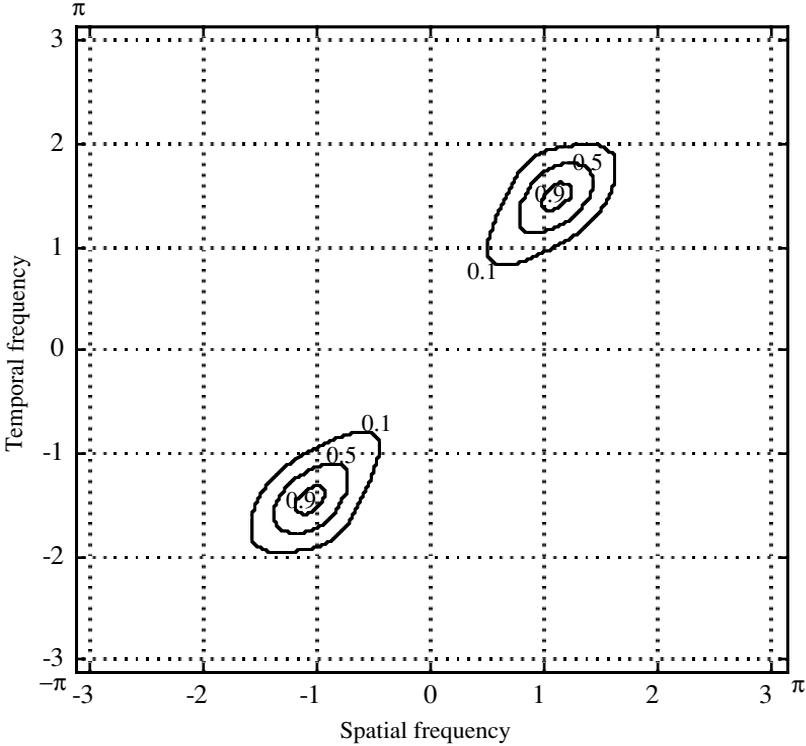
Figure 4.10: Response of filter for broadband beamformation. The slopes of the radial lines are a=1 (45°) and b=2 (63.4^{o}). $\omega_1 = 0$ and $\omega_2 = 0.6\pi$. The maximum energy in the passband is 94% ($\lambda_{\max} = 0.94$). 16 sensors and 16 delayed samples. $\Delta x = 1$ and $\Delta t = 1$.

$$\mathbf{S}_f(\omega) = S_0(\omega)\mathbf{a}(\omega\frac{d}{c}\sin\theta)\mathbf{a}^H(\omega\frac{d}{c}\sin\theta) + \mathbf{S}_\eta(\omega)$$

where both $\mathbf{S}_f(\omega)$ and $\mathbf{S}_\eta(\omega)$ are MxM matrices but $S_0(\omega)$ is a scalar. We like to find a weight vector $\mathbf{w}$ acting on the array output such that it minimizes the power output of the array and is transparent to all waves propagating through a narrow cone with a cone angle $\Delta\theta$ and its axis pointing in the direction of θ_0. Thus, the beamwidth is made intentionally wider to allow for possible variation in the direction of arrival. This model is useful

when the direction of the incoming wavefront is likely to be different from the assumed or given direction. Translated into a mathematical statement we obtain

$$\mathbf{w}^H \mathbf{S}_f(\omega)\mathbf{w} = \min$$
$$\mathbf{w}^H \Gamma \mathbf{w} = 1$$

(4.27a)

where

$$\Gamma = \frac{1}{\Delta\theta} \int_{\theta_0 - \frac{\Delta\theta}{2}}^{\theta_0 + \frac{\Delta\theta}{2}} \mathbf{a}(\omega \frac{d}{c} \sin\theta)\mathbf{a}^H(\omega \frac{d}{c} \sin\theta)d\theta \quad (4.27b)$$

and $\mathbf{S}_f(\omega)$ is the array signal spectrum under the assumption that the source bearing lies in the range $\theta_0 \pm \Delta\theta$.

The constrained minimization problem specified in (4.27) is solved by the Lagrange method,

$$\mathbf{w}^H \mathbf{S}_f(\omega)\mathbf{w} + \lambda(1 - \mathbf{w}^H \Gamma \mathbf{w}) = \min \qquad (4.28)$$

Differentiating (4.28) with respect to $\mathbf{w}$ and setting the derivative to zero we obtain

$$\mathbf{S}_f \mathbf{w} = \lambda \Gamma \mathbf{w} \qquad (4.29a)$$

or

$$\lambda^{-1}\mathbf{w} = \mathbf{S}_f^{-1}\Gamma \mathbf{w} \qquad (4.29b)$$

From (4.29b) it is clear that $\mathbf{w}$ is an eigenvector of $\mathbf{S}_f^{-1}\Gamma$ and the corresponding eigenvalue is λ^{-1}. Note that from (4.29) $\mathbf{w}^H \mathbf{S}_f \mathbf{w} = \lambda$, that is, equal to the output power of the array weighted by vector $\mathbf{w}$. In order that the array output power is minimum we must select the largest eigenvalue of $\mathbf{S}_f^{-1}\Gamma$ and the corresponding eigenvector as the weight vector, $\mathbf{w}$.

Special Case: Let $\Delta\theta = 0$, that is, the beamwidth is zero. Therefore,

$$\Gamma = \mathbf{a}(\omega \frac{d}{c} \sin\theta_0)\mathbf{a}^H(\omega \frac{d}{c} \sin\theta_0)$$

Equation (4.29b) now becomes

$$\lambda^{-1}\mathbf{w} = \mathbf{S}_f^{-1}\mathbf{a}(\omega\frac{d}{c}\sin\theta_0)\mathbf{a}^H(\omega\frac{d}{c}\sin\theta_0)\mathbf{w} \tag{4.30}$$

By premultiplying both sides by $\mathbf{a}^H(\omega\frac{d}{c}\sin\theta_0)$ we find that

$$\lambda^{-1} = \mathbf{a}^H(\omega\frac{d}{c}\sin\theta_0)\mathbf{S}_f^{-1}\mathbf{a}(\omega\frac{d}{c}\sin\theta_0) \tag{4.31a}$$

It turns out that

$$\mathbf{w} = \frac{\mathbf{S}_f^{-1}\mathbf{a}(\omega\frac{d}{c}\sin\theta_0)}{\mathbf{a}^H(\omega\frac{d}{c}\sin\theta_0)\mathbf{S}_f^{-1}\mathbf{a}(\omega\frac{d}{c}\sin\theta_0)} \tag{4.31b}$$

satisfies (4.30). We can express the array output power, which we shall call as Capon spectrum,

$$s_{Cap}(\omega,\theta_0) = \frac{1}{\mathbf{a}^H(\omega\frac{d}{c}\sin\theta_0)\mathbf{S}_f^{-1}\mathbf{a}(\omega\frac{d}{c}\sin\theta_0)} \tag{4.32}$$

Capon [14], who first suggested the above measure of spectrum, however, called it maximum likelihood spectrum. It is also known as the minimum variance distortionless response (MVDR) beamformer or a linearly constrained minimum variance (LCMV) beamformer [15]. Since θ is related to the spatial frequency, $u = \omega\frac{d}{c}\sin\theta$, $s_{Cap}(\omega,\theta)$ is indeed a (ω,k) spectrum as a function of θ or u.

4.3.1 <u>Resolution</u>: The Capon spectrum has a better resolution compared to the BT ω-k spectrum. We shall demonstrate this by considering two uncorrelated wavefronts in the presence of white noise. The spectral matrix is given by

$$\mathbf{S}_f(\omega) = s_0\mathbf{a}_0(\omega,\theta_0)\mathbf{a}_0^H(\omega,\theta_0) + s_1\mathbf{a}_1(\omega,\theta_1)\mathbf{a}_1^H(\omega,\theta_1) + \sigma_\eta^2\mathbf{I} \tag{4.33}$$

where θ_0 and θ_1 are directions of arrival and s_0 and s_1 are powers of two plane wavefronts and σ_η^2 is noise variance. The inverse of the spectral matrix in (4.33) may be computed following the procedure described in [16].

$$\mathbf{S}_f^{-1}(\omega) = \mathbf{V}^{-1} - s_1 \frac{\mathbf{V}^{-1}\mathbf{a}_1(\omega,\theta_1)\mathbf{a}_1^{H}(\omega,\theta_1)\mathbf{V}^{-H}}{1 + s_1\mathbf{a}_1^{H}(\omega,\theta_1)\mathbf{V}^{-1}\mathbf{a}_1(\omega,\theta_1)} \tag{4.34a}$$

where

$$\mathbf{V}^{-1} = \frac{1}{\sigma_\eta^2}\left[\mathbf{I} - \frac{\dfrac{s_0}{\sigma_\eta^2}\mathbf{a}_0(\omega,\theta_0)\mathbf{a}_0^{H}(\omega,\theta_0)}{1 + \dfrac{s_0}{\sigma_\eta^2}M}\right] \tag{4.34b}$$

Using (4.34a) in (4.32) we obtain the Capon spectrum for the two source model,

$$s_{Cap}(\omega,\theta) = \frac{\sigma_\eta^2}{M - \dfrac{\dfrac{s_0}{\sigma_\eta^2}|\mathbf{a}^{H}\mathbf{a}_0|^2}{1 + \dfrac{s_0}{\sigma_\eta^2}M} - \dfrac{\dfrac{s_1}{\sigma_\eta^2}\left|\mathbf{a}^{H}\mathbf{a}_1 - \dfrac{\dfrac{s_0}{\sigma_\eta^2}\mathbf{a}^{H}\mathbf{a}_0\mathbf{a}_0^{H}\mathbf{a}_1}{1 + \dfrac{s_0}{\sigma_\eta^2}M}\right|^2}{1 + \dfrac{s_1}{\sigma_\eta^2}\left(M - \dfrac{\dfrac{s_0}{\sigma_\eta^2}|\mathbf{a}_0^{H}\mathbf{a}_1|^2}{1 + \dfrac{s_0}{\sigma_\eta^2}M}\right)}} \tag{4.35}$$

where for the sake of compactness we have dropped the arguments of the vectors $\mathbf{a}$, $\mathbf{a}_0$, and $\mathbf{a}_1$. When the steering vector points to one of the sources, for example, when $\mathbf{a} = \mathbf{a}_0$

$$s_{Cap}(\omega,\theta_0) = \frac{\sigma_\eta^2}{\dfrac{M}{1 + \dfrac{s_0}{\sigma_\eta^2}M} - \dfrac{\dfrac{s_1}{\sigma_\eta^2}\left|\mathbf{a}_0^{H}\mathbf{a}_1 - \dfrac{\dfrac{s_0}{\sigma_\eta^2}M\mathbf{a}_0^{H}\mathbf{a}_1}{1 + \dfrac{s_0}{\sigma_\eta^2}M}\right|^2}{1 + \dfrac{s_1}{\sigma_\eta^2}\left(M - \dfrac{\dfrac{s_0}{\sigma_\eta^2}|\mathbf{a}_0^{H}\mathbf{a}_1|^2}{1 + \dfrac{s_0}{\sigma_\eta^2}M}\right)}} \approx s_0 + \frac{\sigma_\eta^2}{M} \tag{4.36a}$$

and when $\mathbf{a} = \mathbf{a}_1$

$$s_{Cap}(\omega,\theta_1) = \sigma_\eta^2 \frac{1 + \dfrac{s_1}{\sigma_\eta^2}\left(M - \dfrac{\dfrac{s_0}{\sigma_\eta^2}\left|\mathbf{a}_0^H\mathbf{a}_1\right|^2}{1 + \dfrac{s_0}{\sigma_\eta^2}M}\right)}{M - \dfrac{\dfrac{s_0}{\sigma_\eta^2}\left|\mathbf{a}_1^H\mathbf{a}_0\right|^2}{1 + \dfrac{s_0}{\sigma_\eta^2}M}} \approx s_1 + \frac{\sigma_\eta^2}{M} \qquad (4.36b)$$

The approximation shown in (4.36) is valid for $\left|\mathbf{a}_1^H\mathbf{a}_0\right| << M$. From the above it is clear that when the wavefronts are well resolved the peak amplitude approximately equals the power of the source. The noise power is reduced by a factor equal to the number of sensors.

We will examine the resolution power of the Capon spectrum. Consider again two equal power wavefronts incident at angles θ_0 and θ_1. The peaks corresponding to two wavefronts are resolved when a valley is formed in between them. Let $s_{Cap}(\omega,\tilde{\theta})$ be the spectrum at $\tilde{\theta}$ midway between θ_0 and θ_1. Define the ratio ρ as

$$\rho = \frac{s_{Cap}(\omega,\theta_0)}{s_{Cap}(\omega,\tilde{\theta})}$$

$$= \frac{\left\{1 + \dfrac{s_1 M}{\sigma_\eta^2}\left(1 - \alpha\left|\dfrac{\mathbf{a}_0^H\mathbf{a}_1}{M}\right|^2 - (1+\alpha)\left|\dfrac{\tilde{\mathbf{a}}^H\mathbf{a}_1}{M}\right|^2 + 2\alpha\,\mathrm{Re}\left[\dfrac{\tilde{\mathbf{a}}^H\mathbf{a}_1\mathbf{a}_1^H\mathbf{a}_0\mathbf{a}_0^H\tilde{\mathbf{a}}}{M^3}\right]\right) - \alpha\left|\dfrac{\tilde{\mathbf{a}}^H\mathbf{a}_1}{M}\right|^2\right\}}{1 - \alpha\left|\dfrac{\mathbf{a}_0^H\mathbf{a}_1}{M}\right|^2} \qquad (4.37a)$$

Assuming $\dfrac{s_1 M}{\sigma_\eta^2} >> 1$ (4.37a) simplifies to

$$\approx \frac{\left\{1 + \dfrac{s_1 M}{\sigma_\eta^2}\left(1 - \dfrac{\left|\mathbf{a}_0^H \mathbf{a}_1\right|^2}{M} - 2\dfrac{\left|\tilde{\mathbf{a}}^H \mathbf{a}_1\right|^2}{M} + 2\operatorname{Re}\left[\dfrac{\tilde{\mathbf{a}}^H \mathbf{a}_1 \mathbf{a}_1^H \mathbf{a}_0 \mathbf{a}_0^H \tilde{\mathbf{a}}}{M^3}\right]\right)\right\}}{1 - \dfrac{\left|\mathbf{a}_0^H \mathbf{a}_1\right|^2}{M}},$$

<div align="right">(4.37b)</div>

where $\alpha = \dfrac{s_1 M}{1 + s_1 M} \approx 1$ for $\dfrac{s_1 M}{\sigma_\eta^2} \gg 1$. A valley is formed iff $\rho > 1$. Let $\mathbf{a}_0 = \mathbf{a}_1 = \tilde{\mathbf{a}}$, that is, when two wavefronts merge into a single wavefront we notice that $\rho = 1$ for all snr, which means that these two wavefronts can never be resolved. Next, let $\dfrac{\left|\mathbf{a}_0^H \mathbf{a}_1\right|^2}{M} = \dfrac{\left|\tilde{\mathbf{a}}^H \mathbf{a}_1\right|^2}{M} \approx 0$ which means that the wavefronts are well separated. Then, $\rho = 1 + \dfrac{s_1 M}{\sigma_\eta^2} > 1$ except when $s_1 = 0$. The wavefronts can then always be resolved. All the above conclusions follow from common sense. We now consider two wavefronts with DOAs $\pm \dfrac{\Delta\theta}{2}$ respectively and compute ρ for different $\dfrac{s_1 M}{\sigma_\eta^2}$ and $\Delta\theta$. A plot of $\Delta\theta$ for which ρ is just greater than one as a function of array snr is shown in fig. 4.11.

4.3.2 <u>Robust Beamformation</u>: The sensitivity of beamformation to errors in the sensor position and other phase errors has been demonstrated in figs. 4.7 and 4.8. These drawbacks may be reduced through an appropriate choice of weighting coefficients. In this section we shall show how such coefficients can be obtained [17] following a constraint used in deriving Capon's filter and the associated spectrum. Let $\mathbf{w} = col\{w_0, w_1, \ldots w_{M-1}\}$ be a coefficient vector. The array response may be expressed as

$$H(\omega) = \mathbf{w}^H \mathbf{a}(\omega \frac{d}{c} \sin\theta) \tag{4.38}$$

We shall model two types of phase errors, namely, those caused by position errors and those caused by all other sources of errors lumped into a single phase
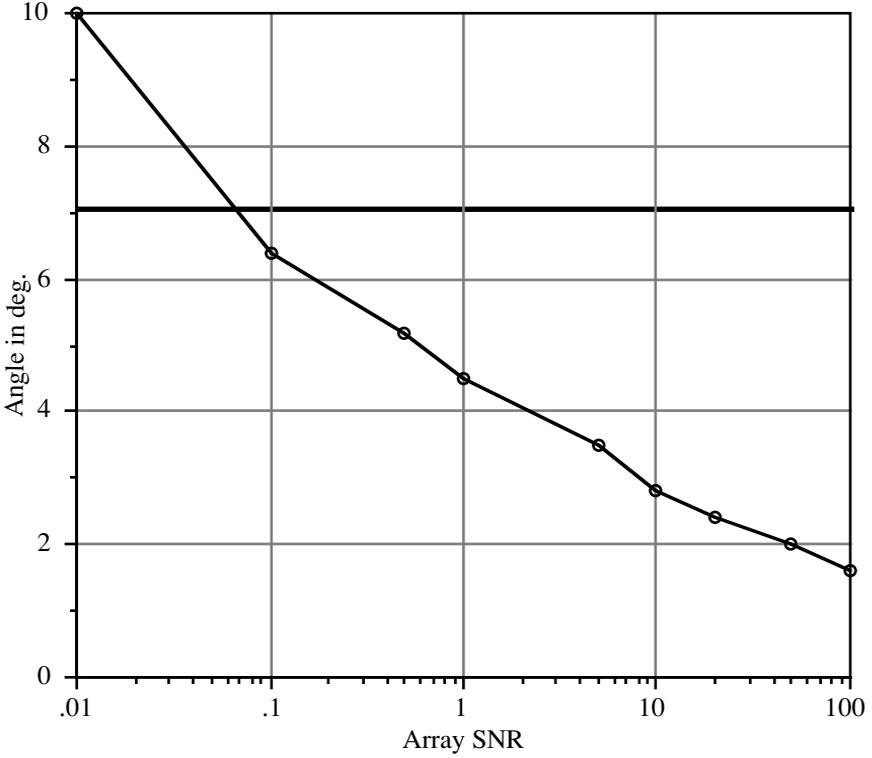
Figure 4.11: Resolution properties of the Capon spectrum as a function of array snr. A 16 sensor ULA is assumed. The angles of incidence are 30° and 30°+ angular separation as shown on the y-axis. Simple beamformation (BT ω-k Spectrum) yields a resolution of 7.18^O shown in the figure by a thick line for comparison. Resolution gain by the Capon Spectrum is possible only for high array snr.

error. The steering vector given by (4.19). We have assumed that the first sensor is a reference sensor and hence there is no position error. Let H_0 be the desired response, for example, equal to 1 and $\tilde{H}$ be corrupted response,

$$\tilde{H} = \mathbf{w}^H \tilde{\mathbf{a}} \tag{4.39}$$

The weighting coefficients are selected so as to minimize the mean square difference between H_0 and $\tilde{H}$, defined as

$$mse = \int ... \int \Omega(\zeta_0, \zeta_1, ..., \zeta_{M-1}) \left| H_0 - \mathbf{w}^H \tilde{\mathbf{a}} \right|^2 d\zeta_0 d\zeta_1 ... d\zeta_{M-1} \tag{4.40a}$$

where $\Omega(\zeta_0, \zeta_1, \ldots, \zeta_{M-1})$ is the probability density function of the random variables appearing in (4.40a). We can rewrite (4.40a) in a compact form

$$mse = \mathbf{w}^H \mathbf{Q} \mathbf{w} - (H_0 \mathbf{P}^H \mathbf{w} + H_0^H \mathbf{w}^H \mathbf{P}) + |H_0|^2 \qquad (4.40b)$$

where

$$\mathbf{P} = \int \ldots \int \Omega(\zeta_0, \zeta_1, \ldots, \zeta_{M-1}) \tilde{\mathbf{a}} d\zeta_0 d\zeta_1 \ldots d\zeta_{M-1}$$

$$\mathbf{Q} = \int \ldots \int \Omega(\zeta_0, \zeta_1, \ldots, \zeta_{M-1}) \tilde{\mathbf{a}} \tilde{\mathbf{a}}^H d\zeta_0 d\zeta_1 \ldots d\zeta_{M-1}$$

The mean square error is minimum for $\mathbf{w}_0$ which is a solution of the following equation:

$$\mathbf{Q} \mathbf{w}_0 = \mathbf{P} \qquad (4.41)$$

We shall rewrite (4.40b) in terms of $\mathbf{w}_0$

$$mse = (\mathbf{w}_0 - \mathbf{w})^H \mathbf{Q} (\mathbf{w}_0 - \mathbf{w}) + |H_0|^2 - \mathbf{w}_0^H \mathbf{Q} \mathbf{w}_0 \qquad (4.42)$$

An increased robustness in beamformation is sought by requiring that the output power of the beamformer be minimum [17],

$$\text{Output power} = \mathbf{w}^H \mathbf{S}_f \mathbf{w} = \min \qquad (4.43a)$$

subject to a quadratic constraint on the weight vector, namely,

$$(\mathbf{w}_0 - \mathbf{w})^H \mathbf{Q} (\mathbf{w}_0 - \mathbf{w}) \leq \varepsilon^2 \qquad (4.43b)$$

where $\varepsilon^2 = mse - |H_0|^2 + \mathbf{w}_0^H \mathbf{Q} \mathbf{w}_0$ is a prescribed number which represents an excess error over the minimum that can be achieved by satisfying (4.41). Note that $\mathbf{S}_f$ is a spectral matrix of the array output. Using the standard primal-dual method we can solve the constrained optimization problem [18]. The solution is given by

$$\mathbf{w} = \mathbf{w}_0 - (\mathbf{S}_f + \lambda \mathbf{Q})^{-1} \mathbf{S}_f \mathbf{w}_0 \qquad (4.44)$$

where λ is a Lagrange multiplier and it is given as a root of the following rational function:

$$\mathbf{w}_0^H \mathbf{S}_f (\mathbf{S}_f + \lambda \mathbf{Q})^{-1} \mathbf{Q} (\mathbf{S}_f + \lambda \mathbf{Q})^{-1} \mathbf{S}_f \mathbf{w}_0 = \varepsilon^2 \qquad (4.45)$$

A simple approach to solve for λ is to plot λ vs ε^2 for the given $\mathbf{S}_f$, $\mathbf{P}$ and $\mathbf{Q}$ and pick a value of λ for a prescribed ε^2. It is demonstrated in [17] that for a circular array consisting of two concentric rings with random interring spacing the array gain remains practically unaltered if the actual spacing is well within the bounds used in the design of the weight coefficients. However, the sidelobe characteristics of the response function of the weight coefficients are not known.

4.3.3 <u>High Resolution Capon Spectrum</u>: The resolution capability of Capon's frequency wavenumber spectrum may be improved by noise cancellation by subtracting an estimated white noise power from the diagonal elements of the spectral matrix [19] and stretching the eigenvalue spread of the spectral matrix [20]. A predetermined quantity is subtracted from the diagonal elements of the spectral matrix, thus increasing the ratio between the maximum and minimum eigenvalues. This process is called stretching the eigenvalue spread. The spectral matrix must, however, remain positive definite. Consider a model of P plane waves and white background noise. The spectral matrix, given in (4.12b), is subjected to stretching of the eigenvalue spread by subtracting a fixed number σ_0^2 from the diagonal elements,

$$\begin{aligned} \tilde{\mathbf{S}}_f(\omega) &= \mathbf{S}_f(\omega) - \sigma_0^2 \mathbf{I} \\ &= \mathbf{v} \Gamma_0 \mathbf{v}^H \end{aligned} \qquad (4.46)$$

where

$$\begin{aligned} \Gamma_0 &= diag\{\alpha_m + \sigma_\eta^2 - \sigma_0^2, m = 1,2,...,M-1\} \\ &= diag\{\gamma_m, m = 1,2,...,M-1\} \end{aligned}$$

We now introduce an improved Capon spectrum by using the spectral matrix whose eigenvalue spread has been stretched in (4.32) to obtain

$$\tilde{s}_{Cap}(\omega,\theta) = \frac{1}{\mathbf{a}^H(\omega\frac{d}{c}\sin\theta)\tilde{\mathbf{S}}_f^{-1}\mathbf{a}(\omega\frac{d}{c}\sin\theta)}$$

$$= \frac{1}{\mathbf{a}^H\mathbf{v}\Gamma_0^{-1}\mathbf{v}^H\mathbf{a}} = \frac{1}{\sum\limits_{m=0}^{M-1}\frac{1}{\gamma_m}\left|\mathbf{a}^H\mathbf{v}_m\right|^2} \qquad (4.47)$$

By selecting σ_0^2 as close to σ_η^2 as possible we can make $\gamma_m \approx 0$, m=P,P+1, ..., M-1 and $\gamma_m \approx \alpha_m$, $m = 0,1,...,P-1$. Next we note that from (4.14c) $\mathbf{a}^H\mathbf{v}_m = 0$ for m=P, P+1, ...,M-1 and when $\mathbf{a}$1 is equal to one of the direction vectors of the incident wavefronts. Conversely, when $\mathbf{a}$ does not belong to that set of direction vectors, $\mathbf{a}^H\mathbf{v}_m \neq 0$. Let us rewrite (4.47) as

$$\tilde{s}_{Cap}(\omega,\theta) = \frac{1}{\sum\limits_{m=0}^{P-1}\frac{1}{\gamma_m}\left|\mathbf{a}^H\mathbf{v}_m\right|^2 + \sum\limits_{m=P}^{M-1}\frac{1}{\gamma_m}\left|\mathbf{a}^H\mathbf{v}_m\right|^2} \qquad (4.48)$$

In the denominator of (4.48) the second term dominates whenever $\mathbf{a}$ does not belong to the set of direction vectors of the incident wavefronts and vice versa. Hence,

$$\tilde{s}_{Cap}(\omega,\theta) = s_{Cap}(\omega,\theta), \quad \theta = \theta_i \ i = 0,1,...,P-1$$

and

$$\tilde{s}_{Cap}(\omega,\theta) \approx 0, \qquad\qquad \theta \neq \theta_i$$

As an example, consider a single source, that is, P=1, then $\left|\mathbf{a}^H\mathbf{v}_0\right|^2 = M$ and $\alpha_0 = Mp_0$. On account of (4.14c) $\mathbf{a}^H\mathbf{v}_m = 0$ (but $\gamma_m \neq 0$) for m=1, 2,..., M-1. Therefore, $\tilde{s}_{Cap}(\omega,\theta_0) = p_0$ (power of the incident wavefront). For all other values of θ $\tilde{s}_{Cap}(\omega,\theta) \approx 0$. To demonstrate the resolution power of the high resolution Capon spectrum we consider a 16 sensor ULA and two wavefronts incident at angles 30$^{\text{o}}$ and 35$^{\text{o}}$ and 0dB snr. The resolution of the Capon spectrum as shown in fig. 4.12 has dramatically improved when the eigenvalue spread is increased from 30 to 1000. The above result is for an error-free spectral matrix, that is, with infinite data. To study the effect of finite data
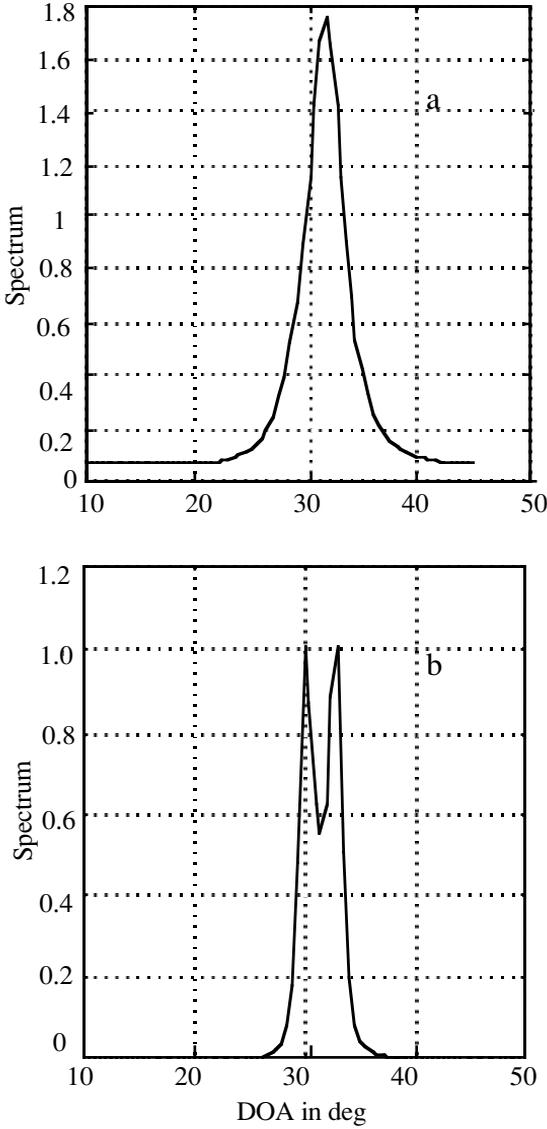
Figure 4.12: (a) Capon spectrum. (b) High resolution Capon spectrum. Eigenvalue spread is 1000. 16 sensor ULA with sensor spacing of $\lambda/2$ and two uncorrelated wavefronts incident at 30° and 33° were assumed. The high resolution Capon spectrum yields correct amplitude and DOA information.

a numerical experiment was carried out [20] on an eight sensor ULA with two wavefronts incident at angles 45° and 53° in presence of white background

noise (-5dB). The spectral matrix was computed by averaging over ten independent segments. The Capon spectrum is barely able to resolve the peaks but the high resolution Capon spectrum shows two clear peaks (see fig. 2 in [20]). The minimum resolvable angular separation between two uncorrelated wavefronts as a function of snr is shown in fig. 4.13. The spectral matrix was computed using 10 data segments. The eigenvalue spread was set at 2000 (note that the Rayleigh resolution is equal to 16.6^o). Further, a study of the bias, the mean square error (mse) and the probability of resolution was also carried out in the above experiment. The results are summarized in table 4.2. The total bias is a sum of the bias (magnitude) in both peaks. Similarly, the total standard deviation is a sum of the standard deviations of both peaks. The probability of resolution was computed by noting the number of times the peaks were clearly resolved in a hundred trials. Finally, by increasing the eigenvalue spread the spectral peaks become sharper but soon instability sets in while inverting the modified spectral matrix.

§4.4 **Maximum Entropy ω-k Spectrum**:

The maximum entropy (ME) spectrum is another example of the nonlinear spectrum estimation method, originally developed for time series. The basic idea is to find a frequency wavenumber spectrum which is consistent with the observed spectral matrix of the array signal but is maximally noncommittal on the wavefield which has not been observed simply because the array happened to be finite. The requirement of being maximally noncommittal is translated into maximization of entropy [2]. As in time series, there is an implied prediction of the wavefield outside the array aperture. Many of the properties of the ME spectrum of time series naturally hold good in the wavefield analysis. We shall elaborate in this section some of these aspects of the ME spectrum.

4.4.1 <u>Maximum Entropy</u>: We shall start with an assumption that the spectral matrix of the array signal is available. An element of the spectral matrix and the frequency wavenumber spectrum are related through (4.9a), reproduced here for convenience

$$S_{mn}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_f(\omega, u) e^{-jd(m-n)u} du \qquad (4.9a)$$

The entropy gain is given by

$$\Delta H = \frac{1}{4\pi} \int_{-\infty}^{\infty} \log S_f(\omega, u) du \qquad (4.49)$$
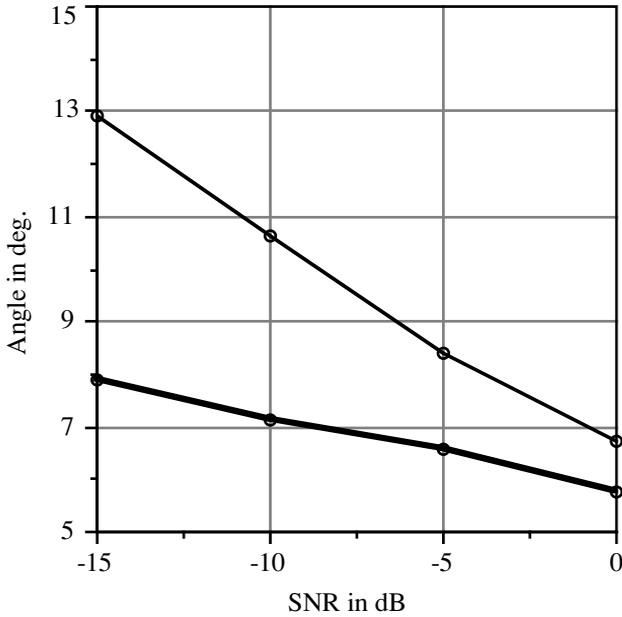
Figure 4.13: Resolution capability of high resolution Capon spectrum as a function of snr. Thick line: High resolution Capon spectrum and Thin line: Capon spectrum.

Method	Total bias	Total std. dev.	Probability of resolution
Capon spectrum	1.42^0	0.86^0	58.3%
High resolution Capon spectrum	0.69^0	0.80^0	92.2%

Table 4.2: Comparison of performance of Capon spectrum and its high resolution version. Parameters: 8 sensor ULA, two wavefronts with angular separation = 6^0 incident on broadside, snr=0 dB, number of segments used =50, number of trials=500 and eigenvalue spread= 2000.

It is proposed to find $S_f(\omega, u)$ which satisfies (4.9a) and maximizes the entropy gain (4.49). We shall keep the temporal frequency ω fixed throughout. The above statement of estimation using the principle of maximum entropy is exactly similar to that in the estimation of time series [2]. The approach to maximization is essentially the same as in the time series case. We shall only briefly outline the approach leaving out all the details which may be found in

[2]. Maximization of entropy is achieved when the spectrum can be expressed in a parametric form,

$$S_f(\omega, u) = \cfrac{1}{\displaystyle\sum_{p=-p_0}^{p=p_0} \lambda_p e^{jup}}$$

(4.50)

where λ_p, $p=0, \pm 1, \pm 2, \pm p_0$ are Lagrange coefficients. Let us further express the denominator in (4.50) as

$$S_f(\omega, u) = \cfrac{1}{H(u) H^H(u)}$$

$$= \cfrac{1}{\displaystyle\sum_{m=0}^{p_0} h_m e^{jum} \sum_{m=0}^{p_0} h_m e^{-jum}}$$

(4.51)

The phase of $H(u)$ is yet to be specified. We will choose $H(u)$ as a minimum phase function whose all poles and zeros are to the left of imaginary axis in the complex frequency plane. The coefficients $h_m, m = 0, 1, ..., p_0$ may be obtained by solving the following equation,

$$\mathbf{S}_f \mathbf{H} = \frac{1}{h_0} \delta$$

(4.52a)

where $\mathbf{S}_f$ is a spectral matrix, $\mathbf{H} = col\{h_0, h_1, h_2, ..., h_{p_0}\}$ and $\delta = col\{1, 0, 0, ..., 0\}$. The solution of (4.52a) is given by

$$\mathbf{H} = \frac{1}{h_0} \mathbf{S}_f^{-1} \delta$$

(4.52b)

Notice that $\mathbf{S}_f^{-1} \delta$ refers to the first column of $\mathbf{S}_f^{-1}$ and h_0 is the first element of that column. Using (4.52b) in (4.51) we can express the maximum entropy spectrum as

$$S_{ME}(\omega,u) = \frac{h_0^2}{\left| \mathbf{a}^H \mathbf{S}_f^{-1} \delta \right|^2} \tag{4.52c}$$

where $\mathbf{a} = col\{1, e^{ju}, e^{j2u}, \dots, e^{jp_0 u}\}$ is the steering vector.

We shall look at an alternate approach which will lead to an equation identical to (4.52b). At a fixed temporal frequency the output of an array (ULA) may be expressed as a sum of complex sinusoids,

$$f_m(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \sum_{p=0}^{p_0} dG_p(\omega) e^{j\omega(t - m\frac{d}{c}\sin\theta_p)} \tag{4.53a}$$

Let

$$dF_m(\omega) = \sum_{p=0}^{P} dG_p(\omega) e^{-jm\frac{\omega d}{c}\sin\theta_p}, \qquad m = 0, 1, \dots, M-1 \tag{4.53b}$$

where we have assumed that P plane wavefronts are incident on a ULA. The sources are assumed to radiate stationary stochastic but uncorrelated signals. Clearly, (4.53b) is a sum of P random spatial complex sinusoids. A sum of P random sinusoids (real) are known as a deterministic random process and it can be predicted without error from $2P$ past samples [2]. In the case of P complex sinusoids we will require P past samples for prediction. Error-free prediction is not possible when there is background noise. The prediction equation is given by

$$dF_m(\omega) + \sum_{p=1}^{P} h_p dF_{m-p}(\omega) = \eta_m \tag{4.54a}$$

where η_m is the prediction error. We express (4.54a) in a matrix form,

$$\mathbf{H}^H d\mathbf{F}_m(\omega) = \eta_m \tag{4.54b}$$

where $d\mathbf{F}_m(\omega) = col\{d\mathbf{F}_m, d\mathbf{F}_{m-1}, \dots, d\mathbf{F}_{m-p_0}\}$ and $h_0 = 1$. The prediction error is given by

$$\mathbf{H}^H \mathbf{S}_f(\omega) \mathbf{H} = \sigma_\eta^2 \tag{4.55}$$

which we wish to minimize, subject to the constraint that $h_0 = 1$ or $\mathbf{H}^H \delta = 1$. This leads to the following equation:

$$\mathbf{S}_f(\omega)\mathbf{H} = \sigma_\eta^2 \delta$$

or

$$\mathbf{H} = \sigma_\eta^2 \mathbf{S}_f^{-1}(\omega)\delta \tag{4.56}$$

which is identical to (4.52b), except for a scale factor. Using (4.56) in (4.55) the minimum prediction error is equal to $\sigma_\eta^2\big|_{\min} = \dfrac{1}{\delta^H \mathbf{S}_f^{-1}(\omega)\delta}$. The prediction filter vector corresponding to minimum error is given by

$$\mathbf{H} = \frac{\mathbf{S}_f^{-1}(\omega)\delta}{\delta^H \mathbf{S}_f^{-1}(\omega)\delta} \tag{4.57}$$

The maximum entropy spectrum defined in (4.51) may be expressed as follows:

$$S_{ME}(\omega,u) = \frac{\left(\delta^H \mathbf{S}_f^{-1}(\omega)\delta\right)^2}{\left|\mathbf{a}^H \mathbf{S}_f^{-1}(\omega)\delta\right|^2} \tag{4.58}$$

Note that the numerator of (4.58) is equal to the first element of the first column of $\mathbf{S}_f^{-1}$. The maximum entropy spectrum given by (4.52c) and the spectrum obtained by minimizing the prediction error, (4.58), are identical. Thus, the maximum entropy principle leads to a simple interpretation in the form of linear prediction.

4.4.2 <u>Resolution:</u> We shall now examine some of the properties of the maximum entropy spectrum and compare them with those of the Capon spectrum. As earlier we shall consider two wavefronts incident on a ULA. The spectral matrix is given by (4.33). Using the inversion formula (4.34) in (4.52c) we obtain an expression for the maximum entropy spectrum for two wavefronts in presence of white noise.

$$s_{ME}(\omega,\theta) = \frac{h_0^2 \sigma_\eta^4}{den} \tag{4.59}$$

where

$$\text{den}=\left|1-\frac{\frac{s_0}{\sigma_\eta^2}\mathbf{a}^H\mathbf{a}_0}{1+\frac{s_0}{\sigma_\eta^2}M}-\frac{\frac{s_1}{\sigma_\eta^2}\left[\mathbf{a}^H\mathbf{a}_1-\frac{\frac{s_0}{\sigma_\eta^2}\mathbf{a}^H\mathbf{a}_0\mathbf{a}_0^H\mathbf{a}_1}{1+\frac{s_0}{\sigma_\eta^2}M}\right]\left[1-\frac{\frac{s_0}{\sigma_\eta^2}\mathbf{a}_1^H\mathbf{a}_0}{1+\frac{s_0}{\sigma_\eta^2}M}\right]}{1+\frac{s_1}{\sigma_\eta^2}\left(M-\frac{\frac{s_0}{\sigma_\eta^2}\left|\mathbf{a}_0^H\mathbf{a}_1\right|^2}{1+\frac{s_0}{\sigma_\eta^2}M}\right)}\right|^2$$

where all symbols are as in (4.35). An example of maximum entropy of two uncorrelated wavefronts incident at 30° and 35° on a 16 sensor ULA is shown in fig 4.14a and the corresponding Capon spectrum is shown in fig. 4.14b. In both cases we have assumed white background noise and snr=1.0.

Let us now evaluate (4.59) at $\theta=\theta_0$ and $\theta=\theta_1$ under the assumption that $\left|\mathbf{a}_0^H\mathbf{a}_1\right|^2\approx0$. We obtain

$$s_{ME}(\omega,\theta_0)=h_0^2(\sigma_\eta^2+s_0M)^2 \tag{4.60a}$$

$$s_{ME}(\omega,\theta_1)=h_0^2(\sigma_\eta^2+s_1M)^2 \tag{4.60b}$$

The height of the spectral peak grows with the array signal-to-noise ratio, increasing to infinity as $sM\rightarrow\infty$. This is demonstrated in fig. 4.15 for the two wavefront model, whose spectrum is plotted in fig. 4.14.

From figs. 4.14a and 4.14b it may be conjectured that the maximum entropy spectrum has a better resolution property than that of the Capon spectrum. The depth of the valley for the maximum entropy spectrum is much larger than the one for the Capon spectrum. We have carried out a series computation to find out the minimum snr required to resolve two equal amplitude wavefronts separated by a specified angular distance. The criterion for resolution was formation of a nascent valley between the spectral peaks. While this is not a quantitative criterion it serves the purpose of comparison. The results are shown in fig. 4.16 which may be compared with fig. 4.11 for the Capon spectrum. Clearly, the maximum entropy spectrum has a better resolution capability, but its peak amplitude does not bear any simple relation to the actual spectrum value. In the time series context it was shown in [21] that the Capon spectrum and maximum entropy spectrum of different orders are related,
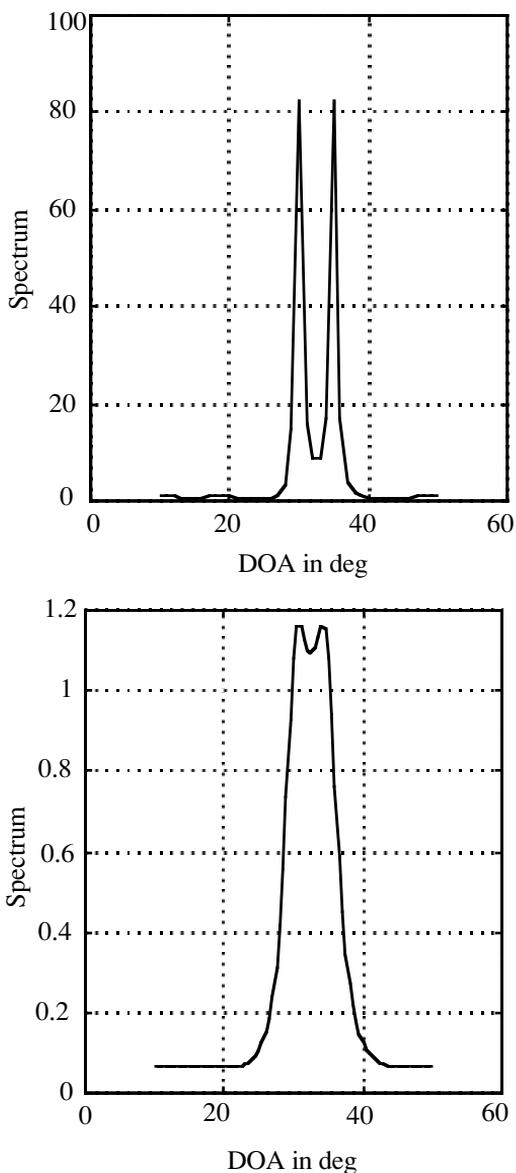
Figure 4.14: (a) Maximum entropy spectrum and (b) Capon spectrum. Two unit amplitude plane wavefronts are incident at angles 30° and 35° on a 16 sensor ULA. The amplitude of the peaks of Capon spectrum is close to the actual amplitude but the amplitude of ME spectrum is much higher. According to (4.53), valid for large separation, the computed amplitude is equal to 200.
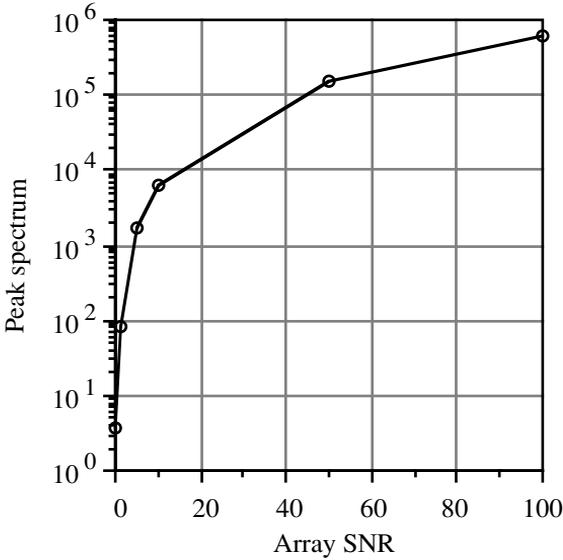
Figure 4.15: Peak spectrum as a function of array snr, that is, Msnr. Two equal amplitude uncorrelated wavefronts incident at 30° and 35° on a 16 sensor ULA.

$$\frac{1}{S_{Cap}(\omega)} = \frac{1}{M} \sum_{m=0}^{M-1} \frac{1}{S_{ME}^m(\omega)} \qquad (4.61)$$

where M stands for the size of the covariance matrix. In array processing, M stands for array size.

4.4.3 <u>Finite Data Effects</u>: So far we have tacitly assumed that the spectral matrix is known and the incident wavefield and noise confirm with the assumed model; for example, the wavefronts are planar and uncorrelated and the noise is white. In practice, however, the spectral matrix needs to be computed from the available data. Since the spectral matrix is a statistically defined quantity involving the operation of expectation, there is bound to be some error in its estimation when only finite length data is available. The effect of the errors in the spectral matrix on wavenumber spectrum has been a subject of investigation by many researchers [22, 23]. Here we shall briefly outline the important results. The mean and variance of the BT frequency wavenumber spectrum (linear), the Capon frequency wavenumber spectrum (nonlinear), are tabulated in
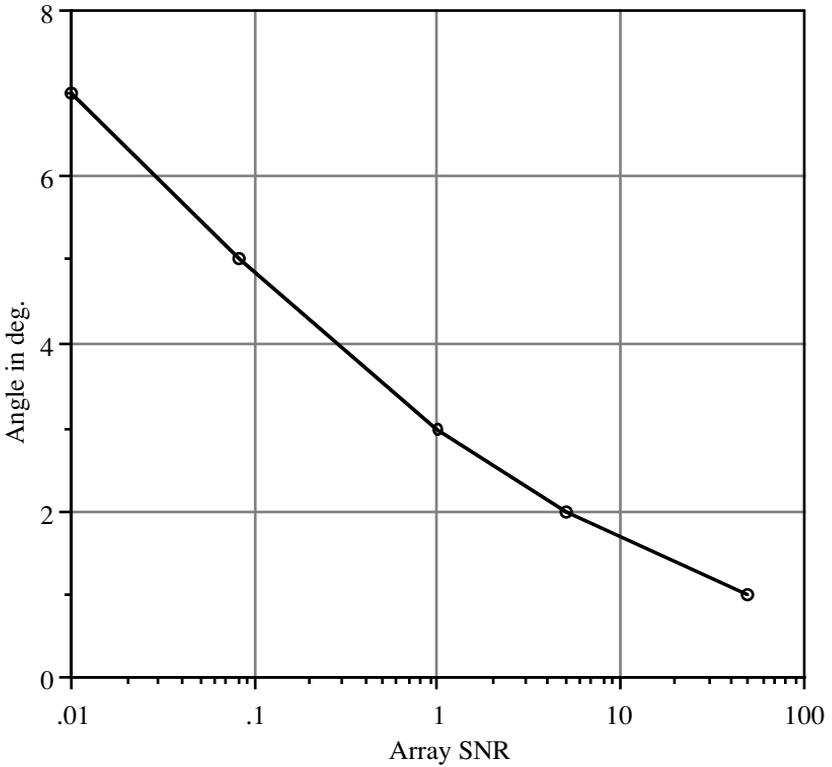
Figure 4.16: Resolution property of ME spectrum. Two equal amplitude uncorrelated wavefronts are incident on a 16 sensor ULA. The directions of arrival are 30° and 30°+ the angular separation, as shown on the y-axis.

Method	Mean	Variance
BT spectrum	$S(\omega,\theta)$	$\dfrac{S^2(\omega,\theta)}{N}$
Capon spectrum	$\dfrac{N-M+1}{N}S_{Cap}(\omega,\theta)$	$\dfrac{N-M+1}{N^2}S_{Cap}^2(\omega,\theta)$

Table 4.3 Mean and variance of frequency wavenumber spectrum estimated from finite data. N stands for number of independent snapshots and M for array size.

table 4.3. Unfortunately, we do not have simple expressions for mean and variance of ME spectrum estimates; only experimental results are available [22]. The ME spectrum is more variable and hence it needs much larger data to

stabilize. It is reported in [22] that to get a stable ME spectrum the number of snapshots must be far greater than M and $M^2 SNR$. the effect of finite data will result in (a) a loss in resolution, that is, closely spaced wavefronts cannot be resolved; (b) a shift in the position of peaks, that is, an erroneous estimate of the direction of arrival.

4.4.4 <u>Iterative Inversion of Spectral Matrix</u>: The most important computational step in Capon and ME methods is the computation of an inverse of the spectral matrix (covariance matrix for narrowband signal). The spectral matrix is computed as an average of the outer product of the frequency snapshots. The output of an array is divided into equal duration and overlapping segments and each one is Fourier transformed. A frequency snapshot of an array, analogous to the time snapshot, is an output of an array at a fixed (temporal) frequency. Let $\mathbf{F}_i(\omega)$ be the ith frequency snapshot obtained by Fourier transforming the ith segment of the array output. The spectral matrix is estimated as

$$\hat{\mathbf{S}}_f(\omega) = \frac{1}{N} \sum_1^N \mathbf{F}_i(\omega)\mathbf{F}_i^H(\omega)$$

When a new time snapshot arrives a new segment is formed with the newly received snapshot and the past snapshots to form a required length segment and then a new frequency snapshot is formed. The spectral matrix is updated by incorporating the outer product of the newly formed frequency snapshot,

$$\hat{\mathbf{S}}_f^{N+1}(\omega) = \frac{1}{N+1}\left[\sum_1^N \mathbf{F}_i(\omega)\mathbf{F}_i^H(\omega) + \mathbf{F}_{N+1}(\omega)\mathbf{F}_{N+1}^H(\omega) \right]$$

$$= \frac{N}{N+1}\hat{\mathbf{S}}_f^N(\omega) + \frac{1}{N+1}\mathbf{F}_{N+1}(\omega)\mathbf{F}_{N+1}^H(\omega)$$

(4.62a)

The recursion may be commenced with an initial value $\hat{\mathbf{S}}_f^1(\omega) = \mathbf{F}_1(\omega)\mathbf{F}_1^H(\omega)$. We can obtain a recursive expression for the (ω, k) spectrum by using (4.62a) in (4.16)

$$S^{N+1}(\omega,\theta) = \left[\begin{array}{l} \dfrac{N}{N+1}S^N(\omega,\theta) + \\[2mm] \dfrac{1}{N+1}\left| \mathbf{a}^H(\omega\dfrac{d}{c}\sin\theta)\mathbf{F}_{N+1}(\omega) \right|^2 \end{array} \right]$$

(4.62b)

Using the matrix inversion formula given in (4.34) we can get a recursive relation between the inverse of spectral matrix $\hat{\mathbf{S}}_f^N(\omega)$ and $\hat{\mathbf{S}}_f^{N+1}(\omega)$

$$\left[\hat{\mathbf{S}}_f^{N+1}(\omega) \right]^{-1} =$$

$$\frac{N+1}{N} \left\{ \left[\hat{\mathbf{S}}_f^N(\omega) \right]^{-1} - \frac{\dfrac{1}{N} \left[\hat{\mathbf{S}}_f^N(\omega) \right]^{-1} \mathbf{F}_{N+1}(\omega) \mathbf{F}_{N+1}^H(\omega) \left[\hat{\mathbf{S}}_f^N(\omega) \right]^{-H}}{1 + \dfrac{1}{N} \mathbf{F}_{N+1}^H(\omega) \left[\hat{\mathbf{S}}_f^N(\omega) \right]^{-1} \mathbf{F}_{N+1}(\omega)} \right\} \quad (4.62c)$$

The above recursion can be commenced only after $\hat{\mathbf{S}}_f^N(\omega)$ becomes full rank. This will require a minimum of M frequency snapshots. There is yet another recursive approach to spectral matrix inversion. It is based on diagonalization of the spectral matrix through a transformation,

$$\Gamma = \mathbf{Q}\hat{\mathbf{S}}_f^N(\omega)\mathbf{Q}^H \quad (4.63a)$$

or equivalently,

$$\hat{\mathbf{S}}_f^N(\omega) = \mathbf{Q}\Gamma\mathbf{Q}^H \quad (4.63b)$$

where $\Gamma = \text{diag}\{\gamma_1, \gamma_1, \ldots \gamma_M\}$ and $\mathbf{Q}$ is the upper triangular matrix of prediction coefficients which are computed through a recursive algorithm [24]. From (4.63b) we can express the inverse of the spectral matrix as

$$\left[\hat{\mathbf{S}}_f^N(\omega) \right]^{-1} = \mathbf{Q}\Gamma^{-1}\mathbf{Q}^H \quad (4.64)$$

§4.5 **Exercises:**

1. The spatial undersampling has resulted into aliasing of the frequency wavenumber spectrum as illustrated in figs. 4.2 and 4.3. Now consider temporal undersampling. Sketch the (ω, k) spectrum of a stochastic plane wave which has been undersampled temporally.

2. Apply the principle of Rayleigh resolution to wideband signals. In chapter 2 we have shown how the DOA of an incident wideband signal can be estimated from the position of the spectral peaks of the transfer functions. Show that for resolution the wavefronts must be separated by an angle greater than $\Delta\theta$, where

$$\Delta\theta = \frac{\tan\theta}{M}$$

3. Show that the power output of a beamformer given by (4.16) is always greater than the Capon spectrum. Use Schwarze inequality (see, p.19, [23]).

4. The noise variance is often required to be known but in practice this is not likely to be true. It has to be estimated from the observed data. One possible approach is to use the partitioning of the spectral matrix as on page 215. Show that

$$\sigma_\eta^2 = \frac{tr(\mathbf{H}_2\pi)}{tr(\pi)}$$

where $\pi = \mathbf{I}_{M-P} - \mathbf{G}_2\mathbf{G}_2^\#$ and $\mathbf{G}_2^\#$ is a pseudoinverse of $\mathbf{G}_2$.

5. A plane wavefront is incident on a ULA from a known direction θ_0 in presence of coloured background noise. Consider the following noise reduction strategy. The array is steered to the known direction along with a weight vector. The steering vector is $\mathbf{w}_1 = \Gamma\mathbf{w}_0$ where $\mathbf{w}_0$ is an unknown weight vector and

$$\Gamma = diag\left\{1, e^{-j\omega\frac{d}{c}\sin\theta_0}, \ldots, e^{-j(M-1)\omega\frac{d}{c}\sin\theta_0}\right\}$$

It is proposed to find $\mathbf{w}_0$ which shall minimize the noise variance and at the same time preserve the signal power. Show that $\mathbf{w}_0$ is equal to the eigenvector corresponding to the least eigenvalue of $\Gamma^H\mathbf{S}_\eta\Gamma$ where $\mathbf{S}_\eta$ is the spectral matrix of the noise.

6. It was shown on page 222 that the weight vector which maximizes the snr is the eigenvector corresponding to the largest eigenvalue of a pencil matrix $[\mathbf{c}_s, \mathbf{c}_\eta]$. Let a single wavefront be incident on the array. The covariance matrix, when there is no noise, is $\mathbf{c}_s = s_0\mathbf{a}_0\mathbf{a}_0^H$. Show that $\mathbf{w} = \alpha\mathbf{c}_\eta^{-1}\mathbf{a}_0$ where α is a normalizing constant and the maximum snr is equal to $s_0\mathbf{a}_0^H\mathbf{c}_\eta^{-1}\mathbf{a}_0$.

7. Let $\mathbf{B}$ be a MxQ (Q<M) matrix satisfying a property $\mathbf{B}^H\mathbf{B} = \mathbf{I}$. We define a reduced array output $\mathbf{G}_i(\omega) = \mathbf{B}^H\mathbf{F}_i(\omega)$ where $\mathbf{F}_i(\omega)$ is a frequency snapshot (see 4.62a). Show that the eigenstructure of the spectral matrix of the reduced output is identical to that of the spectral matrix of the normal array output, in particular, $\mathbf{B}^H\mathbf{a} \perp \mathbf{v}_\eta$. This property forms the basis for the beamspace subspace method where $\mathbf{B}$ acts as a spatial filter to restrict the incident energy to a preselected angular sector [26, 27].

8. Let the columns of $\mathbf{B}$ be the eigenvectors corresponding to the significant eigenvalues of $\mathbf{Q}$ in (1.70). Assume that noise sources are distributed over an

angular sector $\pm\Delta$ ($\theta_0 = 0$). Show that the reduced noise (as defined in exercise 7 above) becomes white.

References

1. S. Haykin (Ed), *Nonlinear methods of spectrum analysis*, Springer-Verlag, Berlin, 1979.

2. P. S. Naidu, *Modern Spectrum Analysis of Time Series*, CRC Press, Boca Raton, FL. 1996.

3. G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins Univ. Press, Baltimore, 1983.

4. S. Marcos, A. Marsal, and M. Benidie, The propagation method for source bearing estimation, Signal Proc., vol.42, pp. 121-138, 1995.

5. J. Munier and G. Y. Deliste, Spatial analysis using new properties of the cross spectral matrix, IEEE, Trans. Signal Proc., vol.37, pp. 1991.

6. H. P. Bucker, Use of calculated sound fields and matched field detection to locate sound sources in shallow water, J. Acoust. Soc. of Am., vol. 59, pp. 368-373, 1976.

7. A. Tolstoy, *Matched field processing*, WorldScientific, Singapore, 1995.

8. I. D. Rathjen, G. Boedecker, and M. Siegel, Omnidirectional beam forming for linear antennas by means of interpolated signals, IEEE Jour of Ocean Eng. vol. OE-10, pp. 360-368, 1985.

9. F. J. Harris, On the use of windows for harmonic analysis with the discrete Fourier transform, Proc IEEE, vol. 66, pp. 51-83, 1978.

10. R. Monzingo and T. Miller, *Introduction to Adaptive Arrays*, Wiley and Sons, New York, 1980.

11. A. Eberhard, An optimal discrete window for calculation of power spectra, IEEE Trans., AU-21, pp. 37-43, 1973.

12. K. Yao, Maximum energy windows with constrained spectral values, Signal Processing, pp. 157-168, 1986.

13. D. Korompis, K. Yao, and F. Lorenzelli, Broadband maximum energy array with user imposed spatial and frequency constraints, IEEE, ICASSP, pp. IV-529-532, 1994.

14. J. Capon, High resolution frequency-wavenumber spectrum analysis, Proc. IEEE, 57, pp. 1408-1418, 1969.

15. B. D. Van Veen and K. M. Buckley, Beamforming: A versatile approach to spatial filtering, IEEE ASSP Mag., pp.4-24, April 1988.

16. H. Cox, Resolving power and sensitivity to mismatch of optimum array processors, J. Acoust. Soc. of Am., vol. 54, pp. 771-785, 1973.

17. M. H. Er and A Cantoni, A unified approach to the design of robust narrowband antenna processor, IEEE Trans. on Antenna and Propagation, vol. 38, pp. 17-23, 1990.

18. D. G. Luenberger, Optimization by vector space methods, Wiley, New York, 1969.

19. J. Munier and G. R. Deliste, Spatial analysis in passive listening using adaptive techniques, Proc. IEEE, vol. 75, pp. 1458-1471, 1987.

20. P. S. Naidu and V. V. Krishna, Improved maximum likelihood (IML) spectrum estimation: Performance analysis, Jour. I. E. T. E, vol. 34, pp. 383-390, 1988.

21. J. P. Burg, The relationship between maximum entropy spectra and maximum likelihood spectra, Geophysics, vol. 37, pp. 375-376, 1972.

22. S. R. De Graaf, and D. H. Johnson, Capability of array processing algorithms to estimate source bearings, IEEE Trans ASSP-33, pp. 1368-1379, 1985.

23. S. U. Pillai, Array Signal Processing, Springer-Verlag, New York, 1989.

24. T. S. Durrani and K. C. Sharman, Eigenfilter approaches to adaptive array processing, IEE Proc., vol. 130, Pts F & H, pp. 22-28, 1983.

25. S. A. Nakhamkin, L. G. Tyurikov, and A. V. Malik, Construction models and decomposition algorithm of seismic fields from the characteristic numbers and vectors of spectral matrices, Izvestia, Earth Physics, No. 12, pp. 785-791, 1975.

26. G. Bienvenu and L Kopp, Decreasing high resolution method sensitivity by conventional beamformer preprocessing, IEEE, ICASSP-84, pp.33.2.1-33.2.4, 1984.

27. P. Forster and G. Vezzosi, Application of spheroidal sequences to array processing, IEEE, ICASSP-87, pp. 2268-2271, 1987.

Chapter Five

Source Localization: Subspace Methods

The location parameters are estimated directly without having to search for peaks as in frequency-wavenumber spectrum (the approach described in chapter 4). In open space the direction of arrival (DOA), that is, azimuth or elevation or both, is estimated using the subspace properties of the spatial covariance matrix or spectral matrix. MUSIC is a well known algorithm where we define a positive quantity which becomes infinity whenever the assumed parameter(s) is equal to the true parameter. We shall call this quantity as a spectrum even though it does not possess the units of power as in the true spectrum. The MUSIC algorithm in its original form does involve scanning and searching, often very fine scanning lest we may miss the peak. Later extensions of the MUSIC, like root MUSIC, ESPRIT, etc. have overcome this limitation of the original MUSIC algorithm. When a source is located in a bounded space, such as a duct, the wavefront reaching an array of sensors is necessarily nonplanar due to multipath propagation in the bounded space. In this case all three position parameters can be measured by means of an array of sensors. But the complexity of the problem of localization is such that a good prior knowledge of the channel becomes mandatory for successful localization. In active systems, since one has control over the source, it is possible to design waveforms which possess the property that is best suited for localization; for example, a binary phase shift key (BPSK) signal with its narrow autocorrelation function is best suited for time delay estimation. Source tracking of a moving source is another important extension of the source localization problem.

§5.1 **Subspace Methods** (Narrowband):
Earlier we showed an interesting property of eigenvalues and eigenvectors of a spectral matrix of a wavefield which consists of uncorrelated plane waves in the presence of white noise. Specifically, equations (4.14a) and (4.14b) form the basis for the signal subspace method for DOA estimation. For convenience, those two equations are reproduced here

$$\lambda_m = \alpha_m + \sigma_\eta^2 \qquad\qquad\qquad (4.14a)$$

$$\mathbf{v}_m^H \mathbf{A} = \mathbf{0}, \qquad \text{for} \qquad m = P, P+1,\ldots,M-1 \qquad (4.14b)$$

Equation (4.14b) implies that the space spanned by the columns of $\mathbf{A}$, that is, the direction vectors of incident wavefronts, is orthogonal to the space spanned by the eigenvectors, $\mathbf{v}_m$, $m = P, P+1,\ldots,M-1$, often known as noise

subspace, $\mathbf{v}_\eta$. The space spanned by the columns of $\mathbf{A}$ is known as signal subspace, $\mathbf{v}_s$. Consider the space spanned by a steering vector,

$$\mathbf{a}(\omega,\theta) = col\left\{1, e^{-j\omega\frac{d}{c}\sin\theta}, e^{-j2\omega\frac{d}{c}\sin\theta}, \ldots, e^{-j(M-1)\omega\frac{d}{c}\sin\theta}\right\}, \text{ as the steering}$$

angle is varied over a range $\pm\dfrac{\pi}{2}$. The intersection of the array manifold with the signal subspace yields the direction vectors of the signals.

5.1.1 <u>MUSIC</u>: On account of (4.14b) a steering vector pointing in the direction of one of the incident wavefronts will be orthogonal to the noise subspace,

$$\mathbf{v}_m^H \mathbf{a}(\omega,\theta) = 0, \quad m = P, P+1, \ldots, M-1 \qquad (5.1)$$

where $\theta = \theta_0, \theta_1, \ldots, \theta_{P-1}$. For narrowband signals the matrix $\mathbf{A}$ is given by

$$\mathbf{A} = \left[\mathbf{a}(\omega_0 \frac{d}{c}\sin\theta_0), \mathbf{a}(\omega_1 \frac{d}{c}\sin\theta_1), \ldots, \mathbf{a}(\omega_{P-1}\frac{d}{c}\sin\theta_{P-1})\right] \quad (5.2)$$

where we take a wavefront with a center frequency ω_p to be incident at an angle θ_p. Further, we assume that the center frequencies of the narrowband signals are known. We define a quadratic function involving a steering vector and noise subspace,

$$S_{Music}(\omega,\theta) = \frac{1}{\mathbf{a}^H(\omega,\theta)\mathbf{v}_\eta \mathbf{v}_\eta^H \mathbf{a}(\omega,\theta)} \qquad (5.3)$$

$S_{Music}(\omega,\theta)$, also known as the eigenvector spectrum, will show sharp peaks whenever $\theta = \theta_0, \theta_1, \ldots, \theta_{P-1}$. The subscript *Music* stands for <u>M</u>ultiple <u>S</u>ignal <u>C</u>lassification. This acronym was coined by Schmidt [1] who discovered the subspace algorithm. At about the same time but independently Bienvenu and Kopp [2] proposed a similar algorithm. Pisarenko [3] had previously published a subspace based algorithm in the context of harmonic analysis of time series. Note that, although we refer to $S_{Music}(\omega,\theta)$ as spectrum, it does not have the units of power; hence it is not a true spectrum. Let us express the steering vector in terms of the product of frequency and time delay $\tau = \dfrac{d}{c}\sin\theta$,

$$\mathbf{a}(\omega\tau) = col\left\{1, e^{j\omega\tau}, e^{j2\omega\tau}, \ldots, e^{j(M-1)\omega\tau}\right\}$$

The peaks of $S_{Music}(\omega,\theta)$ will now be at $\omega\tau = \omega_0\tau_0, \omega_1\tau_1,...,\omega_{P-1}\tau_{P-1}$ and given the frequencies we can estimate the delays but there is one difficulty. It is possible that two wavefronts with different center frequencies may arrive at such angles that $\omega_p \dfrac{d}{c}\sin\theta_p = \omega_{p'} \dfrac{d}{c}\sin\theta_{p'}$, in which case the two direction vectors will become identical, causing a loss of the full column rank property of the matrix $\mathbf{A}$. When the frequencies are unknown, evidently, the delays or DOA's cannot be uniquely estimated.

We now turn to the signal subspace spanned by the eigenvectors corresponding to eigenvalues $\lambda_m = \alpha_m + \sigma_\eta^2$, $m = 0,1,...,P-1$. We shall show that the signal subspace is the same as the space spanned by the columns of the matrix $\mathbf{A}$. Since $\mathbf{A}$ is a full rank matrix its polar decomposition gives

$$\mathbf{A} = \mathbf{TG} \tag{5.4}$$

where $\mathbf{G}$ is a full rank PxP matrix and $\mathbf{T}$ is a MxP matrix satisfying the following property,

$$\mathbf{T}^H\mathbf{T} = \mathbf{I} \qquad \text{(PxP unit matrix)}$$

In (4.14a) we shall replace $\mathbf{A}$ by its polar decomposition (5.4)

$$\mathbf{v}_m^H\mathbf{A}\mathbf{C}_0\mathbf{A}^H\mathbf{v}_m = \alpha_m \tag{5.5a}$$

or in matrix form

$$\mathbf{v}_s^H\mathbf{T}[\mathbf{G}\mathbf{C}_0\mathbf{G}^H]\mathbf{T}^H\mathbf{v}_s = diag\{\alpha_m, m = 0,1,...,P-1\} \tag{5.5b}$$

Let $\mathbf{T} = \mathbf{v}_s$, which is consistent with the assumed properties of $\mathbf{T}$. Equation (5.5b) now reduces to

$$\mathbf{G}\mathbf{C}_0\mathbf{G}^H = diag\{\alpha_m, m = 0,1,...,P-1\} \tag{5.6a}$$

and

$$\mathbf{A} = \mathbf{v}_s\mathbf{G} \tag{5.6b}$$

We can estimate $\mathbf{G}$ from (5.6b) by premultiplying both sides by $\mathbf{v}_s^H$ and obtain

$$\mathbf{G} = \mathbf{v}_s^H \mathbf{A} \tag{5.6c}$$

Thus, the space spanned by $\mathbf{A}$ and $\mathbf{v}_s^H$ are identical. Further, by eliminating $\mathbf{G}$ from (5.6b) and (5.6c) we obtain an interesting result

$$\mathbf{A} = \mathbf{v}_s \mathbf{v}_s^H \mathbf{A} \tag{5.6d}$$

Let us define a complementary orthogonal space $(\mathbf{I} - \mathbf{v}_s \mathbf{v}_s^H)$ which will also be orthogonal to $\mathbf{A}$; therefore, $\mathbf{A}^H (\mathbf{I} - \mathbf{v}_s \mathbf{v}_s^H)\mathbf{A} = 0$. An equivalent definition of Music spectrum may be given using $(\mathbf{I} - \mathbf{v}_s \mathbf{v}_s^H)$,

$$S_{Music}(\omega,\theta) = \frac{1}{\mathbf{a}^H(\omega,\theta)(\mathbf{I} - \mathbf{v}_s \mathbf{v}_s^H)\mathbf{a}(\omega,\theta)} \tag{5.7}$$

<u>Signal Eigenvalues and Source Power</u>: The signal eigenvalues and source power are related, although the relationship is rather involved except for the single source case. Let us first consider a single source case. The covariance matrix is given by $\mathbf{C}_0 = \mathbf{a}_0 s_0 \mathbf{a}_0^H$ which we use in (5.5a) and obtain

$$s_0 = \frac{\alpha_0}{\left| \mathbf{v}_0^H \mathbf{a}_0 \right|^2} \tag{5.8}$$

Since $\mathbf{v}_0$ and $\mathbf{a}_0$ span the same space, $\mathbf{v}_0 = g_0 \mathbf{a}_0$ where g_0 is a constant which may be obtained by requiring that $\mathbf{v}_0^H \mathbf{v}_0 = 1$. We obtain $g_0 = \dfrac{1}{\sqrt{M}}$. Equation (5.8) reduces to $s_0 = \dfrac{\alpha_0}{M}$. Next we consider the two-source case. The covariance function is given by

$$\mathbf{C}_f = \begin{bmatrix} \mathbf{a}_0, \mathbf{a}_1 \end{bmatrix} \begin{bmatrix} s_0 & 0 \\ 0 & s_1 \end{bmatrix} \begin{bmatrix} \mathbf{a}_0, \mathbf{a}_1 \end{bmatrix}^H$$

and using (5.5a) we obtain

$$\mathbf{v}_s^H \mathbf{A} \mathbf{C}_0 \mathbf{A}^H \mathbf{v}_s = diag\{\alpha_0, \alpha_1\} \tag{5.9a}$$

azimuths	power #1 source	power #2 source
30º, 40º	1.00	1.00
30º, 35º	1.00	1.00
30º, 33º	1.00	1.00
30º, 31º	1.00	1.00
30º, 30.5º	1.00	1.00

Table 5.1 Estimation of power given the azimuths of two equal power wavefronts incident on a 16 sensor ULA. Power estimates are error-free even when the wavefronts are closely spaced.

$$\mathbf{C}_0 = \left[\mathbf{v}_s^H \mathbf{A}\right]^{-1} diag\{\alpha_0, \alpha_1\} \left[\mathbf{A}^H \mathbf{v}_s\right]^{-1} \qquad (5.9b)$$

Equation (5.9b) has been used to estimate the powers of two equal power (snr=1.0) wavefronts incident on a 16 sensor ULA and the results are tabulated in Table 5.1. The estimation is error-free right down to a half-a-degree separation; however, this good performance deteriorates in the presence of model and estimation errors.

Aliasing: We pointed out earlier (p.211) that the sensor spacing in a ULA must be less than half the wavelength so that there is no aliasing in the frequency-wavenumber spectrum. The same requirement exists in all high resolution methods, namely the Capon spectrum, Maximum entropy spectrum and Music spectrum. The basic reason for aliasing lies in the fact that the direction vector is periodic. Consider anyone column of the matrix $\mathbf{A}$,

$$\mathbf{a}_p(\omega\tau) = col\left\{1, e^{-j\omega_p \frac{d}{c}\sin\theta_p}, e^{-j2\omega_p \frac{d}{c}\sin\theta_p}, \ldots, e^{-j(M-1)\omega_p \frac{d}{c}\sin\theta_p}\right\}$$

Now let $d = \dfrac{\lambda_p}{\sin\theta_p} - \delta_p$ where $0 \leq \delta_p \leq \dfrac{\lambda_p}{2\sin\theta_p}$, then $\mathbf{a}_p(\omega\tau)$

becomes

$$\mathbf{a}_p(\omega_p \frac{d}{c}\sin\theta_p) = col\left\{1, e^{-j2\pi\frac{d}{\lambda_p}\sin\theta_p}, e^{-j4\pi\frac{d}{\lambda_p}\sin\theta_p}, \ldots, e^{-j(M-1)2\pi\frac{d}{\lambda_p}\sin\theta_p}\right\}$$

$$= col\left\{1, e^{j2\pi\frac{\delta_p}{\lambda_p}\sin\theta_p}, e^{j4\pi\frac{\delta_p}{\lambda_p}\sin\theta_p}, \ldots, e^{j(M-1)2\pi\frac{\delta_p}{\lambda_p}\sin\theta_p}\right\}$$

$$= \mathbf{a}_p(\omega_p\frac{-\delta_p}{c}\sin\theta_p)$$

We can find an angle $\hat{\theta}_p$ such that $\frac{d}{\lambda_p}\sin\hat{\theta}_p = -\frac{\delta_p}{\lambda_p}\sin\theta_p$ and hence

$$\mathbf{a}_p(\omega_p\frac{d}{c}\sin\theta_p) = \mathbf{a}_p(\omega_p\frac{d}{c}\sin\hat{\theta}_p)$$

An aliased spectral peak would appear at $\hat{\theta}_p$ which is related to θ_p,

$$\hat{\theta}_p = \sin^{-1}\left(\frac{\delta_p}{d}\sin\theta_p\right) \tag{5.10}$$

As an example, consider an array with sensor spacing, $\frac{d}{\lambda_p} = 1$ and a wavefront incident at an angle $\theta_p = 60°$. For this choice of array and wave parameters $\delta_p = 0.1547$ and from (5.10) we get the angle where the aliased peak will be located, $\hat{\theta}_p = -7.6993°$. The wave number spectrum computed by all four methods is shown in fig. 5.1.

Aliasing is on account of periodicity of a direction vector which in turn is caused by periodicity present in a ULA. Thus, to avoid aliasing, it would be necessary to break this periodicity; for example, we may space the sensors nonuniformly. In a circular array, though sensors are uniformly spaced (e.g. UCA), the time delays are nonuniform; therefore a UCA will yield an alias-free spectrum [4]. This is demonstrated in fig. 5.2 where we consider a wavefront which is incident at $60°$ (with respect to x-axis) on a circular array consisting of 16 sensors uniformly spread over a circle of radius 8λ. The Capon spectrum is shown for this case. The aliasing phenomenon is not encountered in random arrays where the sensors are spaced at random intervals. But as shown in chapter 2 the random array possesses a highly nonlinear phase response.
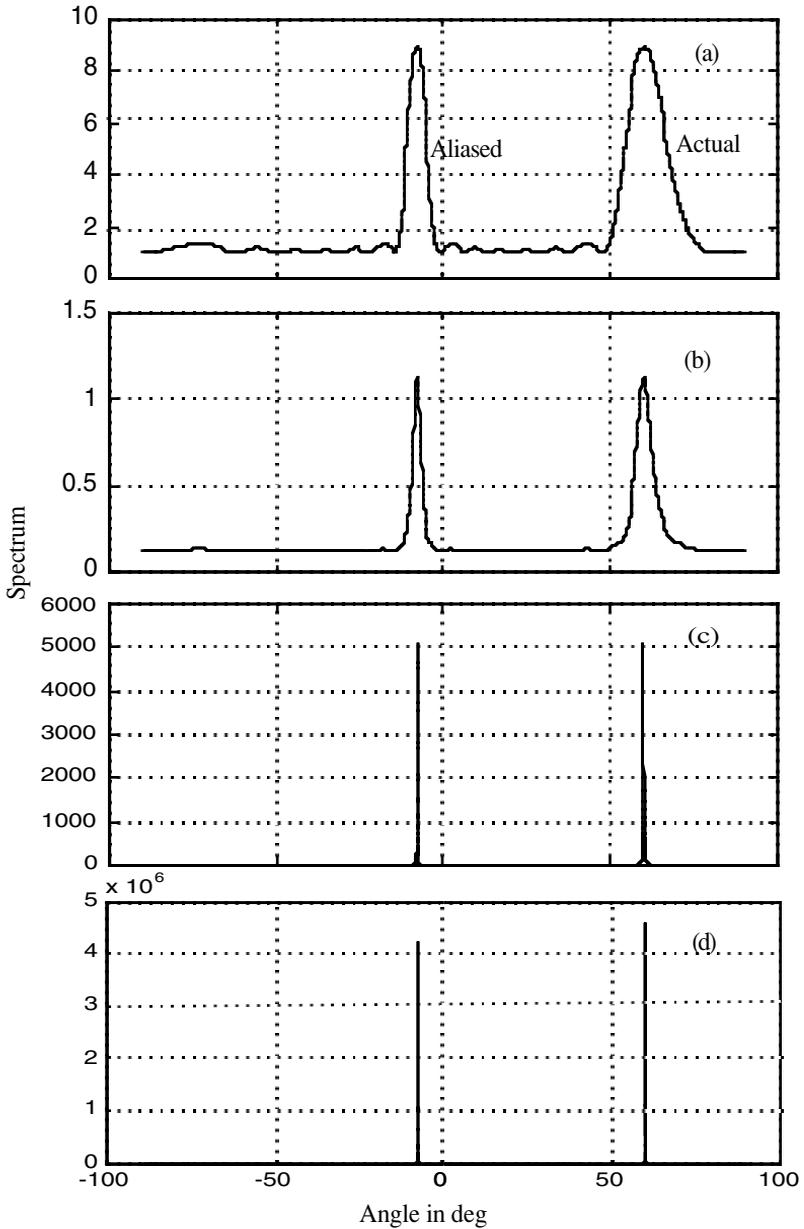
Figure 5.1 The aliasing effect due to undersampling of the wavefield (d=λ). All four methods of spectrum estimation have been used. (a) Bartlett spectrum, (b) Capon spectrum, (c) Maximum entropy spectrum and (d) Music spectrum. While the actual peak is at 60° the aliased peak appears at -7.69^o.

5.1.2 <u>Correlated Sources</u>: We have so far assumed that the source matrix $\mathbf{S}_0$ is a full rank matrix. When sources are fully uncorrelated $\mathbf{S}_0$ is a diagonal matrix with the nonzero elements representing the power of the sources. The source matrix is naturally full rank. Let us now consider a situation where the sources are partially or fully correlated. We model the source matrix as

$$\mathbf{S}_0 = \mathbf{s}\,\rho\,\mathbf{s}^H \tag{5.11}$$

where

$$\mathbf{s} = diag\left\{\sqrt{s_0},\sqrt{s_1},\ldots,\sqrt{s_{P-1}}\right\}$$

$s_0, s_1, \ldots, s_{P-1}$ represent the power of P sources and ρ is the coherence matrix whose $(m,n)^{th}$ element represents the normalized coherence between the mth and n^{th} sources. The signal eigenvalues of the spectral matrix for P=2 are given by [5]

$$\lambda_0 = \frac{M}{2}(s_0 + s_1) + M\sqrt{s_0 s_1}\,\operatorname{Re}\{\rho_{12}\}\psi +$$

$$\frac{M}{2}\left\{\left[(s_0 + s_1) + 2\sqrt{s_0 s_1}\,\operatorname{Re}\{\rho_{12}\}\psi\right]^2 - 4s_0 s_1(1 - |\psi|^2)(1 - |\rho_{12}|^2)\right\}^{\frac{1}{2}} + \sigma_\eta^2$$

$$\lambda_1 = \frac{M}{2}(s_0 + s_1) + M\sqrt{s_0 s_1}\,\operatorname{Re}\{\rho_{12}\}\psi -$$

$$\frac{M}{2}\left\{\left[(s_0 + s_1) + 2\sqrt{s_0 s_1}\,\operatorname{Re}\{\rho_{12}\}\psi\right]^2 - 4s_0 s_1(1 - |\psi|^2)(1 - |\rho_{12}|^2)\right\}^{\frac{1}{2}} + \sigma_\eta^2$$

$$\tag{5.12a}$$

where

$$\psi(M) = \frac{\sin(\pi\frac{d}{\lambda}M(\sin\theta_0 - \sin\theta_1))}{M\sin(\pi\frac{d}{\lambda}(\sin\theta_0 - \sin\theta_1))}$$

The sum of the signal eigenvalues, that is,

$$(\lambda_0 - \sigma_\eta^2) + (\lambda_1 - \sigma_\eta^2) =$$

$$M(s_0 + s_1) + 2M\sqrt{s_0 s_1}\,\operatorname{Re}\{\rho_{12}\}\psi(M) \tag{5.12b}$$
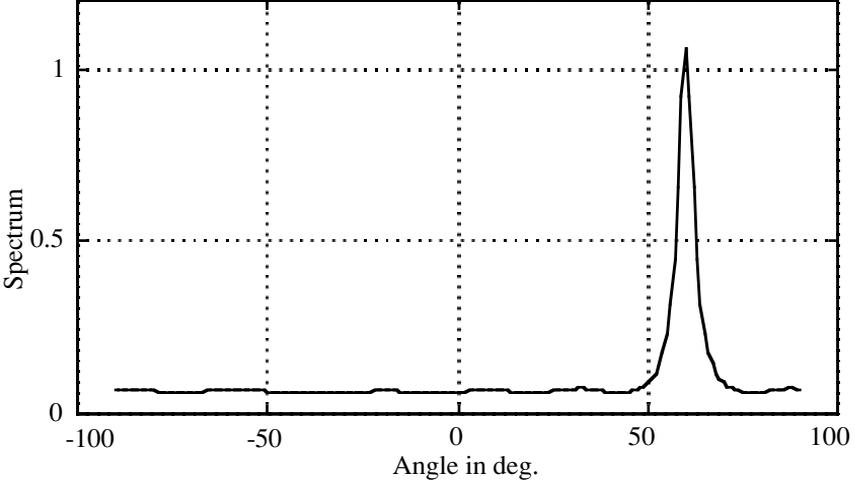
Figure 5.2: No aliasing effect is seen with a circular array. Sixteen sensor UCA with radius=8λ (sensor spacing=3.12l) is used. A plane wavefront is incident on the array at 60°.

represents coherent addition of power. For uncorrelated sources $\rho_{12} = 0$ (5.12a) reduces to

$$\lambda_0 = \frac{M}{2}(s_0 + s_1) + \frac{M}{2}\left\{(s_0 + s_1)^2 - 4s_0 s_1(1 - |\psi(M)|^2)\right\}^{\frac{1}{2}} + \sigma_\eta^2$$

$$\lambda_1 = \frac{M}{2}(s_0 + s_1) - \frac{M}{2}\left\{(s_0 + s_1)^2 - 4s_0 s_1(1 - |\psi(M)|^2)\right\}^{\frac{1}{2}} + \sigma_\eta^2$$

(5.13)

Also, note that when the sources are in the same direction $\lambda_0 = M(s_0 + s_1)$ and $\lambda_1 = 0$.

The source spectral matrix may be modified by spatial smoothing of the array outputs. This is achieved by averaging the spectral matrices of subarray outputs over all possible subarrays (see fig. 5.3). The ith subarray (size μ) signal vector at a fixed temporal frequency is given by

$$\mathbf{F}_i = col\left\{F_i(\omega), F_{i+1}(\omega), \ldots, F_{i+\mu-1}(\omega)\right\} \qquad 0 \le i \le M - \mu + 1$$

$$= \mathbf{I}_{i,\mu}\mathbf{F}$$
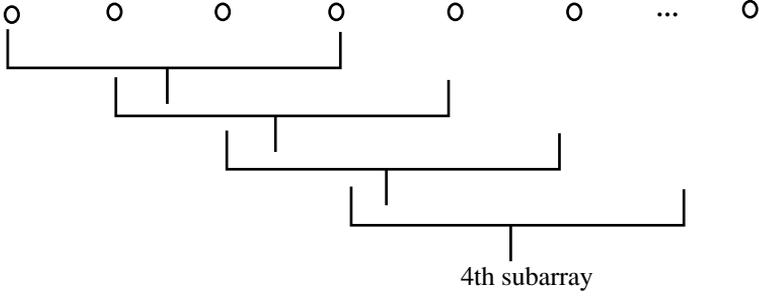
Figure 5.3: Overlapping subarrays are formed as shown above. Each subarray has four sensors. It shares three sensors with its immediate neighbours.

where $\mathbf{I}_{i,\mu}$ is a diagonal matrix,

$$\mathbf{I}_{i,\mu} = diag\left\{\underset{\substack{0 \text{ to } i-1, \quad i \text{ to } i+\mu-1, \quad i+\mu \text{ to } M}}{0,...0, \quad 1,...,1, \quad 0,...,0}\right\}$$

$$\mathbf{F} = col\{F_0(\omega),...,F_{M-1}(\omega)\}$$

The spectral matrix of the ith subarray is now given by $\mathbf{S}_{i,\mu} = E\{\mathbf{F}_i\mathbf{F}_i^H\} = \mathbf{I}_{i,\mu}\mathbf{S}\ \mathbf{I}_{i,\mu}^H$. We shall now average all subarray spectral matrices

$$
\begin{aligned}
\overline{\mathbf{S}} &= \frac{1}{M-\mu+1}\sum_{i=0}^{M-\mu+1}\mathbf{S}_{i,\mu} = \frac{1}{M-\mu+1}\sum_{i=0}^{M-\mu+1}\mathbf{I}_{i,\mu}\mathbf{S}\ \mathbf{I}_{i,\mu}^H \\
&= \frac{1}{M-\mu+1}\sum_{i=0}^{M-\mu+1}\mathbf{I}_{i,\mu}\mathbf{A}(\omega)\mathbf{S}_0(\omega)\mathbf{A}^H\ \mathbf{I}_{i,\mu}^H + \sigma_\eta^2\mathbf{I}
\end{aligned}
\tag{5.14a}
$$

where we have used the spectral matrix of the signal model plane waves in the presence of white noise (4.12b). We can show that

$$
\begin{aligned}
\mathbf{I}_{i,\mu}\mathbf{A}(\omega) &= \mathbf{I}_{i,\mu}\big[\mathbf{a}_0,\mathbf{a}_1,...,\mathbf{a}_{P-1}\big] \\
&= \big[\hat{\mathbf{a}}_0,\hat{\mathbf{a}}_1,...,\hat{\mathbf{a}}_{P-1}\big]diag\left\{e^{-j2\pi i\frac{d}{c}\sin\theta_0},...,e^{-j2\pi i\frac{d}{c}\sin\theta_{P-1}}\right\}
\end{aligned}
\tag{5.14b}
$$

where

$$\hat{\mathbf{a}}_i(\omega\tau) = col\left\{1, e^{-j\omega\frac{d}{c}\sin\theta_i}, e^{-j2\omega\frac{d}{c}\sin\theta_i}, \dots, e^{-j(\mu-1)\omega\frac{d}{c}\sin\theta_i}\right\}$$

We shall use (5.14b) in (5.14a) and obtain

$$\overline{\mathbf{S}} = \hat{\mathbf{A}}\left[\frac{1}{M-\mu+1}\sum_{i=0}^{M-\mu+1}\phi_i\mathbf{S}_0(\omega)\ \phi_i^H\right]\hat{\mathbf{A}}^H + \sigma_\eta^2\mathbf{I} \tag{5.15a}$$

where

$$\hat{\mathbf{A}} = \left[\hat{\mathbf{a}}_0, \hat{\mathbf{a}}_1, \dots, \hat{\mathbf{a}}_{P-1}\right]$$

$$\phi_i = diag\left\{e^{-j2\pi i\frac{d}{c}\sin\theta_0}, e^{-j2\pi i\frac{d}{c}\sin\theta_1}, \dots, e^{-j2\pi i\frac{d}{c}\sin\theta_{P-1}}\right\}$$

The quantity inside the square brackets in (5.15a) may be computed by actual multiplication followed by summation,

$$[\ ]_{mn} = \sqrt{s_m s_n}\rho_{mn}\frac{1}{M-\mu+1}\sum_{i=0}^{M-\mu+1}e^{-j2\pi\frac{d}{c}i(\sin\theta_m - \sin\theta_n)}$$

$$= \sqrt{s_m s_n}\rho_{mn}\psi(M-\mu+1)e^{-j\pi\frac{d}{c}(M-\mu)(\sin\theta_m - \sin\theta_n)} \tag{5.15b}$$

From (5.15b) it follows that the coherence after spatial smoothing may be written as

$$\overline{\rho}_{mn} = \rho_{mn}\psi(M-\mu+1)e^{-j\pi\frac{d}{c}(M-\mu)(\sin\theta_m - \sin\theta_n)} \tag{5.16}$$

The magnitude of the off-diagonal terms in the coherence matrix, ρ (see (5.11)), is reduced by a factor, $\psi(M-\mu+1)$, which is indeed small ($\ll 1$) for large $(M-\mu+1)$. This contributes to the increase of the rank of the smoothed coherence matrix. It is shown in [6] that ρ becomes full rank when $M-\mu+1 \geq P$.

Let us now examine the effect of the spatial smoothing on the eigenvalues for the two source case. From

$$\overline{\lambda}_0 = \frac{\mu}{2}(s_0 + s_1) + \mu\sqrt{s_0 s_1}\,\mathrm{Re}\{\tilde{\rho}_{01}\}\psi(\mu)\psi(M - \mu + 1) +$$

$$\frac{\mu}{2}\left\{\begin{array}{l}\left[(s_0 + s_1) + 2\sqrt{s_0 s_1}\,\mathrm{Re}\{\tilde{\rho}_{01}\}\psi(\mu)\psi(M - \mu + 1)\right]^2 \\ -4s_0 s_1(1 - |\psi(\mu)|^2)(1 - |\rho_{01}\psi(M - \mu + 1)|^2)\end{array}\right\}^{\frac{1}{2}} + \sigma_\eta^2$$

$$\overline{\lambda}_1 = \frac{\mu}{2}(s_0 + s_1) + \mu\sqrt{s_0 s_1}\,\mathrm{Re}\{\tilde{\rho}_{01}\}\psi(\mu)\psi(M - \mu + 1) -$$

$$\frac{\mu}{2}\left\{\begin{array}{l}\left[(s_0 + s_1) + 2\sqrt{s_0 s_1}\,\mathrm{Re}\{\tilde{\rho}_{01}\}\psi(\mu)\psi(M - \mu + 1)\right]^2 \\ -4s_0 s_1(1 - |\psi(\mu)|^2)(1 - |\rho_{01}\psi(M - \mu + 1)|^2)\end{array}\right\}^{\frac{1}{2}} + \sigma_\eta^2$$

$$(5.17)$$

where $\tilde{\rho}_{01} = \rho_{01}e^{j\pi\frac{d}{\lambda}(M - \mu)(\sin\theta_0 - \sin\theta_1)}$. The sum of the signal eigenvalues is given by

$$(\overline{\lambda}_0 - \sigma_\eta^2) + (\overline{\lambda}_1 - \sigma_\eta^2) =$$

$$M(s_0 + s_1) + 2M\sqrt{s_0 s_1}\,\mathrm{Re}\{\tilde{\rho}_{01}\}\psi(\mu)\psi(M - \mu + 1)$$

If we select μ such that

$$\pi\frac{d}{\lambda}(M - \mu + 1)|(\sin\theta_0 - \sin\theta_1)| \approx \pi$$

then $\psi(M - \mu + 1) \approx 0$, $\overline{\lambda}_0 = Ms_0$ and $\overline{\lambda}_1 = Ms_1$ and the desired subarray size would be

$$\mu \approx (M + 1) - \frac{2}{|(\sin\theta_0 - \sin\theta_1)|}$$

Let $|(\sin\theta_0 - \sin\theta_1)| \approx \frac{1}{\mu}$. Then the desired subarray size is

$$\mu \approx \frac{1}{3}(M + 1) \tag{5.18}$$
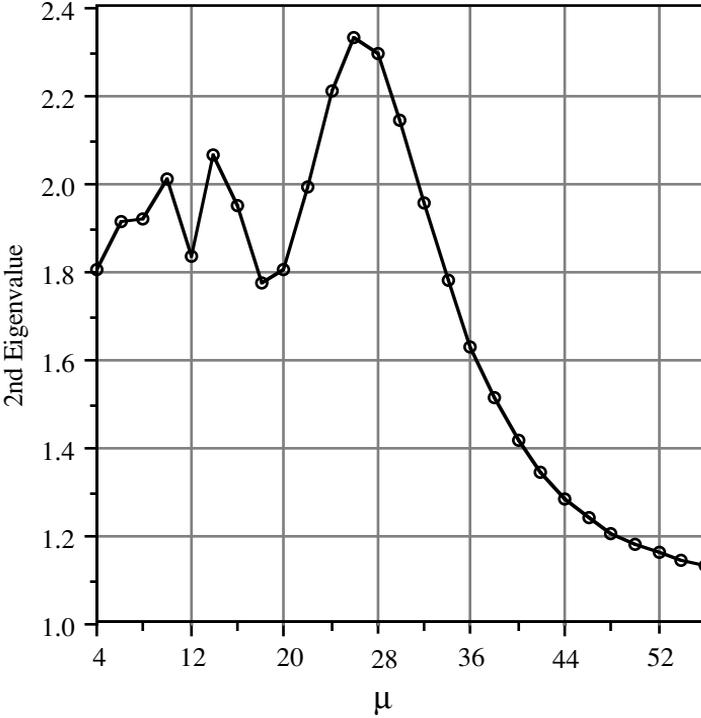
Figure 5.4: The 2nd eigenvalue relative to noise variance is shown as a function of the subarray size (μ).

This result is very close to that given in [7]. We have evaluated $\dfrac{\overline{\lambda_1}}{\sigma_\eta^2}$ as a function of the subarray size for two perfectly correlated equal power waves incident at angles θ_0 and θ_1 such that $\sin\theta_0 - \sin\theta_1 = \dfrac{1}{\mu}$ on a ULA with 64 sensors and array snr =10. The results are shown in fig. 5.4.

5.1.3 <u>Direct Estimation</u>: In MUSIC it is required to scan the entire angular range of $\pm\dfrac{\pi}{2}$. Since the spectral peaks are extremely fine the scanning must be done at a very fine interval; otherwise there is a high risk of missing the peak altogether. This is a serious drawback of the subspace methods based on scanning. Alternatively, there are direct methods which enable us to estimate the DOA directly. We shall describe three such methods, namely, Pisarenko's method, minimum norm, and root Music and show how they are interrelated.

Pisarenko's Method: Let the array size be $M = P + 1$ where P stands for the number of uncorrelated sources. The noise subspace is now spanned by a single eigenvector, $\mathbf{v}_{M-1}$. Hence, the direction vector, $\mathbf{a}$, will be orthogonal to $\mathbf{v}_{M-1}$,

$$\mathbf{v}_{M-1}^H \mathbf{a} = 0 \tag{5.19}$$

The direction vector of the pth source may be expressed as

$$\mathbf{a}_p = col\left\{ 1, e^{-j2\pi\frac{d}{\lambda}\sin\theta_p}, \ldots, e^{-j2\pi(M-1)\frac{d}{\lambda}\sin\theta_p} \right\}$$

$$= col\left\{ 1, z_p, z_p^2, \ldots, z_p^{M-1} \right\}, \quad p = 0, 1, \ldots, P-1$$

where $z_p = e^{-j2\pi\frac{d}{\lambda}\sin\theta_p}$. Consider a polynomial

$$\mathbf{v}_{M-1}^H \left[1, z, z^2, \ldots, z^{M-1} \right]^T = 0 \tag{5.20}$$

whose roots are indeed $z_0, z_1, \ldots, z_{P-1}$. In the complex plane all roots lie on a unit circle. The angular coordinate of a root is related to the DOA. For example, let the pth root be located at φ_p. Then $\varphi_p = \pi\sin\theta_p$. This method was first suggested by Pisarenko [3], many years before MUSIC was invented !

Root Music: The Pisarenko's method has been extended taking into account the noise space spanned by two or more eigenvectors [8]. The extended version is often known as root Music. Define a polynomial,

$$S(z) = \sum_{m=P}^{M-1} \mathbf{v}_m^H \left[1, z, z^2, \ldots, z^{M-1} \right]^T \tag{5.21}$$

The roots of $S(z)$ or $D(z) = S(z)S(\frac{1}{z})$ lying on the unit circle will correspond to the DOAs of the sources and the remaining M-P roots will fall inside the unit circle (and also at inverse complex conjugate positions outside the circle).

In the minimum norm method a vector $\mathbf{a}$ is found which is a solution of $\mathbf{a}^H \mathbf{S}_f \mathbf{a} = \min$ under the constraint that $\mathbf{a}^H \mathbf{a} = 1$. It turns out that the solution is an eigenvector of $\mathbf{S}_f$ corresponding to the smallest eigenvalue.

Radial coordinate	Angular coordinate (in radians)	Estimation error variance=.01	
		Radial	Angular
33.3162	0.9139	2.4854	-1.8515
2.2755	2.5400	1.5593	-0.7223
1.5793	2.9561	1.3047	-3.0944
0.8500	-1.7578	0.7244	-2.2490
0.8809	-0.9386	0.9336	1.6835
1.0000	0.5455	1.0050	0.5647
1.0000	0.7600	1.0057	0.7864

Table 5.2: Direct estimation of DOA by computing signal zeros of a polynomial given by (5.20) or (5.21). The results shown in columns one and two are for error free spectral matrix and those in columns three and four are for a spectral matrix with random errors (zero mean and 0.01 variance).

Thus, the minimum norm method belongs to the same class of direct methods of DOA estimation initiated by Pisarenko [3]. An example of DOA obtained by computing zeros on the unit circle under the ideal condition of no errors in the spectral matrix is shown in table 5.2. Two equal power signal wavefronts are assumed to be incident on eight sensor ULA at angles 10° and 14° . The DOA estimates from the error-free spectral matrix are exact (columns one and two in table 5.2), but in the presence of even a small error in the estimate of the spectral matrix, a significant error in DOA estimation may be encountered (see columns three and four in the table).

The zeros of the polynomial defined in any one of the above methods are important from the point of DOA estimation. In particular, the zeros which fall on the unit circle or close to it represent the DOAs of the incident wavefronts. These zeros are often called signal zeros and the remaining zeros, located deep inside the unit circle, are called the noise zeros. In MUSIC, a peak in the spectrum represents the DOA of the wavefront. But the height of the peak is greatly influenced by the position of the signal zero. The peak is infinite when the zero is right on the unit circle but it rapidly diminishes when the zero moves away from the unit circle. The peak may be completely lost, particularly in the presence of another signal zero in the neighborhood but closer to the unit circle. Hence, it is possible that, while the spectral peaks remain unresolved, the signal zeros are well separated. The shift of the signal zero may be caused by errors in the estimation of the spectral matrix from finite data. Let a signal zero at z_i be displaced to $\hat{z}_i$. The displacement, both radial and angular, is given by

$$\Delta z_i \approx (z_i - \hat{z}_i) \approx \delta r \, e^{j(u+\delta u)}$$

where we have assumed that $\delta u \ll 1$. Perturbation analysis reported in [9] for time series shows that the mean square error in δr and δu are related

$$\frac{E\{|\delta r_i|^2\}}{E\{|\delta \theta_i|^2\}} = 2N\left(\frac{\cos(\theta_i)}{\frac{\omega d}{c}}\right)^2 \leq \frac{2N}{\pi^2} \qquad (5.22)$$

where N stand for the number of time samples (or snapshots). From (5.22) it follows that $E\{|\delta r_i|^2\} \gg E\{|\delta \theta_i|^2\}$, particularly when the wavefronts are incident close to broadside. On account of the above finding the magnitude of a spectral peak in MUSIC is likely to be more adversely affected by the errors due to finite data.

Subspace Rotation: The direct estimation of DOA discussed in the previous section requires a ULA. This itself is a restriction in many practical situations. A new principle of DOA estimation, known as subspace rotation, exploits a property of a dipole sensor array, where all dipoles are held in the same direction, which is the subspace spanned by the direction vectors pertaining to the upper sensors in the dipole array and those pertaining to lower sensors are related through a rotation. The principle is explained in chapter 2 (see page 120). The basic starting equation is (2.57), which we reproduce here for convenience,

$$\mathbf{S}_{\tilde{f}} = \begin{bmatrix} \mathbf{A} \\ \mathbf{A}\Gamma \end{bmatrix} \mathbf{S}_f \begin{bmatrix} \mathbf{A} \\ \mathbf{A}\Gamma \end{bmatrix}^H + \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \sigma_{\eta_1}^2 \\ \sigma_{\eta_2}^2 \end{bmatrix} \qquad (2.57)$$

where $\mathbf{S}_{\tilde{f}}$ is the spectral matrix of the dipole array sensors,

$$\mathbf{S}_{\tilde{f}} = \begin{bmatrix} \mathbf{S}_{f_1 f_1} & \mathbf{S}_{f_1 f_2} \\ \mathbf{S}_{f_2 f_1} & \mathbf{S}_{f_2 f_2} \end{bmatrix} = \begin{bmatrix} \mathbf{A}\mathbf{S}_f \mathbf{A}^H & \mathbf{A}\mathbf{S}_f \Gamma^H \mathbf{A}^H \\ \mathbf{A}\Gamma\mathbf{S}_f \mathbf{A}^H & \mathbf{A}\mathbf{S}_f \mathbf{A}^H \end{bmatrix}$$

Note that $\mathbf{S}_{f_1 f_1}$ is the spectral matrix of the upper sensor outputs and $\mathbf{S}_{f_2 f_2}$ is the spectral matrix of the lower sensor outputs. $\mathbf{S}_{f_1 f_2}$ is the cross-spectral matrix between the upper sensor and lower sensor outputs. From the

eigenvalues of $\mathbf{S}_{f_1 f_1}$ we can estimate the noise variance $\sigma^2_{\eta_1}$ (and $\sigma^2_{\eta_2}$ from $\mathbf{S}_{f_2 f_2}$) and subtract it from the spectral matrix. Define $\mathbf{S}^0_{f_1 f_1} = \mathbf{S}_{f_1 f_1} - \sigma^2_{\eta_1}$. Consider the generalized eigenvector (GEV) problem of the matrix pencil $\left\{ \mathbf{S}^0_{f_1 f_1}, \mathbf{S}_{f_1 f_2} \right\}$,

$$\mathbf{S}^0_{f_1 f_1} \mathbf{v} = \gamma \mathbf{S}_{f_1 f_2} \mathbf{v} \tag{5.23}$$

where γ is a generalized eigenvalue and $\mathbf{v}$ is a corresponding eigenvector. Consider the following:

$$
\begin{aligned}
\mathbf{v}^H \left[\mathbf{S}^0_{f_1 f_1} - \gamma \mathbf{S}_{f_1 f_2} \right] \mathbf{v} &= \mathbf{v}^H \left[\mathbf{A}\mathbf{S}_f \mathbf{A}^H - \gamma \mathbf{A}\mathbf{S}_f \mathbf{\Gamma}^H \mathbf{A}^H \right] \mathbf{v} \\
&= \mathbf{v}^H \mathbf{A}\mathbf{S}_f \left[\mathbf{I} - \gamma \mathbf{\Gamma}^H \right] \mathbf{A}^H \mathbf{v} \\
&= \mathbf{v}^H \mathbf{A}\mathbf{S}_f \left[\mathbf{I} - \gamma \mathbf{\Gamma}^H \right] \mathbf{A}^H \mathbf{v} \\
&= \mathbf{v}^H \mathbf{Q} \mathbf{v} = 0
\end{aligned}
\tag{5.24}
$$

Since $\mathbf{S}_f$ is full rank as the sources are assumed to be uncorrelated and $\mathbf{A}$ is assumed to be full rank, then, for the right hand side of (5.24) to be true, we must have $\gamma = e^{-j\frac{\omega}{c}\Delta.\delta_i}$. The rank of $\mathbf{Q}$ will be reduced by one and a vector $\mathbf{v}$ can be found to satisfy (5.24). Thus the generalized eigenvalues of the matrix pencil $\left\{ \mathbf{S}^0_{f_1 f_1}, \mathbf{S}_{f_1 f_2} \right\}$ are $\gamma_i = e^{-j\frac{\omega}{c}\Delta.\delta_i}, i = 0,1,...,P-1$ from which we can estimate $\theta_i, i = 0,1,...,P-1$. This method of DOA estimation is known as ESPRIT (Estimation of signal parameters via rotation invariance technique) [10]. An example of DOA obtained via ESPRIT is in table 5.3. The eigenvalues of the pencil matrix (5.23) under the ideal condition of no errors in the spectral matrix as well as with errors in the spectral matrix are shown in table 5.3. Two equal power signal wavefronts are assumed to be incident on eight sensor ULA at angles 10º and 14º. The eigenvalues estimated from error free spectral matrix are exact (columns one and two in table 5.3), but in the presence of a small error in the estimation of the spectral matrix, a significant error in DOA estimation may be encountered (see columns three and four in the table).

The generalized eigenvector $\mathbf{v}_i$, corresponding to eigenvalue γ_i, possesses an interesting property. Let us write (5.24) in expanded form

Radial coordinate	Angular coordinate (in radians)	Estimation error variance=.01	
		Radial	Angular
4.6642	1.8540	3.4640	-2.1554
2.0198	0.7276	2.5202	0.4402
0.8182	-2.3272	1.7050	-0.4834
0.2108	-2.6779	0.5738	-1.6783
0.5224	-0.3753	0.4586	1.6290
0.5193	1.1199	0.6753	0.9060
1.0000	0.5455	0.9969	0.5533
1.0000	0.7600	1.0087	0.7720

Table 5.3: Generalized eigenvalues of pencil matrix (5.23).

$$\left[\mathbf{v}_i^H \mathbf{a}_0, \ldots, \mathbf{v}_i^H \mathbf{a}_{i-1}, \mathbf{v}_i^H \mathbf{a}_{i+1}, \ldots, \mathbf{v}_i^H \mathbf{a}_{P-1} \right] \hat{\mathbf{S}}_f$$

$$\begin{bmatrix} 1 - \gamma_i e^{j\frac{\omega d}{c}\sin\theta_0} & \cdots & 0 \\ & & \\ 0 & \cdots & 1 - \gamma_i e^{j\frac{\omega d}{c}\sin\theta_{P-1}} \end{bmatrix} \begin{bmatrix} \mathbf{v}_i^H \mathbf{a}_0 \\ . \\ . \\ . \\ \mathbf{v}_i^H \mathbf{a}_{i-1}, \\ \mathbf{v}_i^H \mathbf{a}_{i+1} \\ . \\ . \\ . \\ \mathbf{v}_i^H \mathbf{a}_{P-1} \end{bmatrix} = 0$$

where $\hat{\mathbf{S}}_f$ represents the source matrix where the ith column and the ith row are deleted. Since $\gamma_i \neq e^{-j\frac{\omega_0}{c}\Delta.\delta_k}, i \neq k$, the diagonal matrix will be full rank and also $\hat{\mathbf{S}}_f$ is full rank by assumption, we must have $\mathbf{v}_i^H \mathbf{a}_{k,} = 0$ for all $k \neq i$. Thus, the generalized signal eigenvector of the pencil matrix $\left\{ \mathbf{S}_{f_1 f_1}^0, \mathbf{S}_{f_1 f_2} \right\}$ is orthogonal to all direction vectors except the ith direction vector. We shall exploit this property for signal separation and estimation later in chapter 6.

5.1.4 <u>Diffused Source</u>: A point source, on account of local scattering or fluctuating medium or reflections from an uneven surface, may appear as a diffused source, that is, the main source surrounded by many secondary sources. A sensor array will receive a large number of rays, all of which would have actually started from the same point (main source) but have traveled different paths due to scattering or refraction or reflection. Such a model was considered in chapter 1 (p. 49). Here we shall explore the possibilities of localizing such a diffused source, in particular, a diffused source due to local scattering (see fig. 1.29). The model of the diffused source signal is given by (1.65)

$$\mathbf{f}(t) = \tilde{\mathbf{a}}s(t) + \eta(t) \qquad (5.25)$$

where $\tilde{\mathbf{a}} = \sum_{l=0}^{L-1} \alpha_l e^{-j\omega_c \delta t_l} \mathbf{a}(\theta + \delta\theta_l)$ where α_l, δt_l and $\delta\theta_l$ are for the lth ray, complex amplitude, time delay and direction of arrival with respect to a direct ray from the main source, respectively. Assume that there are L rays reaching the sensor array. For the direct ray we shall assume that $\alpha_0 = 1$, $\delta t_0 = 0$ and $\delta\theta_0 = 0$. The covariance matrix of the array (ULA) output is given by (see eq. 1.70)

$$\mathbf{c}_f \approx \sigma_s^2 L \sigma_\alpha^2 \mathbf{D}(\theta) \mathbf{Q} \mathbf{D}^H(\theta) + \sigma_\eta^2 \mathbf{I} \qquad (5.26)$$

where we have assumed that the scatterers are uniformly distributed in the angular range $\pm\Delta$. The $\mathbf{Q}$ matrix for the assumed uniform distribution is given by

$$\{\mathbf{Q}\}_{mn} = \frac{\sin 2\pi \dfrac{d}{\lambda}(m-n)\Delta\cos\theta}{2\pi \dfrac{d}{\lambda}(m-n)\Delta\cos\theta}$$

and

$$\mathbf{D} = diag\left[1, e^{j\frac{2\pi d}{\lambda}\sin\theta}, \ldots, e^{j\frac{2\pi(M-1)d}{\lambda}\sin\theta} \right]$$

The eigendecomposition of $\mathbf{Q}$ shows some interesting properties [11]. There are r, where $r = \left[2\dfrac{d}{\lambda} M\Delta\cos\theta \right]$, significant eigenvalues close to unity ([x] stands for the largest integer greater than x). The remaining (M-r) eigenvalues

are insignificant. Let $\mathbf{v}_\eta$ be a matrix whose columns are M-r eigenvectors corresponding to M-r insignificant eigenvalues (cf. the noise subspace in the Music algorithm). Now it follows that

$$\mathbf{v}_\eta^H \left[\mathbf{c}_f - \sigma_\eta^2 \mathbf{I} \right] \mathbf{v}_\eta = \mathbf{0} \quad \text{(a matrix M-r x M-r zeros)} \qquad (5.27)$$

Using (5.26) in (5.27) we obtain

$$\sigma_s^2 L \sigma_\alpha^2 \sum_{m=0}^{r-1} \lambda_m \mathbf{v}_\eta^H \mathbf{D}(\theta) \mathbf{e}_m \mathbf{e}_m^H \mathbf{D}^H(\theta) \mathbf{v}_\eta = \mathbf{0} \qquad (5.28)$$

where we have used eigendecomposition of $\mathbf{Q} \approx \sum_{m=0}^{r-1} \lambda_m \mathbf{e}_m \mathbf{e}_m^H$. Here λ_m (≈ 1) are significant eigenvalues of $\mathbf{Q}$. Since $\sigma_{f_0}^2$, and $\lambda_m > 0$ for all m<r, the following must hold good for all m<r:

$$\left| \mathbf{v}_\eta^H \mathbf{D}(\theta) \mathbf{e}_m \right|^2 = 0 \qquad (5.29)$$

for all m<r. The azimuth of the center of the cluster can be estimated using the orthogonality property demonstrated in (5.29). The eigenvectors $\mathbf{e}_m$, discrete prolate spheroidal sequence (DPSS), are obtained by eigendecomposition of the $\mathbf{Q}$ matrix. But in $\mathbf{Q}$ there is an unknown parameter pertaining to the width of the cluster, namely, $\Delta \cos\theta$ which has to be estimated. We have in chapter 1 (page 59) indicated that the rank of the covariance matrix (noise free) is closely related to this unknown parameter.

5.1.5 <u>Adaptive Subspace</u>: Adaptive methods for the estimation of the eigenvector corresponding to the minimum (or maximum) eigenvalue have been described by many researchers[12-14]. The method is based on inverse power iteration [15]

$$\hat{\mathbf{v}}^0 = [1, 0, ..., 0]^T$$
$$\tilde{\mathbf{v}}^{k+1} = \mathbf{S}_f^{-1}(\omega) \hat{\mathbf{v}}^k \qquad (5.30)$$
$$\hat{\mathbf{v}}^{k+1} = \frac{\tilde{\mathbf{v}}^{k+1}}{\sqrt{\tilde{\mathbf{v}}^{k+1^T} \tilde{\mathbf{v}}^{k+1}}}$$

where $\hat{\mathbf{v}}^k$ is an eigenvector at k^{th} iteration and the last equation is meant for normalization. To estimate the eigenvector corresponding to the largest

eigenvalue the algorithm shown in (5.30) may be used with the difference that the spectral matrix, instead of its inverse, is used.

An attempt to extend the inverse power iteration method to the estimation of the entire subspace (signal or noise subspace) has also been reported [13]. All the above methods require computation of the gradient of a cost function which is minimized. The gradient computation is slow; furthermore, in the presence of noise, it is unstable. Additionally, the rate of convergence of a gradient method is known to be very slow [14]. The alternate approach is to use the well-known power method for estimation of an eigenvector corresponding to the minimum (or maximum) eigenvalue [15]. The power method is easy to understand and easy to implement on a computer. We shall briefly describe the power method for estimation of the entire signal or noise subspace. Let $\hat{\mathbf{v}}_m^N, m = 1, 2, \dots P$ be the eigenvectors based on the data till time N, that is, the eigenvectors of the spectral matrix, $\hat{\mathbf{S}}_f^N(\omega)$. When a new sample arrives the eigenvectors are updated using the following recursive equation followed by Gram-Schmidt orthonormalization (GSO),

$$\hat{\mathbf{S}}_f^{N+1}\hat{\mathbf{v}}_m^N = \mathbf{g}_m^{N+1}$$
$$\hat{\mathbf{v}}_m^{N+1} = GSO\{\mathbf{g}_m^{N+1}\}, \quad m = 1, 2, \dots P$$

(5.31)

The choice of initial eigenvectors to start the recursion is important for rapid convergence. A suggested choice in [14] is to orthonormalize P initial frequency snapshots and use them as the initial eigenvectors. The starting value of the spectral matrix is chosen as

$$\hat{\mathbf{S}}_f^P = (\mathbf{I} + \mathbf{F}_0 \mathbf{F}_0^H)$$

The Gram-Schmidt orthonormalization process involves finding p orthonormal vectors which are linearly equivalent to data vectors [50]. Let $\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_P$ be a set of frequency snapshots and $\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_P$ be the set of linearly equivalent orthonormal vectors which satisfy the following conditions

$$\mathbf{I}_k \perp \mathbf{F}_r \ for \ r < k$$

$$|\mathbf{I}_k|^2 = 1 \ for \ all \ k$$

(5.32)

The Gram-Schmidt method determines the orthonormal vectors by solving a system of linear equations

$$\mathbf{I}_1 = \gamma_1^1 \mathbf{F}_1$$

$$\mathbf{I}_2 = \gamma_1^2 \mathbf{F}_1 + \gamma_2^2 \mathbf{F}_2$$

$$\mathbf{I}_3 = \gamma_1^2 \mathbf{F}_1 + \gamma_2^2 \mathbf{F}_2 + \gamma_3^2 \mathbf{F}_3$$

$$\dots \qquad\qquad (5.33)$$

$$\mathbf{I}_P = \gamma_1^P \mathbf{F}_1 + \gamma_2^P \mathbf{F}_2 + \dots + \gamma_P^P \mathbf{F}_P$$

§5.2 **Subspace Methods** (Broadband):

Since the location parameters are independent of frequency, we ought to get identical location estimates at different frequencies. But, on account of errors in the estimation of the spectral matrix, the location estimates are likely to be different at different frequencies. The errors in the estimation of a spectral matrix are likely to be uncorrelated random variables. By suitably combining the spectral matrix estimates obtained at different frequencies, it is hoped that the error in the final location estimate can be reduced. This is the motivation for broadband processing. In the process of combining all spectral matrices it is important that the basic structure of the spectral matrices should not be destroyed. For example, a spectral matrix at all signal frequencies is of rank one when there is only one source or rank two when there are two uncorrelated sources, and so on. Consider a ULA and a single source. The spectral matrix at frequency ω is given by $\mathbf{S}(\omega) = \mathbf{a}(\omega, \theta) S_0(\omega) \mathbf{a}^H(\omega, \theta) + \sigma_\eta^2 \mathbf{I}$. Notice that the frequency dependence of direction vectors is explicitly shown. Let us consider a generalized linear transformation of the form, $\mathbf{T}(\omega)\mathbf{S}(\omega)\mathbf{T}^H(\omega)$, where $\mathbf{T}(\omega)$ is a transformation matrix, yet to be chosen. The transformation matrix must map the direction vector into another direction vector at a chosen frequency and the background noise continues to remain white. The transformed spectral matrix is given by

$$\mathbf{T}(\omega)\mathbf{S}(\omega)\mathbf{T}^H(\omega) =$$
$$\mathbf{T}(\omega)\mathbf{a}(\omega, \theta) S_0(\omega) \mathbf{a}^H(\omega, \theta)\mathbf{T}^H(\omega) + \sigma_\eta^2 \mathbf{T}(\omega)\mathbf{T}^H(\omega) \qquad (5.34)$$

where the transformation matrix $\mathbf{T}(\omega)$ must possess the following properties:

$$\mathbf{T}(\omega)\mathbf{a}(\omega, \theta) = \mathbf{a}(\omega_0, \theta)$$
$$\mathbf{T}(\omega)\mathbf{T}^H(\omega) = \mathbf{I} \qquad (5.35)$$

where ω_0 is a selected frequency. Using (5.35) in (5.34) and averaging over a band of frequencies we obtain

$$\overline{\mathbf{S}}(\omega_0) = \sum_{i \in bw} \mathbf{T}(\omega_i)\mathbf{S}(\omega_i)\mathbf{T}^H(\omega_i)$$

$$= \mathbf{a}(\omega_0,\theta) \sum_{i \in bw} S_0(\omega_i)\mathbf{a}^H(\omega_0,\theta) + \sigma_\eta^2 \mathbf{I} \qquad (5.36)$$

where bw stands for the signal bandwidth. In (5.36) we observe that the signal power spread over a band of frequencies has been focused at one frequency, namely, ω_0. This idea of focusing of energy has been actively pursued in [16, 17] for DOA estimation.

In the direction vector there is a parameter, namely, sensor spacing which may be changed according to the frequency such that $\dfrac{\omega d}{c}$ remains constant. This will require the sensor spacing at frequency ω_i should be equal to $d_i = \dfrac{\omega_0 d_0}{\omega_i}, i \in bw$. In practice the required change in the physical separation is difficult to achieve, but resampling of the wavefield through interpolation is possible. This approach to focusing was suggested in [18]. In chapter 2 we introduced the spatio temporal covariance matrix (STCM) which contains all spatio temporal information present in a wavefield. It has been extensively used for source localization [19]. We shall probe into the signal and noise subspace structure of STCM of ULA as well as UCA and show how the structure can be exploited for source localization.

5.2.1 <u>Wideband Focusing</u>: Consider P uncorrelated point sources in a far field and a ULA for DOA estimation. As in (4.12b) the spectral matrix of an array signal may be expressed as $\mathbf{S}_f(\omega) = \mathbf{A}(\omega,\theta)\mathbf{S}_0(\omega)\mathbf{A}^H(\omega,\theta)$ where the columns of $\mathbf{A}(\omega,\theta)$ matrix are the direction vectors. We seek a transformation matrix $\mathbf{T}(\omega)$ which will map the direction vectors at frequency ω into direction vectors at the preselected frequency ω_0 as in (5.35). There is no unique solution but a least squares solution is possible,

$$\mathbf{T}(\omega) \approx \mathbf{A}(\omega_0,\theta)\left[\mathbf{A}^H(\omega,\theta)\mathbf{A}(\omega,\theta)\right]^{-1}\mathbf{A}^H(\omega,\theta) \qquad (5.37)$$

The transformation matrix given by (5.37) depends upon the unknown azimuth information. However, it is claimed in [20] that approximate estimates obtained through beamformation or any other simple approach are adequate. As an example, consider a single source case. For this, $\mathbf{A}(\omega,\theta) =$

$$col\left\{1,e^{-j\omega\frac{d}{c}\sin\theta},...,e^{-j\omega(M-1)\frac{d}{c}\sin\theta}\right\}$$ and the transformation matrix given by

(5.37) is equal to $\mathbf{T}(\omega)\approx\frac{1}{N}\mathbf{A}(\omega_0,\theta)\mathbf{A}^H(\omega,\theta)$. To show that this is not a unique answer, consider the following transformation matrix,

$$\mathbf{T}(\omega)=diag\left\{1,e^{-j(\omega_0-\omega)\frac{d}{c}\sin\theta},...,e^{-j(\omega_0-\omega)(M-1)\frac{d}{c}\sin\theta}\right\} \qquad (5.38a)$$

Clearly, using (5.38a) we can also achieve the desired transformation. In (5.38a) let $\omega_0=0$; the transformation matrix becomes

$$\mathbf{T}(\omega)=diag\left\{1,e^{j\omega\frac{d}{c}\sin\theta},...,e^{j\omega(M-1)\frac{d}{c}\sin\theta}\right\} \qquad (5.38b)$$

which we shall use as a filter on the array output. The filtered output is given by

$$\mathbf{f}_1(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathbf{T}(\omega)\mathbf{F}(\omega)e^{j\omega t}d\omega$$

$$=col\left\{f(t),f(t-\frac{d}{c}\sin\theta),...,f(t-\frac{d}{c}(M-1)\sin\theta)\right\} \qquad (5.39)$$

which indeed is the same as the progressively delayed array output in beamformation processing. Let us next compute the covariance function of the filtered output,

$$\mathbf{c}_{f_1}=E\left\{\mathbf{f}_1(t)\mathbf{f}_1^H(t)\right\}$$

$$=E\left\{\begin{bmatrix}f(t-m\frac{d}{c}\sin\theta),m=0,1,...,M-1\end{bmatrix}\begin{bmatrix}f(t-m\frac{d}{c}\sin\theta),m=0,1,...,M-1\end{bmatrix}^H\right\}$$

$$=\mathbf{c}_f((m-n)\frac{d}{c}(\sin\theta_0-\sin\theta))=\begin{bmatrix}\mathbf{c}_f\end{bmatrix}_{mn}$$

where $\mathbf{c}_f$ is the covariance matrix of the signal emitted by the source. $\mathbf{c}_{f_1}$ is known as the steered covariance matrix [21]. When $\theta = \theta_0$ all elements of $\mathbf{c}_{f_1}$ are equal to a constant equal to the total signal power, that is, $\frac{1}{2\pi}\int_{-\infty}^{\infty}S_f(\omega)d\omega$. Thus, in beamformation we seem to focus all signal power to zero frequency. When there is more than one source we should compute a series of steered covariance matrices over a range of azimuth angles. Whenever a steering angle matches with one of the directions of arrival the steered covariance matrix will display a strong 'dc' term, equal to the power of the source.

The main drawback of wideband focusing is the requirement that the DOAs of incident wavefronts must be known, at least approximately, but the resulting estimate is likely to have large bias and variance [20]. To overcome this drawback an alternate approach has been suggested in [22]. Let $\mathbf{A}$ and $\mathbf{B}$ be two square matrices and consider a two sided transformation, $\mathbf{TBT}^H$, which is closest to $\mathbf{A}$. It is shown in [22] that this can be achieved if $\mathbf{T} = \mathbf{v}_A\mathbf{v}_B^H$ where $\mathbf{v}_A$ is a matrix whose columns are eigenvectors of $\mathbf{A}$ and $\mathbf{v}_B$ is similarly the eigenvector matrix of $\mathbf{B}$. Now let $\mathbf{A}$ and $\mathbf{B}$ be the spectral matrices at two frequencies, $\mathbf{A} = \mathbf{S}_f(\omega_0)$ and $\mathbf{B} = \mathbf{S}_f(\omega_i)$. To transform $\mathbf{S}_f(\omega_i)$ into $\mathbf{S}_f(\omega_0)$ the desired matrix is

$$\mathbf{T}(\omega_0,\omega_i) = \mathbf{v}_s(\omega_0)\mathbf{v}_s^H(\omega_i) \qquad (5.40)$$

where $\mathbf{v}_s$ is a matrix whose columns are signal eigenvectors. It is easy to show that $\mathbf{T}(\omega_0,\omega_i)\mathbf{T}^H(\omega_0,\omega_i) = \mathbf{I}$. Applying the transformation on a spectral matrix at frequency ω_i we get

$$\mathbf{T}(\omega_0,\omega_i)\mathbf{S}_f(\omega_i)\mathbf{T}^H(\omega_0,\omega_i)$$
$$= \mathbf{v}_s(\omega_0)\mathbf{v}_s^H(\omega_i)\mathbf{S}_f(\omega_i)\mathbf{v}_s(\omega_i)\mathbf{v}_s^H(\omega_0)$$
$$= \mathbf{v}_s(\omega_0)\lambda(\omega_i)\mathbf{v}_s^H(\omega_0)$$

where

$$\lambda(\omega_i) = diag\{\lambda_0(\omega_i),\lambda_1(\omega_i),...,\lambda_{P-1}(\omega_i),0,...,0\},$$

are the eigenvalues of the noise free spectral matrix, $\mathbf{S}_f(\omega_i)$. Next, we average all transformed matrices and show that

$$\overline{\mathbf{S}}_f(\omega_0) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{T}(\omega_0, \omega_i) \mathbf{S}_f(\omega_i) \mathbf{T}^H(\omega_0, \omega_i)$$

$$= \mathbf{v}_s(\omega_0) \frac{1}{N} \sum_{i=1}^{N} \lambda(\omega_i) \mathbf{v}_s^H(\omega_0) \qquad (5.41)$$

$$= \mathbf{v}_s(\omega_0) \overline{\lambda} \mathbf{v}_s^H(\omega_0)$$

$\overline{\mathbf{S}}_f(\omega_0)$ is the focused spectral matrix whose eigenvectors are the same as those of $\mathbf{S}_f(\omega_0)$ (before focusing) but its eigenvalues are equal to the averaged eigenvalues of all spectral matrices taken over the frequency band. We may now use the focused spectral matrix for DOA estimation.

5.2.2 <u>Spatial Resampling</u>: The basic idea is to resample the wavefield sensed by a fixed ULA so as to create a virtual ULA with a different sensor spacing depending upon the frequency. The spectral matrices computed at different frequencies are then simply averaged. Let us take a specific case of two wideband uncorrelated sources in the far field. The spectral matrix is given by

$$\mathbf{S}_f(\omega) = \begin{bmatrix} \mathbf{a}(\omega \dfrac{d}{c} \sin\theta_0) S_0(\omega) \mathbf{a}^H(\omega \dfrac{d}{c} \sin\theta_0) + \\ \mathbf{a}(\omega \dfrac{d}{c} \sin\theta_1) S_1(\omega) \mathbf{a}^H(\omega \dfrac{d}{c} \sin\theta_1) \end{bmatrix} + \sigma_\eta^2 \mathbf{I}$$

where $S_0(\omega)$ and $S_1(\omega)$ are spectra of the first and the second source respectively. We compute the spectral matrix at two different frequencies, ω_0 and ω_1, with different sensor spacings, d_0 and d_1, where

$$d_1 = d_0 \frac{\omega_0}{\omega_1}$$

and form a sum.

$$\mathbf{S}_f(\omega_0) + \mathbf{S}_f(\omega_1) =$$

$$\begin{Bmatrix} \mathbf{a}(\omega_0 \dfrac{d_0}{c} \sin\theta_0) S_0(\omega_0) \mathbf{a}^H(\omega_0 \dfrac{d_0}{c} \sin\theta_0) \\ + \mathbf{a}(\omega_0 \dfrac{d_0}{c\,0} \sin\theta_1) S_1(\omega_0) \mathbf{a}^H(\omega_0 \dfrac{d_0}{c} \sin\theta_1) \end{Bmatrix} + \sigma_\eta^2 \mathbf{I}$$

$$+\left\{\begin{array}{l} \mathbf{a}(\omega_1 \frac{d_1}{c}\sin\theta_0)S_0(\omega_1)\mathbf{a}^H(\omega_1 \frac{d_1}{c}\sin\theta_0) \\ +\mathbf{a}(\omega_1 \frac{d_1}{c}\sin\theta_1)S_1(\omega_1)\mathbf{a}^H(\omega_1 \frac{d_1}{c}\sin\theta_1) \end{array}\right\}+\sigma_\eta^2\mathbf{I}$$

$$= \mathbf{a}(\omega_0 \frac{d_0}{c}\sin\theta_0)\left[S_0(\omega_0)+S_0(\omega_i)\right]\mathbf{a}^H(\omega_0 \frac{d_0}{c}\sin\theta_0)$$

$$+\mathbf{a}(\omega_0 \frac{d_0}{c}\sin\theta_1)\left[S_1(\omega_0)+S_1(\omega_i)\right]\mathbf{a}^H(\omega_0 \frac{d_0}{c}\sin\theta_1)+2\sigma_\eta^2\mathbf{I}$$

$$\mathbf{S}_f(\omega_0)+\mathbf{S}_f(\omega_1) =$$

$$\left\{\begin{array}{l} \mathbf{a}(\omega_0 \frac{d_0}{c}\sin\theta_0)S_0(\omega_0)\mathbf{a}^H(\omega_0 \frac{d_0}{c}\sin\theta_0) \\ +\mathbf{a}(\omega_0 \frac{d_0}{c_0}\sin\theta_1)S_1(\omega_0)\mathbf{a}^H(\omega_0 \frac{d_0}{c}\sin\theta_1) \end{array}\right\}+\sigma_\eta^2\mathbf{I}$$

$$+\left\{\begin{array}{l} \mathbf{a}(\omega_1 \frac{d_1}{c}\sin\theta_0)S_0(\omega_1)\mathbf{a}^H(\omega_1 \frac{d_1}{c}\sin\theta_0) \\ +\mathbf{a}(\omega_1 \frac{d_1}{c}\sin\theta_1)S_1(\omega_1)\mathbf{a}^H(\omega_1 \frac{d_1}{c}\sin\theta_1) \end{array}\right\}+\sigma_\eta^2\mathbf{I} \qquad (5.42)$$

$$= \mathbf{a}(\omega_0 \frac{d_0}{c}\sin\theta_0)\left[S_0(\omega_0)+S_0(\omega_i)\right]\mathbf{a}^H(\omega_0 \frac{d_0}{c}\sin\theta_0)$$

$$+\mathbf{a}(\omega_0 \frac{d_0}{c}\sin\theta_1)\left[S_1(\omega_0)+S_1(\omega_i)\right]\mathbf{a}^H(\omega_0 \frac{d_0}{c}\sin\theta_1)+2\sigma_\eta^2\mathbf{I}$$

Let us assume that N virtual arrays, each with different sensor spacing, have been created by a resampling process. We can generalize (5.42),

$$\frac{1}{N}\sum_{i=0}^{N-1}\mathbf{S}_f(\omega_i) = \mathbf{a}(\omega_0 \frac{d_0}{c}\sin\theta_0)\frac{1}{N}\sum_{i=0}^{N-1}\mathbf{S}_0(\omega_i)\mathbf{a}^H(\omega_0 \frac{d_0}{c}\sin\theta_0)$$

$$+ \mathbf{a}(\omega_0 \frac{d_0}{c}\sin\theta_1)\frac{1}{N}\sum_{i=0}^{N-1}\mathbf{S}_1(\omega_i)\mathbf{a}^H(\omega_0 \frac{d_0}{c}\sin\theta_1)+\sigma_\eta^2\mathbf{I}$$

$$(5.43)$$

The above procedure for focusing may be extended to P sources; some of them may be correlated. While the idea of resampling is conceptually simple

its implementation would require us to perform interpolation of the wavefield in between physical sensors. Fortunately, this is possible because in a homogeneous medium the wavefield is limited to a spatial frequency range,

$u^2 + v^2 \leq \dfrac{\omega^2}{c^2}$ (see chapter 1, p. 14). In two dimensions, the spatial frequency

range is $-\dfrac{\omega}{c} \leq u \leq \dfrac{\omega}{c}$; hence the wavefield is spatially bandlimited. This fact

has been exploited in [21] for resampling.

The interpolation filter is a lowpass filter with its passband given by

$-\dfrac{\omega}{c} \leq u \leq \dfrac{\omega}{c}$. Consider a ULA with sensor spacing equal to d_0 ($d_0 = \dfrac{\lambda}{2}$).

The maximum temporal frequency will be $\omega_{max} = \pi \dfrac{c}{d_0}$. The low pass filter

is given by

$$h(m) = \frac{\sin \omega_{max} \dfrac{(x - md_0)}{c}}{\omega_{max} \dfrac{(x - md_0)}{c}}, \quad m = 0, \pm 1, \ldots, \pm \infty \qquad (5.44)$$

where x denotes the point where interpolation is desired, for example,

$x = m'd_k$ where $d_k = \dfrac{d_0 \omega_{max}}{\omega_k}$.

5.2.3 Spatio Temporal Covariance Matrix (STCM): In chapter 2 we introduced the concept of extended direction vector (chapter 2, p. 129) and STCM using the extended direction vector and the source spectrum (2.67) which we reproduce here for convenience,

$$\mathbf{C}_{STCM} =$$

$$E\left\{\mathbf{f}_{stacked}(t)\mathbf{f}_{stacked}^{H}(t)\right\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{h}(\omega, \varphi_0)\mathbf{h}^{H}(\omega, \varphi_0)S_{f_0}(\omega)d\omega \qquad (2.67)$$

In this subsection we shall show how the STCM can be used for the direction of arrival estimation of a broadband source. For this it is necessary that the eigenstructure, in particular, the rank of STCM, will have to be ascertained. This has been done for a ULA in [23] and for a UCA in [24, 4]. We shall assume that the source spectrum is a smooth function and that it may be approximated by a piecewise constant function

$$S_{f_0}(\omega) \approx \sum_{l=-L}^{l=L} S_{0l} rect(\omega - \Delta\omega l) \tag{5.45}$$

where

$$rect(\omega - \Delta\omega l) = \frac{1}{\Delta\omega} \quad for \quad \Delta\omega(l - \frac{1}{2}) \leq \omega \leq \Delta\omega(l + \frac{1}{2})$$
$$= 0 \quad otherwise$$

and S_{0l} $(S_{0,l} = S_{0,-l})$ is the average spectrum in the lth frequency bin, $\Delta\omega l$. Using (5.45) in (2.67) we obtain

$$\mathbf{C}_{STCM} = \frac{1}{2\pi} \sum_{l=-L}^{l=L} S_{0l} \frac{1}{\Delta\omega} \int_{-\Delta\omega(l-\frac{1}{2})}^{\Delta\omega(l-\frac{1}{2})} \mathbf{h}(\omega,\varphi_0)\mathbf{h}^H(\omega,\varphi_0)d\omega \tag{5.46}$$

where we have divided the frequency interval $\pm\pi$ into 2L+1 nonoverlapping frequency bins $(= \dfrac{2\pi}{\Delta\omega})$. Using the definition of the extended direction vector given in (2.67) in the integral in (5.46) we obtain the following result:

$$\frac{1}{2\pi\Delta\omega} \int_{-\Delta\omega(l-\frac{1}{2})}^{\Delta\omega(l-\frac{1}{2})} \mathbf{h}(\omega,\varphi_0)\mathbf{h}^H(\omega,\varphi_0)d\omega = \mathbf{g}_l \mathbf{Q} \mathbf{g}_l^H$$

where

$$\mathbf{g}_l =$$

$$diag\left\{ \begin{array}{l} e^{-j\omega\tau_0}, e^{-j\omega\tau_1}, ..., e^{-j\omega\tau_{M-1}}; e^{-j\omega(\Delta t+\tau_0)}, e^{-j\omega(\Delta t+\tau_1)}, ...,\\ e^{-j\omega(\Delta t+\tau_{M-1})}; ...; e^{-j\omega((N-1)\Delta t+\tau_0)}, e^{-j\omega((N-1)\Delta t+\tau_1)}, ...,\\ e^{-j\omega((N-1)\Delta t+\tau_{M-1})} \end{array} \right\} \Bigg|_{\omega=\Delta\omega l}$$

and

$$[\mathbf{Q}]_{\alpha,\alpha'} = \sin c\left[((n-n')\Delta t + \tau_m - \tau_{m'})\frac{\Delta\omega}{2} \right]$$

where $\alpha = m + n \times M$, $\alpha' = m' + n' \times M$, $n, n' = 0, 1, \ldots, N-1$ and $m, m' = 0, 1, \ldots, M-1$. Note that the time delays $(\tau_0, \tau_1, \ldots, \tau_{M-1})$ depend upon the array geometry and, of course, on the azimuth; for example, for a linear array, the time delays are $(0, \dfrac{d}{c}\sin\varphi, \ldots, (M-1)\dfrac{d}{c}\sin\varphi)$ and for a circular array the time delays are $(\delta\cos\varphi, \delta\cos(\dfrac{2\pi}{M} - \varphi), \ldots,$

$\delta\cos(\dfrac{2\pi}{M}(M-1) - \varphi))$ where $\delta = \dfrac{a}{c}$ and a is the radius of the circular array. Equation (5.46) may be expressed as

$$\mathbf{C}_{STCM} = \sum_{l=-L}^{l=L} S_{0l}\mathbf{g}_l\mathbf{Q}\mathbf{g}_l^H \tag{5.47}$$

The rank of $\mathbf{C}_{STCM}$ is determined by the rank of matrix $\mathbf{Q}$. It follows from the results in [11, 24] that 99.99% energy is contained in the first $\left[((N-1)\Delta t + \dfrac{2a}{c})\dfrac{\Delta\omega}{2\pi} + 1\right]$ eigenvalues of $\mathbf{Q}$ where [x] denotes the next integer greater than x. A comparison of the theoretical and numerically determined rank of $\mathbf{C}_{STCM}$ is given in fig. 5.5 for a circular array. For a linear array the corresponding number is given by $\left[((N-1)\Delta t + \dfrac{(M-1)d\cos\varphi}{c})\dfrac{\Delta\omega}{2\pi} + 1\right]$ [19]. Note that the dimension of the signal subspace is approximately equal to the time bandwidth product.

When the size of the observation space, that is, dimensions of $\mathbf{C}_{STCM}$, is larger than the rank of $\mathbf{C}_{STCM}$, $NM > \left[((N-1)\Delta t + \dfrac{2a}{c})\dfrac{\Delta\omega}{2\pi} + 1\right]$, there exists a null subspace, a subspace of the observation space, of dimension equal to

$$Dim\{\mathbf{v}_{null}\} = NM - \left[((N-1)\Delta t + \dfrac{2a}{c})\dfrac{\Delta\omega}{2\pi} + 1\right]$$

where $\mathbf{v}_{null}$ eigenvectors correspond to the insignificant eigenvalues. Note that the null subspace thus defined will not be an ideal null space with zero power; some residual power (<0.01%) will be present.
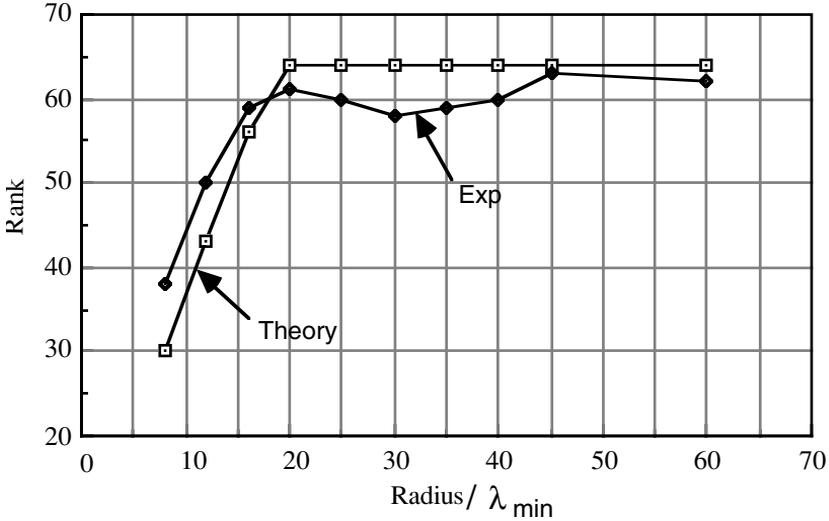
Figure 5.5: The effective rank of signal only STCM as a function of the radius of a circular array. The drop in the rank at radius $30\lambda_{min}$ is due to a transition from smooth spectrum to line spectrum. We have considered a sixteen sensor equispaced circular array with other parameters N=4, B=0.8 (normalized bandwidth) and one source.(From [4]. © 1994, With permission from Elsevier Science)

Now, consider the following quadratic,

$$
\mathbf{v}_i^H \mathbf{C}_{STCM} \mathbf{v}_i = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{v}_i^H \mathbf{h}(\omega, \varphi_0) \mathbf{h}^H(\omega, \varphi_0) \mathbf{v}_i S_{f_0}(\omega) d\omega
$$

$$
= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \mathbf{v}_i^H \mathbf{h}(\omega, \varphi_0) \right|^2 S_{f_0}(\omega) d\omega \approx 0
$$

(5.48)

Assume that $S_{f_0}(\omega) > 0$ over some frequency band. For (5.48) to hold good we must have in that frequency band

$$
\left| \mathbf{v}_i^H \mathbf{h}(\omega, \varphi_0) \right|^2 \approx 0 \quad i \in null\, space
$$

(5.49)

Equation (5.49) is the basic result used for estimation of φ_0. For the purpose of locating the null we can define a parametric spectrum as in narrowband MUSIC.

5.2.4 <u>Number of Sources</u>: We have so far assumed that the number of sources is known *a priori* or can be estimated from the knowledge of the significant and the repeated eigenvalues of the spectral matrix. The need for this information arises in the estimation of signal and noise subspaces. In time series analysis the equivalent problem is estimation of the number of sinusoids or estimation of the order of a time series model. A lot of effort has gone into this problem in time series analysis [25]. In this section we shall describe a test known as the sphericity test and its modern versions for estimating the number of repeated eigenvalues. Let P plane wavefronts be incident on an M sensor array. A covariance matrix of size $M \times M$ (M>P) has been computed using a finite number of N snapshots. Let $\hat{\lambda}_1 > \hat{\lambda}_2 > \hat{\lambda}_3 > ... > \hat{\lambda}_M$ be the eigenvalues of the estimated spectral matrix. We begin with the hypothesis that there are p wavefronts; then a subset of smaller eigenvalues, $\hat{\lambda}_{p+1} > \hat{\lambda}_{p+2} > ... > \hat{\lambda}_M$, is equal (or repeated). The null hypothis, that the smallest eigenvalues has a multiplicity of M-p, is tested starting at p=0 till the test fails. Mauchley [26] suggested a test known as the sphericity test to verify the equality of all eigenvalues of an estimated covariance matrix. Define a likelihood ratio statistic

$$\Gamma(\hat{\lambda}_{p+1}, \hat{\lambda}_{p+2},...,\hat{\lambda}_M) = \ln\left[\frac{\left(\frac{1}{M-p}\sum_{i=p+1}^{M}\hat{\lambda}_i\right)^{M-p}}{\prod_{i=p+1}^{M}\hat{\lambda}_i}\right]^N \qquad (5.50a)$$

where N stands for number of snapshots. The log likelihood ratio is then compared with a threshold γ. Whenever

$$\Gamma(\hat{\lambda}_{p+1}, \hat{\lambda}_{p+2},...,\hat{\lambda}_M) > \gamma$$

the test is said to have failed. Modified forms of the sphericity tests have been suggested in References [27, 28].

The choice of γ is subjective. Alternate approaches which do not require a subjective judgment have been proposed [29]. The number of wavefronts is given by that value of p in the range 0 to $M-1$ where, either

$$\Gamma(\hat{\lambda}_{p+1}, \hat{\lambda}_{p+2},...,\hat{\lambda}_M) + p(2M-p) \qquad (5.50b)$$

or

$$\Gamma(\hat{\lambda}_{p+1}, \hat{\lambda}_{p+2}, \ldots, \hat{\lambda}_M) + \frac{p}{2}(2M - p)\ln N \qquad (5.50c)$$

is minimum. The first form is known as Akaike information criterion (AIC) and the second form is known as minimum description length (MDL). The expression consists of two parts; the first part is the log likelihood ratio which decreases monotonically and the second part is a penalty term which increases monotonically as p increases from 0 to $M-1$. The penalty term is equal to free adjusted parameters. The log likelihood ratio plot around p=P undergoes a rapid change of slope, which is responsible for the occurrence of a minimum at p=P. This is demonstrated in fig. 5.6.

The performance of AIC and MDL criteria has been studied in the context of detection of plane waves incident on a linear array in [30]. They define a probability of error as

$$prob. of\ error = prob\big(p_{\min} < P \big| H_P\big) + prob\big(p_{\min} > P \big| H_P\big)$$

where $p_{\min}$ stands for the position of the minimum of (5.50) and H_P denotes the hypothesis that the true number of signal sources is P. The probability of error for the MDL criterion monotonically goes to zero for a large number of snapshots or high snr. In contrast, the probability of error for the AIC criterion tends to a small finite value. However, for a small number of snapshots or low snr the performance of AIC is better than that of MDL.

§5.3 **Coded Signals**:
In communication a message bearing signal is specifically designed to carry maximum information with minimum degradation. In analog communication, amplitude modulated and frequency modulated signals are often used while in digital communication each pulse representing a bit is modulated with a sine wave (frequency shift keying) or with a pseudorandom sequence (spread spectrum). From the signal processing point, as these signals belong to a class of cyclostationary processes whose covariance function is periodic, but the background noise is simple ordinary stochastic process, these differences have been exploited in array processing for the direction arrival estimation. Sometimes, it is necessary to localize an active source for improved communication as in dense wireless telephone user environment. Here we would like to emphasize how a specially coded signal leads us to newer approaches to localization. We shall demonstrate with three different types of coded signals, namely, (i) multitone signal where the tones are spaced at prescribed frequency intervals but with random phases, (ii) binary phase shift keying (BPSK) signal and (iii) cyclostationary signals.
5.3.1 <u>Multitone Signal</u>: Consider a circular boundary array (see chapter 2) of M equispaced, omnidirectional, wideband sensors (fig. 5.7) with a target which is
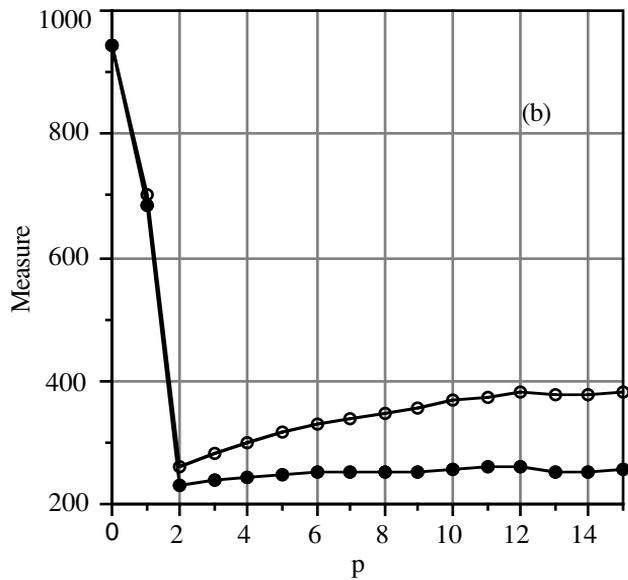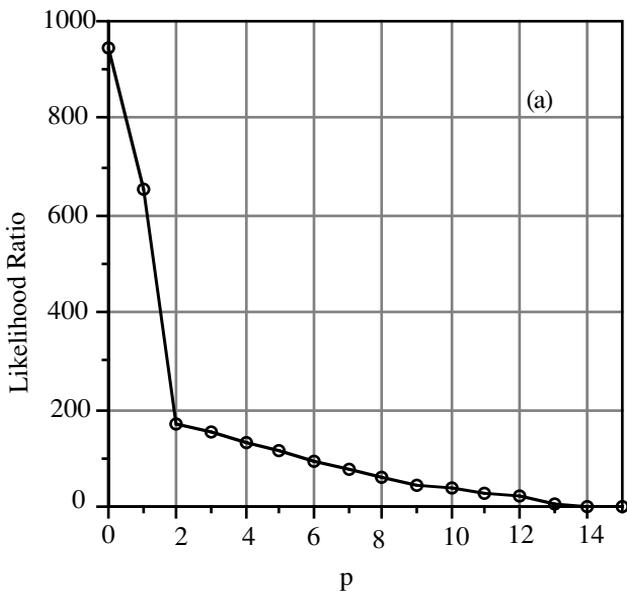
Figure 5.6: (a) Log likelihood ratio as a function of p, number of assumed wavefronts, (b) AIC (filled circles) and MDL (empty circles) as a function of p. Note the minimum at p=2. 20 frequency snapshots. 16 sensor ULA with $\lambda/2$ spacing. Two uncorrelated wavefronts are incident at angles 24° and 32°. snr=10 db per source.

surrounded by the array. A source is assumed at a point with polar coordinates (r, φ) and it radiates a signal which is a sum of harmonically related random sinusoids,

$$f(t) = \sum_{k=0}^{Q-1} \alpha_k e^{-j(\omega_0 + k\Delta\omega)t} \tag{5.51}$$

where α_k is the complex amplitude of kth sinusoid. The output of the ith sensor is given by $f_i(t) = f(t + \Delta\tau_i) + \eta_i(t)$, where $\Delta\tau_i$ is time delay at the ith sensor with respect to a fictitious sensor at the center of the array. Each sensor output is tapped at N time instants and arranged in a vector form,

$$\mathbf{f}_i = col\{f_i(0), f_i(1), \ldots f_i(N-1)\}$$

The output data vector of the ith sensor may be written in a compact form shown below:

$$\mathbf{f}_i = a(r_i)\mathbf{H}\mathbf{A}_i \varepsilon + \eta_i \tag{5.52}$$

where

$$\mathbf{H} = [\mathbf{h}_0, \mathbf{h}_1, \ldots \mathbf{h}_{Q-1}]$$
$$\mathbf{h}_k = col\{1, e^{j\omega_k}, \ldots e^{j(N-1)\omega_k}\}$$
$$\omega_k = \omega_0 + k\Delta\omega$$
$$\mathbf{A}_i = diag\{e^{j\omega_0\Delta\tau_i}, e^{j\omega_1\Delta\tau_i}, \ldots e^{j\omega_{Q-1}\Delta\tau_i}\}$$
$$\varepsilon = col[\alpha_0, \alpha_1, \ldots \alpha_{Q-1}]$$

Next, we shall stack up all data vectors into one large vector $\mathbf{F}$ of dimension MNx1. It may be expressed as

$$\mathbf{F} = \mathbf{TDE} + \eta \tag{5.53}$$

where

$$\mathbf{T} = diag\{\mathbf{H}, \mathbf{H}, \ldots \mathbf{H}\}, (\text{NMxMQ}),$$
$$\mathbf{D} = diag\{a(r_0)\mathbf{A}_0, a(r_1)\mathbf{A}_1, \ldots a(r_{M-1})\mathbf{A}_{M-1}\}, (\text{MQxMQ}) \text{ and}$$
$$\mathbf{E} = [\varepsilon^T, \varepsilon^T, \varepsilon^T, \ldots]^T, (\text{MQx1})$$

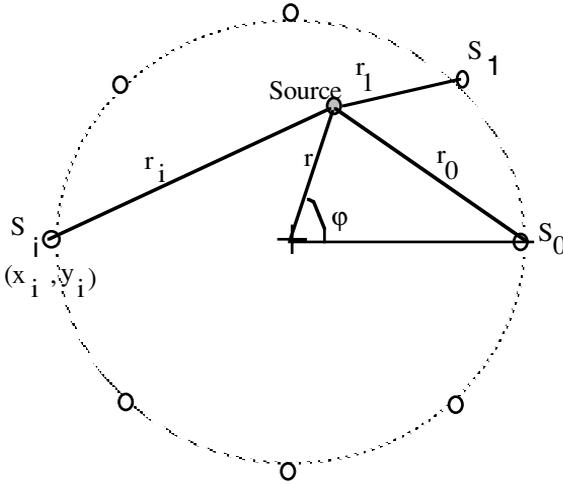The covariance matrix (STCM) of $\mathbf{F}$ is given by

Figure 5.7: A circular array of sensors and a source inside the circle. The range (r) and azimuth (φ) of the source and all time delays are measured with respect to the center of the array as shown.

$$\mathbf{c}_f = \mathbf{TD}\Gamma_0 \mathbf{D}^H \mathbf{T}^H + \sigma_\eta^2 \mathbf{I}$$

where $\Gamma_0 = E\{\mathbf{EE}^H\} = \mathbf{1} \otimes \Gamma$ where $\mathbf{1}$ is a square matrix of size MxM whose elements are all equal to 1, $\Gamma = diag\{\gamma_0, \gamma_1, \ldots \gamma_{Q-1}\}$ and $\gamma_0, \gamma_1, \ldots \gamma_{Q-1}$ are powers of the random sinusoids. Symbol $\otimes$ stands for Kronecker product. We will assume hereafter that the noise variance σ_η^2 is known or has been estimated from the array output when there is no signal transmission or by averaging the noise eigenvalues of the covariance matrix and that it has been subtracted from the covariance matrix.

Let us consider the structure of the mth column of the STCM. By straightforward multiplication it can be shown that the mth column of the covariance matrix is given by

$$\mathbf{c}_m = \left[\mathbf{c}_m^{0^T}, \mathbf{c}_m^{1^T}, \ldots, \mathbf{c}_m^{Q-1^T}\right]^T \qquad (5.54)$$

where $\mathbf{c}_m^i = a(r_i)a(r_0)\mathbf{HA}_i\Gamma\mathbf{A}_o^H\mathbf{H}^H\mathbf{u}_m$ and $\mathbf{u}_m$ is a Mx1 vector consisting of all zeros except at mth location where there is one. Note $\mathbf{A}_i\Gamma\mathbf{A}_0^H$ is a diagonal matrix given by

$$\mathbf{A}_i\Gamma\mathbf{A}_0^H = diag\{\gamma_0 e^{j\omega_0\mu_i}, \gamma_1 e^{j\omega_1\mu_i}, \dots \gamma_{Q-1}e^{j\omega_{Q-1}\mu_i}\} \qquad (5.55)$$

where $\mu_i = \Delta\tau_i - \Delta\tau_0$. Further, we note that

$$\mathbf{H}^H\mathbf{u}_m = col\{e^{-jm\omega_0}, e^{-jm\omega_1}, \dots e^{-jm\omega_{Q-1}}\} \qquad (5.56)$$

Using (5.55) and (5.56) in (5.54) we obtain

$$\mathbf{C}_m^i = a(r_i)a(r_0)\times$$
$$\mathbf{H}col\{\gamma_0 e^{j\omega_0(\mu_i-m)}, \gamma_1 e^{j\omega_1(\mu_i-m)}, \dots \gamma_{Q-1}e^{j\omega_{Q-1}(\mu_i-m)}\} \qquad (5.57)$$

Multiply on both sides of (5.57) by $\mathbf{H}^\# = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H$, the pseudo-inverse of **H,** and obtain

$$\mathbf{H}^\#\mathbf{C}_m^i = a(r_i)a(r_0)col\{\gamma_0 e^{j\omega_0(\mu_i-m)}, \gamma_1 e^{j\omega_1(\mu_i-m)}, \dots \gamma_{Q-1}e^{j\omega_{Q-1}(\mu_i-m)}\}$$

We now define a matrix **D** whose columns are $\mathbf{H}^\#\mathbf{C}_m^i$, i=0, 1, 2, ... M-1,

$$\mathbf{D} = \{\mathbf{H}^\#\mathbf{C}_m^0, \mathbf{H}^\#\mathbf{C}_m^1, \dots \mathbf{H}^\#\mathbf{C}_m^{M-1}\}$$
$$= a^2(r_0)diag\{\gamma_0 e^{-j\omega_0 m}, \gamma_1 e^{-j\omega_1 m}, \dots \gamma_{Q-1}e^{-j\omega_{Q-1}m}\} \qquad (5.58)$$

where

$$\mathbf{G} = \begin{bmatrix} 1, b_1 e^{j\omega_0\mu_1}, & \dots b_{M-1}e^{j\omega_0\mu_{M-1}} \\ 1, b_1 e^{j\omega_1\mu_1}, & \dots b_{M-1}e^{j\omega_1\mu_{M-1}} \\ & \dots \\ & \dots \\ 1, b_1 e^{j\omega_{Q-1}\mu_1}, & \dots b_{M-1}e^{j\omega_{Q-1}\mu_{M-1}} \end{bmatrix}$$

Source #1		Source #2	
Range	Azimuth	Range	Azimuth
24.99	50.01	30.00	50.02
(25.00)	(50.00)	(30.00)	(50.00)
89.99	90.00	70.01	-90.00
(90.00)	(90.00)	(70.00)	(-90.00)
9.99	29.99	50.02	150.02
(10.00)	(30.00)	(50.00)	(150.00)
50.01	-60.00	49.99	-60.00
(50.1)	(-60.00)	(50.00)	(-60.00)

Table 5.4: Localization of two sources by a circular (boundary) array of 8 sensors. snr=10 dB. The STCM was averaged over one hundred independent estimates. The wave speed was assumed as 1500 meters/sec. The numbers inside brackets represent the true values (range in meters and azimuth in degrees).

and $b_i = \dfrac{a(r_i)}{a(r_0)}$, i=0,1,...,M-1. Since the location information is present in $\mu_0, \mu_1, \ldots \mu_{M-1}$, our aim naturally is to estimate these from the columns of **G**. Each column may be considered as a complex sinusoid whose frequency can be estimated, provided $|\Delta\omega\mu_i| \le \pi$. For a boundary array, specifically a circular array, this limitation can be overcome. In chapter 2 we have given an algorithm to estimate time delays or the location of the source from the phase estimates.

As a numerical example we have considered two sources each emitting eight random tones. The frequencies emitted by the first source are (0, 200, 400, 600, 800, 1000, 1200, 1400 Hz) and those by the second are (1600, 1800, 2000, 2200, 2400, 2600, 2800, 3000 Hz). An eight sensor UCA of radius 100 meters surrounds both sources. Sixteen delayed snapshots (taps) were used making the size of STCM as 128x128. The results are displayed in table 5.4.

5.3.2 <u>BPSK Signal</u>: Binary phase shift keying signal consists of square pulses of fixed width with amplitude given by a Bernoulli random variable, an outcome of a coin tossing experiment where head equals +1 and tail equals -1. A typical BPSK sequence is shown in fig. 5.8. The analytic representation of BPSK signal is given by

$$c_0(t) = \sum_{n=0}^{L-1} c_{0,n} h_{T_c}(t - nT_c) \qquad (5.59a)$$

where $c_{0,n}$, n=0,1,...,L-1 are Bernoulli random variables and $h_{T_c}(t)$ is a
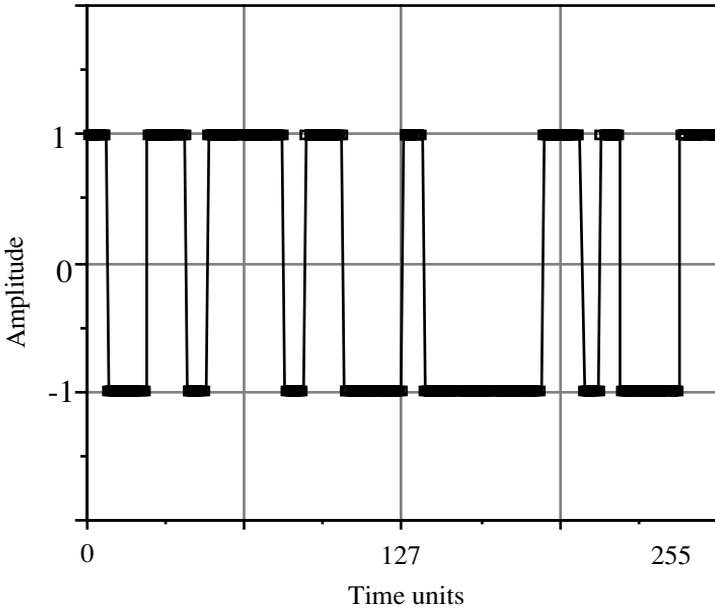
Figure 5.8: BPSK waveform. A train of rectangular pulses with amplitude alternating between +1 and -1.

rectangular function of width T_c. One of the most useful properties of a BPSK sequence is its narrow autocorrelation function with sidelobes whose variance is less than the inverse of the sequence length. We can easily generate many uncorrelated BPSK sequences. BPSK sequences have beenused for time delay measurement [31], digital communication [32], and mobile communication [33] where it is of interest to exploit the spatial diversity for the purpose of increasing the capacity. Here we shall briefly show how the azimuth of a source emitting a known BPSK signal may be estimated.

The signal model considered here is as in [34]. There are Q uncorrelated sources emitting data bits which are encoded with a user specific random sequence, a BPSK sequence, for example.

$$x_m(t) = \sum_{k=0}^{Q} p_k a_m(\theta_k) \sum_l b_{k,l} c_k(t - lT_s - \tau_k) + \eta_m(t) \qquad (5.59b)$$

$$m = 0, 1, 2, \dots M - 1$$

where

$\quad a_m(\theta_k)$: response of mth sensor to a signal coming from k[th] user

p_k: signal amplitude of kth user

$b_{k,l}$: data stream from k^{th} user

$c_k(t)$: random code of kth user

T_S: bit duration

τ_k: delay of a signal from k^{th} user

$\eta_m(t)$: Noise at m^{th} sensor

Q : number of users

There are two different approaches to DOA estimation with BPSK coded signal. In the first approach the usual covariance matrix is computed and in the second approach, due to the fact that the code used by the user of interest is known, the received signal is first cross-correlated with that code. The output of the cross-correlator is then used to compute the covariance matrix. Both approaches yield similar results; though in the second approach the interference from the users of no interest is reduced by a factor proportional to the code length.

Precorrelation Covariance Matrix: The outputs of the sensor array, after removing the carrier, are arranged in a vector form,

$$\mathbf{f}(t) = \sum_{k=0}^{Q} p_k \mathbf{a}(\theta_k) \sum_l b_{k,l} c_k(t - lT_s - \tau_k) + \eta(t) \qquad (5.60a)$$

where $\mathbf{f}(t)$, $\mathbf{a}(\theta_k)$, and $\eta(t)$ are Mx1 vectors. The precorrelation covariance matrix is given by

$$\mathbf{c}_f = E\{\mathbf{f}(t)\mathbf{f}^H(t)\} \qquad (5.60b)$$

Using (5.60a) in (5.60b) we obtain

$$\mathbf{c}_f = \left\{ \begin{matrix} \sigma_\eta^2 \mathbf{I} + \sum_{k=0}^{Q} \sum_{k'=0}^{Q} p_k p_{k'} \mathbf{a}(\theta_k) \mathbf{a}(\theta_{k'}) \\ \left\{ \sum_l \sum_{l'} b_{k,l} c_k(t - lT_s - \tau_k) b_{k',l'} c_{k'}(t - l'T_s - \tau_{k'}) \right\} \end{matrix} \right\} \qquad (5.60c)$$

We shall assume that the data bits coming from different users are independent and codes assigned to different users are uncorrelated. As a result, (5.60c) reduces to

$$\mathbf{c}_f = \left\{ \sum_{k=0}^{Q} p_k^2 \mathbf{a}(\theta_k)\mathbf{a}(\theta_k) \left\{ \sum_{l} b_{k,l}^2 c_k^2(t - lT_s - \tau_k) \right\} \right\} + \sigma_\eta^2 \mathbf{I}$$

But the quantity inside the inner curly brackets is always equal to 1; hence we obtain

$$\mathbf{c}_f = \left\{ \sum_{k=0}^{Q} p_k^2 \mathbf{a}(\theta_k)\mathbf{a}(\theta_k) \right\} + \sigma_\eta^2 \mathbf{I}$$

which may be expressed in a standard form (4.12b) as

$$
\begin{aligned}
\mathbf{c}_f &= \left[\mathbf{a}(\theta_0), \mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_{Q-1}) \right] diag\left\{ p_0^2, p_1^2, \ldots, p_{Q-1}^2 \right\} \\
&\quad \left[\mathbf{a}(\theta_0), \mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_{Q-1}) \right]^H + \sigma_\eta^2 \mathbf{I} \\
&= \mathbf{A}\, diag\left\{ p_0^2, p_1^2, \ldots, p_{Q-1}^2 \right\} \mathbf{A}^H + \sigma_\eta^2 \mathbf{I}
\end{aligned}
\tag{5.60d}
$$

<u>Postcorrelation Covariance Matrix</u>: The outputs of the antenna array, after removing the carrier, are correlated with the desired user code assuming synchronization has been achieved. The postcorrelation array signal vector corresponding to the lth bit is given by

$$\mathbf{g}_0(l) = \frac{1}{T_s} \int_{lT_s}^{(l+1)T_s} \mathbf{f}(t) c_0(t - lT_s)\, dt \tag{5.61}$$

where $\mathbf{f}(t)$ stands for array signal vector without carrier. Using the signal model given by (5.60a), eq. (5.61) may be written as

$$\mathbf{g}_0(l) =$$

$$\frac{1}{T_s} p_0 \mathbf{a}(\theta_0) \int_{lT_s}^{(l+1)T_s} \left[\sum_{j=-\infty}^{\infty} b_{0j} c_0(t - jT_s) \right] c_0(t - lT_s)\, dt$$

$$+ \frac{1}{T_s} \sum_{k=1}^{Q-1} p_k \mathbf{a}(\theta_k) \int_{lT_s}^{(l+1)T_s} \left[\sum_{-\infty}^{\infty} b_{kj} c_k(t - jT_s + \tau_k) \right] c_0(t - lT_s)\, dt \tag{5.62}$$

$$+ \frac{1}{T_s} \int_{lT_s}^{(l+1)T_s} \eta(t) c_0(t - lT_s)\, dt$$

where we have set $\tau_0 = 0$. In (5.62) the first term, equal to b_{0k}, is the signal term. The second term represents the interference. It is evaluated as follows: The integral is split into two parts, lT_s to $lT_s + \tau_k$, and $lT_s + \tau_k$ to $(l+1)T_s$. After a change of variable it reduces to

$$\frac{1}{T_s} \sum_{k=1}^{Q-1} p_k \mathbf{a}(\theta_k) \left[b_{kl-1} \int_0^{\tau_k} c_k(t - T_s - \tau_k) c_0(t) dt + b_{kl} \int_0^{\tau_k} c_k(t - \tau_k) c_0(t) dt \right]$$

$$= \mathbf{b}_l^1 + \mathbf{b}_l^2 \tag{5.63}$$

Finally, the noise term reduces to

$$\eta_l = \frac{1}{T_s} \int_0^{T_s} \eta(t + lT_s) c_0(t) dt$$

The postcorrelation covariance matrix of the array signal is given by

$$\mathbf{C}_{g_0 g_0} = E\{\mathbf{g}_0(l) \mathbf{g}_0^H(l)\}$$
$$= p_0^2 \mathbf{a}(\theta_0) \mathbf{a}^H(\theta_0) + E\{\mathbf{b}_l^1 \mathbf{b}_l^{1H}\} + E\{\mathbf{b}_l^2 \mathbf{b}_l^{2H}\} + E\{\eta_l \eta_l^H\} \tag{5.64a}$$

All cross terms vanish. First, we shall evaluate $E\{\mathbf{b}_l^1 \mathbf{b}_l^{1H}\}$.

$$E\{\mathbf{b}_l^1 \mathbf{b}_l^{1H}\} = \frac{1}{T_s^2} \sum_{k=1}^{Q} p_k^2 \mathbf{a}(\theta_k) \mathbf{a}(\theta_k)^H \rho_1 \tag{5.64b}$$

where

$$\rho_1 = E\left\{ \left[\int_0^{\tau_k} c_k(t + T_s - \tau_k) c_0(t) dt \right]^2 \right\}$$

and the sources are assumed to be uncorrelated. Using the representation of BPSK signal (5.59a) we obtain

$$\rho_1 =$$

$$E\left\{ \left[\int_0^{\tau_k} \sum_{m_0}^{L-1} \sum_{m_k}^{L-1} c_{0,m_0} c_{k,m_k} h_{T_c}(t - m_0 T_c) h_{T_c}(t + T_s - m_k T_c - \tau_k) dt \right]^2 \right\}$$

which, on account of the fact that c_{0,m_0} and c_{k,m_k} are uncorrelated random variables, reduces to

$$\rho_1 = \frac{1}{T_s} \int_0^{T_s} \sum_{m_0}^{L-1} \sum_{m_k}^{L-1} \left[\int_0^{\tau_k} \frac{h_{T_c}(t - m_0 T_c)}{h_{T_c}(t + T_s - m_k T_c - \tau_k)dt} \right]^2 d\tau_k \qquad (5.65a)$$

where τ_k, the arrival time of the signal from the kth user, is assumed to be uniformly distributed over 0 to T_s, the symbol duration. The integral over τ_k may be expressed as a sum over L integrals, one for each chip. When h_{T_c} is a rectangular function the integral with respect to τ_k in (5.65a) can be evaluated in a closed form. After some algebraic manipulations we obtain

$$\rho_1 = \frac{L^2}{3T_s} T_c^3 \qquad (5.65b)$$

Substituting (5.65b) in (5.64b) we obtain

$$E\{\mathbf{b}_l^1 \mathbf{b}_l^{1H}\} = \frac{1}{3L} \sum_{k=1}^{Q} p_k^2 \mathbf{a}(\theta_k) \mathbf{a}(\theta_k)^H \qquad (5.66a)$$

Evaluation of $E\{\mathbf{b}_l^2 \mathbf{b}_l^{2H}\}$ proceeds on the same lines as above. In fact the result is identical, that is,

$$E\{\mathbf{b}_l^2 \mathbf{b}_l^{2H}\} = \frac{1}{3L} \sum_{k=1}^{Q} p_k^2 \mathbf{a}(\theta_k) \mathbf{a}(\theta_k)^H = E\{\mathbf{b}_l^1 \mathbf{b}_l^{1H}\} \qquad (5.66b)$$

Finally, we shall evaluate the noise term in (5.62).

$$E\{\eta_l \eta_l^H\} = \frac{1}{T_s^2} E\left\{ \int_0^{T_s} \int_0^{T_s} \eta(t_1 + lT_s) \eta^H(t_2 + lT_s) c_0(t_1) c_0(t_2) dt_1 dt_2 \right\}$$

$$= \frac{1}{T_s^2} \sum_{m_1=0}^{L-1} \int_0^{T_c} \int_0^{T_c} \frac{E\{\eta(t_1 + lT_c) \eta^H(t_2 + lT_c)\}}{h_{T_c}(t_1 - m_1 T_c) h_{T_c}(t_2 - m_1 T_c) dt_1 dt_2}$$

$$= \frac{\sigma_\eta^2}{L} \mathbf{I} \qquad (5.67)$$

where we assumed that the noise is spatially uncorrelated but temporally correlated over the chip width T_c. Using equations (5.65b), (5.66) and (5.67) in (5.64) the postcorrelation covariance matrix may be expressed as

$$\mathbf{c}_{g_0 g_0} = p_0^2 \mathbf{a}(\theta_0) \mathbf{a}^H(\theta_0) + \frac{2}{3L} \sum_{k=1}^{Q} p_k^2 \mathbf{a}(\theta_k) \mathbf{a}^H(\theta_k) + \frac{\sigma_\eta^2}{L} \mathbf{I} \qquad (5.68)$$

which may be expressed in a standard form (4.12b)

$$\mathbf{c}_{g_0 g_0} = \left[\mathbf{a}(\theta_0), \mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_{Q-1}) \right] diag\left\{ p_0^2, \frac{2}{3L} p_1^2, \ldots, \frac{2}{3L} p_{Q-1}^2 \right\}$$

$$\left[\mathbf{a}(\theta_0), \mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_{Q-1}) \right]^H + \frac{\sigma_\eta^2}{L} \mathbf{I}$$

$$= \mathbf{A} diag\left\{ p_0^2, \frac{2}{3L} p_1^2, \ldots, \frac{2}{3L} p_{Q-1}^2 \right\} \mathbf{A}^H + \frac{\sigma_\eta^2}{L} \mathbf{I} \qquad (5.69)$$

The directions of arrival (DOA) may be estimated following the subspace approach described earlier in §5.1 and §5.2.

5.3.3 Cyclostationary Signals: A signal defined in (5.58) and other similar communication signals possess an important property, namely, its temporal covariance function is periodic. This is a result of transmission of symbols at a fixed rate and also due to the use of a periodic carrier. Indeed the periodicity in the covariance function of the signal, also known as cyclic frequency α, is equal to $l f_b$ where l is an integer and f_b is baud rate, that is, the number of symbols transmitted per second. The baud rate is a unique property associated with each source and it is not affected by propagation. Since the baud rate of a signal of interest (SOI) is known a priori and it is different from that of the other interfering signals, it is possible to distinguish the SOI from the interference and the system noise whose covariance function is known to be aperiodic. In this section we shall describe how a subspace method, i.e., MUSIC, may be devised by exploiting the property of cyclostationarity. The background information on cyclostationary process will not be covered here as such material is already available in a book [35] and in a popular exposition [36]. However, we shall introduce enough material that is essential for the understanding of its application to DOA estimation.

$$\text{Let } \mathbf{f}(t) = \left\{ f_0(t), f_0(t - \frac{d}{c} \sin\theta_0), \ldots, f_0(t - (M-1)\frac{d}{c} \sin\theta_0) \right\}$$

be the output of a M sensor array on which a cyclostationary signal $f_0(t)$ is incident with an angle of incidence θ_0. We define frequency shifted versions of $\mathbf{f}(t)$

$$\mathbf{f}_+(t) = \mathbf{f}(t)e^{j\pi\alpha t}$$
$$\mathbf{f}_-(t) = \mathbf{f}(t)e^{-j\pi\alpha t}$$

(5.70)

and cyclic covariance matrix

$$\mathbf{c}_f^\alpha(\tau) = \frac{1}{T} \sum_{t=-\frac{T}{2}}^{\frac{T}{2}} E\left\{ \mathbf{f}_-(t+\frac{\tau}{2})\mathbf{f}_+^H(t-\frac{\tau}{2}) \right\}$$

$$= \frac{1}{T} \sum_{t=-\frac{T}{2}}^{\frac{T}{2}} E\left\{ \mathbf{f}(t+\frac{\tau}{2})\mathbf{f}^H(t-\frac{\tau}{2}) \right\}e^{-j2\pi\alpha t}$$

$$T \to \infty$$

Let us show how to evaluate the $(k,l)^{\text{th}}$ element of the matrix $\mathbf{c}_f^\alpha(\tau)$.

$$\left[\mathbf{c}_f^\alpha(\tau)\right]_{kl}$$

$$= \frac{1}{T} \sum_{t=-\frac{T}{2}}^{\frac{T}{2}} E\left\{ f_0^*(t+k\frac{d}{c}\sin\theta_0 - \frac{\tau}{2})f_0(t+l\frac{d}{c}\sin\theta_0 + \frac{\tau}{2})e^{-j2\pi\alpha t} \right\}$$

$$T \to \infty$$

$$= c_{f_0}^\alpha(\tau+(k-l)\frac{d}{c}\sin\theta_0)e^{j\pi\alpha(k+l)\frac{d}{c}\sin\theta_0}$$

(5.71a)

When a source is narrowband with center frequency ω_c so that following approximation holds good, $f_0(t+\tau) \approx f_0(t)e^{j\omega_c\tau}$ (see (4.15b)), the $(k,l)^{\text{th}}$ element of the matrix $\mathbf{c}_f^\alpha(\tau)$ may be approximated as $\left[\mathbf{c}_f^\alpha(\tau)\right]_{kl} \approx c_{f_0}^\alpha(\tau)e^{j\omega_c(k-l)\frac{d}{c}\sin\theta_0}e^{j\pi\alpha(k+l)\frac{d}{c}\sin\theta_0}$, which may be further

approximated as $\left[\mathbf{c}_f^\alpha(\tau)\right]_{kl} \approx c_{f_0}^\alpha(\tau)e^{j\omega_c(k-l)\frac{d}{c}\sin\theta_0}$ for $\omega_c \gg 2\pi\alpha$. Using the narrowband approximation and the assumption $\omega_c \gg 2\pi\alpha$ we can express the cyclic covariance matrix as

$$\mathbf{c}_f^\alpha(\tau) \approx \mathbf{a}_0\mathbf{a}_0^H c_{f_0}^\alpha(\tau) \qquad (5.71b)$$

The above relation is quite similar to (4.12b), which was the starting point in the subspace algorithm, for example, in MUSIC. Naturally, based on (5.71b), a subspace algorithm known as Cyclic MUSIC has been proposed in [37]. Although, in deriving (5.71b) we have assumed a single source, it holds good even in the presence of multiple signals with *different* cyclic frequencies and any type of stationary noise.

Let us consider the diagonal terms of the cyclic covariance matrix. From (5.71a) the diagonal terms, k=l, are given by

$$\left[\mathbf{c}_f^\alpha(\tau)\right]_{k=l} = c_{f_0}^\alpha(\tau)e^{j2\pi\alpha l\frac{d}{c}\sin\theta_0} \qquad (5.72a)$$

which we shall express in a matrix form. Let

$$\tilde{\mathbf{c}}_f^\alpha(\tau) = col\left\{\left[\mathbf{c}_f^\alpha(\tau)\right]_{k=l=0},\left[\mathbf{c}_f^\alpha(\tau)\right]_{k=l=1},\ldots,\left[\mathbf{c}_f^\alpha(\tau)\right]_{k=l=M-1}\right\}$$

$$\mathbf{a}(\alpha,\theta_0) = col\left\{1,e^{j2\pi\alpha\frac{d}{c}\sin\theta_0},\ldots,e^{j2\pi\alpha(M-1)\frac{d}{c}\sin\theta_0}\right\}$$

and using these vectors, (5.72a) may be written as

$$\tilde{\mathbf{c}}_f^\alpha(\tau) = c_{f_0}^\alpha(\tau)\mathbf{a}(\alpha,\theta_0) \qquad (5.72b)$$

and for P uncorrelated sources, but with the same cyclic frequency, we obtain

$$\tilde{\mathbf{c}}_f^\alpha(\tau) = \left\{\mathbf{a}(\alpha,\theta_0),\mathbf{a}(\alpha,\theta_1),\ldots,\mathbf{a}(\alpha,\theta_{P-1})\right\}\begin{bmatrix} c_{f_0}^\alpha(\tau) \\ \ldots \\ c_{f_{P-1}}^\alpha(\tau) \end{bmatrix} \qquad (5.72c)$$

Note that in deriving (5.72) we have not used the narrowband approximation which was earlier used in deriving (5.71b). The computer simulation results reported in [38] showed improved results obtained by using (5.72c) over (5.71b)

except when the bandwidth is only a tiny fraction (<0.05) of the carrier frequency. In (5.72b) or (5.72c) the noise term is absent as it is not a cyclostationary process. This is a very important feature of any system using the cyclostationary signals. Whether the noise is correlated or white its cyclostationary autocorrelation function at any cyclic frequency vanishes. Another important feature of great signficance is that the condition of two sources being mutually cyclically uncorrelated is weaker than the condition of mutually uncorrelated. Two cyclostationary signals f(t) and g(t) are said to be uncorrelated if their cross-cyclic covariance function defined as

$$c_{fg}^{\alpha}(\tau) = \frac{1}{T} \sum_{t=-\frac{T}{2}}^{\frac{T}{2}} E\left\{ f^*(t - \frac{\tau}{2}) g(t + \frac{\tau}{2}) \right\} e^{-j2\pi\alpha t}$$

$$T \to \infty$$

(5.73)

vanishes for all τ and α is equal to the cyclic frequency either of f(t) or g(t). To demonstrate this property let us consider $f(t) = s(t)\cos(\omega_1 t + \theta_1)$ and $g(t) = s(t)\cos(\omega_2 t + \theta_2)$ and substituting in (5.73) we obtain

$$c_{fg}^{\alpha}(\tau)=$$

$$c_s(\tau)\frac{1}{2T} \sum_{t=-\frac{T}{2}}^{\frac{T}{2}} \left[\begin{array}{l} \cos((\omega_1 + \omega_2)t - (\omega_1 - \omega_2)\frac{\tau}{2} + \theta_1 + \theta_2) \\ +\cos((\omega_1 - \omega_2)t - (\omega_1 + \omega_2)\frac{\tau}{2} + \theta_1 - \theta_2) \end{array} \right] e^{-j2\pi\alpha t}$$

$$T \to \infty$$

which becomes zero except when $2\pi\alpha = \pm(\omega_1 + \omega_2)$ or $= \pm(\omega_1 - \omega_2)$. Thus, two sources transmitting the same message but with different carriers become cyclically uncorrelated. Generalizing the above result, P sources transmitting even the same message but with different carrier frequencies may become *cyclically uncorrelated*. Further, we notice that even when the carrier frequencies are the same, the sources can become cyclically uncorrelated unless we choose $2\pi\alpha = \pm 2\omega_1$ or 0. As a result, it is possible through an appropriate choice of α to selectively cancel all signals of no interest. As a numerical example, we computed the cyclic cross-correlation function between two sources radiating the same stochastic signal with different carrier frequencies. In the cyclic cross-correlation function is shown as a function of the cyclic frequency. The carrier frequencies are 0.15 Hz and 0.20 Hz. (The Nyquist frequency is 0.5 Hz and 1024 time samples were used.)
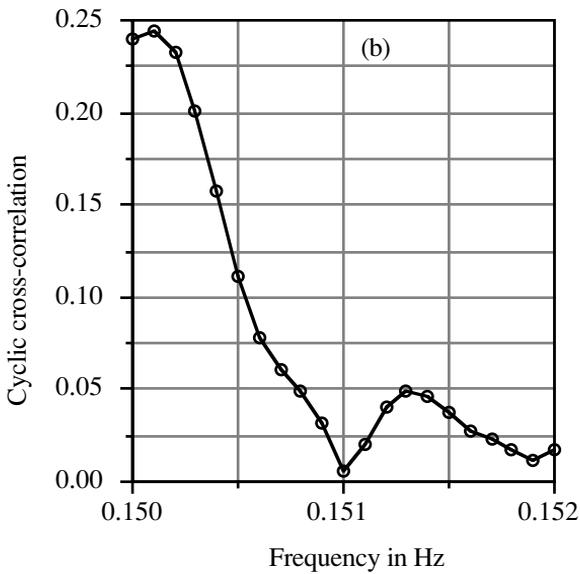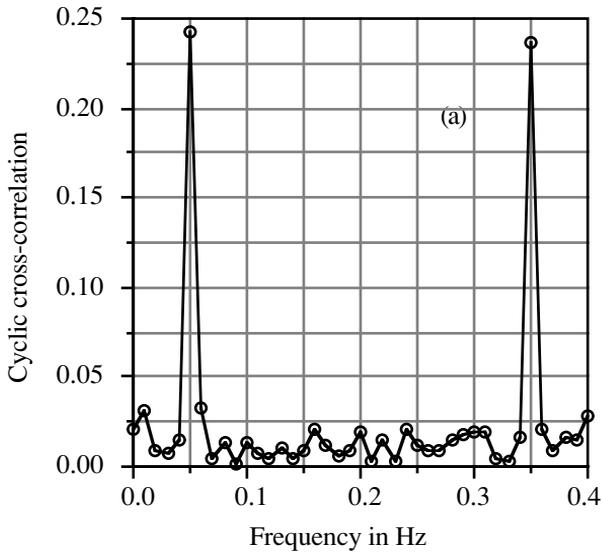
Figure 5.9: Cyclic cross-correlation function at zero lag as a function of (a) cyclic frequency and (b) carrier frequency in Hz.

Note the peaks at the sum and difference frequencies. In the cyclic cross-correlation function is shown as a function of the carrier frequency of the second source.

§5.4 **Array Calibration**:
Throughout we have implicitly assumed that all sensors are ideal, that is, they are omnidirectional point sensors with constant response. Further, the shape of the array is fully known. If we are dealing with a ULA we assume that all sensors are on a perfectly straight line and sensors are equispaced. Any deviations from the assumptions can cause considerable loss of performance, particularly in the use of subspace methods [39]. One way out of this limitation is to actually measure the properties of the array, including its shape, and use this information in the design of the processor. Often this requires a source whose location is known. The properties of the array can be measured, a process known as array calibration. We shall describe two types of calibrations. In the first instance we shall give a method of computing the amplitude and the phase variations. Next we shall describe a method for the shape estimation, which is vital in the use of flexible arrays.

5.4.1 Amplitude and Phase Variation of a Sensor:
From (2.17e) the direction vector may be expressed as a product of two components

$$\mathbf{a}(\theta_0) = \alpha(\theta_0)\phi(\theta_0)$$

where

$$\alpha(\theta_0) = diag\{\alpha_0(\theta_0), \alpha_1(\theta_0), \ldots \alpha_{M-1}(\theta_0)\}$$

$$\phi(\theta_0) = col\{1, e^{-j\omega\frac{d}{c}\sin\theta_0}, e^{-j\omega\frac{2d}{c}\sin\theta_0} \ldots e^{-j\omega\frac{(M-1)d}{c}\sin\theta_0}\}$$

The purpose of array calibration is to estimate $\alpha(\theta_0)$ from actual array observations. We preclude the possibility of direct in-situ measurement of sensor sensitivity (hardware calibration). Recall the relation we had derived between the direction vectors and eigenvectors of $\mathbf{v}_s\mathbf{v}_s^H$ (5.6d). Consider a single source whose direction of arrival is known. For single source (5.6d) may be expressed as

$$\alpha(\theta_0)\phi(\theta_0) = \mathbf{v}_s\mathbf{v}_s^H\alpha(\theta_0)\phi(\theta_0) \qquad (5.74)$$

Clearly $\alpha(\theta_0)\phi(\theta_0)$ is an eigenvector of $\mathbf{v}_s\mathbf{v}_s^H$ and the corresponding eigenvalue is 1. Since the direction of arrival of the source is known we can compute $\phi(\theta_0)$ and demodulate the eigenvector. We can thus estimate $\alpha(\theta_0)$

but for a complex constant. Assume that $\alpha(\theta_0)$ is independent of θ_0. We shall express (5.74) in a different form. Note that

$$\alpha\,\phi(\theta_0) = diag\{\alpha_0,\alpha_1 e^{-j\omega\frac{d}{c}\sin\theta_0},...\alpha_{M-1}e^{-j\omega\frac{(M-1)d}{c}\sin\theta_0}\}$$

$$= diag\{1,e^{-j\omega\frac{d}{c}\sin\theta_0},...e^{-j\omega\frac{(M-1)d}{c}\sin\theta_0}\}col\{\alpha_0,...\alpha_{M-1}\}$$

$$= \phi_d(\theta_0)\alpha_c$$

Using the above result in (5.74) we obtain

$$\alpha_c = \phi_d^H(\theta_0)\mathbf{v}_s\mathbf{v}_s^H\phi_d(\theta_0)\alpha_c \qquad (5.75a)$$

When there are P sources with known DOAs, $\theta_p, p = 0,1,...,P-1$ (5.75a) may be expressed as $\alpha_c = \mathbf{Q}\alpha_c$ where

$$\mathbf{Q} = \sum_{p=0}^{P-1}\phi_d^H(\theta_p)\mathbf{v}_s\mathbf{v}_s^H\phi_d(\theta_p) \qquad (5.75b)$$

and α_c is the eigenvector of $\mathbf{Q}$ corresponding to its largest eigenvalue equal to P [40].

5.4.2 <u>Shape Estimation:</u> In any practical towed array system used in sonar or seismic exploration the position of sensors is continuously monitored. Additionally, it is also possible to estimate the shape from the array output. Here from the point of array signal processing our natural interest is in the second approach. Consider a ULA which has been deformed either on account of towing or ocean currents. Assume that we have a known source radiating a narrowband signal and the background noise is spatially uncorrelated. We have shown in (5.4) the eigenvector of the spatial covariance matrix corresponding to the largest eigenvalue is proportional to the direction vector. In fact, the relation is given by $\mathbf{v}_0 = \dfrac{1}{\sqrt{M}}\mathbf{a}_0$. For deformed array the direction vector may be obtained from (2.34c)

$$\mathbf{a}(\theta_0,\varphi_0) = col\left\{1,e^{-j\omega_c\frac{d}{c}[\gamma_1\sin\theta_0\sin\varphi_0+\varepsilon_1\sin\theta_0\cos\varphi_0+\xi_1\sin\theta_0]},...,\atop e^{-j\omega_c\frac{d}{c}[\gamma_{M-1}\sin\theta_0\sin\varphi_0+\varepsilon_{M-1}\sin\theta_0\cos\varphi_0+\xi_{M-1}\sin\theta_0]}\right\} \qquad (5.76)$$

Therefore, we can relate the eigenvector to the direction vector

$$
\mathbf{v}_0 = \frac{1}{\sqrt{M}} col \left\{ 1, e^{-j\omega_c \frac{d}{c}[\gamma_1 \sin\theta_0 \sin\varphi_0 + \varepsilon_1 \sin\theta_0 \cos\varphi_0 + \xi_1 \sin\theta_0]}, \ldots, e^{-j\omega_c \frac{d}{c}[\gamma_{M-1} \sin\theta_0 \sin\varphi_0 + \varepsilon_{M-1} \sin\theta_0 \cos\varphi_0 + \xi_{M-1} \sin\theta_0]} \right\} \tag{5.77a}
$$

Note that since the source is known, its azimuth and elevation are known a priori. The unknowns are $\gamma_m, \varepsilon_m, m = 0,1,\ldots,M-1$. (Note $\xi_m, m = 0,1,\ldots,M-1$ are dependent on γ_m, ε_m.) To solve for a pair of unknowns we would need one more source, say, with different azimuth and elevation, (θ_1, φ_1). Once again the largest eigenvector may be related to the direction vector of the second source.

$$
\mathbf{v}_1 = \frac{1}{\sqrt{M}} col \left\{ 1, e^{-j\omega_c \frac{d}{c}[\gamma_1 \sin\theta_1 \sin\varphi_1 + \varepsilon_1 \sin\theta_1 \cos\varphi_1 + \xi_1 \sin\theta_1]}, \ldots, e^{-j\omega_c \frac{d}{c}[\gamma_{M-1} \sin\theta_1 \sin\varphi_1 + \varepsilon_{M-1} \sin\theta_1 \cos\varphi_1 + \xi_{M-1} \sin\theta_1]} \right\} \tag{5.77b}
$$

We shall, for the sake of simplicity, assume that the deformed array is in the x-y plane and also place the calibrating source in the x-y plane. Then, a single source is enough for estimation of γ_m and ε_m which now take a form

$$
\gamma_m = \sum_{i=0}^{m} \cos\alpha_i, \quad m = 0,1,\ldots,M-1
$$
$$
\varepsilon_m = \sum_{i=0}^{m} \sin\alpha_i, \quad m = 0,1,\ldots,M-1
$$
(5.78)

Using (5.78) in (5.77b) we obtain the following basic result

$$
\frac{\lambda_c}{2\pi}\left[\angle\{\mathbf{v}_0\}_m - \angle\{\mathbf{v}_0\}_{m-1}\right] + \lambda_c n =
$$
$$
d\cos\alpha_m \cos\varphi_0 + d\sin\alpha_m \sin\varphi_0 = \Delta x_m \cos\varphi_0 + \Delta y_m \sin\varphi_0
$$
(5.79)

where n is an unknown integer and Δx_m and Δy_m are x,y coordinates of mth sensor relative to m-1st sensor. The ambiguity, arising out of the unknown integer, may be resolved through geometrical considerations. Note that the mth sensor must be on a circle of radius d centered at
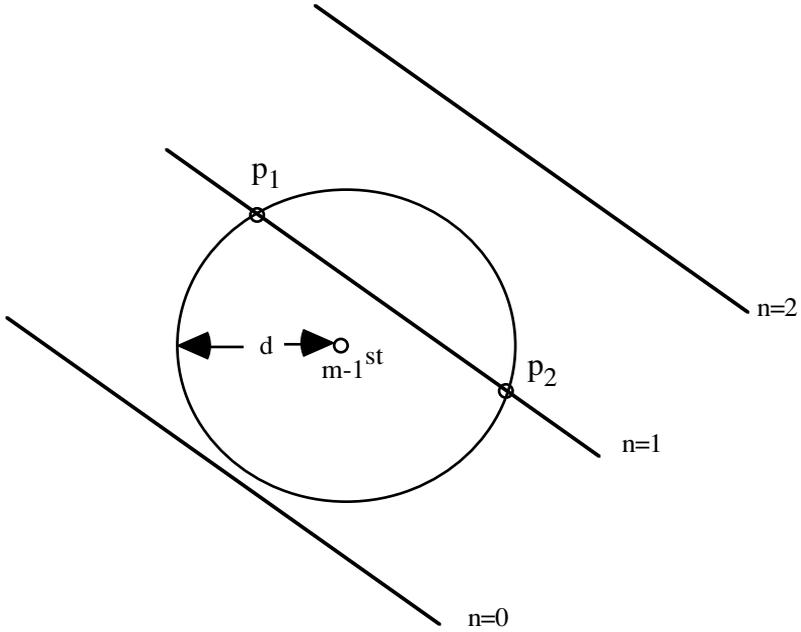
Figure 5.10 The ambiguity in (5.77) is resolved by requiring that the line it represents must intersect the circle of radius d drawn at m-1st sensor. There are two intersection points. The sensor may be at any one of the two intersections.

m-1st sensor. Further the position of the sensor must satisfy (5.79), which, incidentally, is an equation of a straight line [41]. For illustration let the line corresponding to n=1 intersect the circle at two points p_1 and p_2 (see fig. 5.10). The sensor can either be at p_1 or p_2. This ambiguity is resolved by choosing a point which results in minimum array distortion [42].

§5.5 **Source in Bounded Space**:
When a source is in a bounded space a sensor array will receive signals reflected from the boundaries enclosing the space, for example, radar returns from a low flying object [43], an acoustic source in shallow water [44], a speaker in a room [45]. In all these cases the reflections are strongly correlated with the direct signal and come from a set of predictable directions. The complexity of the multipath structure increases with the increasing number of reflecting surfaces, as illustrated in fig. 5.11. We shall in this section briefly consider two simple cases involving one reflecting surface (a low flying object) and two reflecting surfaces (an acoustic source in shallow water) shown in fig. 5.11a & b.

(a)

(b)

(c)
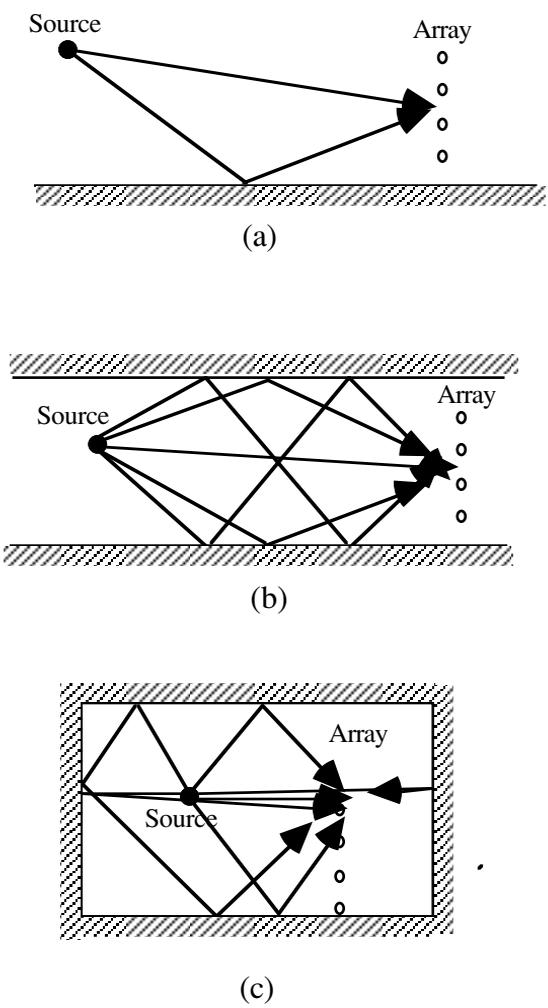
Figure 5.11: Three examples of multipath structure with increasing complexity. Assuming the boundaries are well defined we can predict the multipaths given the source and array locations.

5.5.1 <u>Single Reflecting Surface</u>: Assume that the source emits a stationary stochastic signal. A vertical sensor array is used to receive the signal radiated by the source. The array output in frequency domain may be written in terms of the radiated signal as follows:

$$dF(\omega) = \mathbf{a}_0 dF_0(\omega) + w_1 e^{-j\omega\tau_1}\mathbf{a}_1 dF_0(\omega) + dN(\omega)$$

$$= [\mathbf{a}_0, \mathbf{a}_1][1, w_1 e^{-j\omega\tau_1}]^T dF_0(\omega) + dN(\omega) \tag{5.80}$$

where $dF_0(\omega)$ is the differential of the generalized Fourier transform of stochastic signal emitted by the source, $dN(\omega)$ is the differential of the generalized Fourier transform of the background noise presumed to be spatially and temporally white and $\mathbf{a}_0$, and $\mathbf{a}_1$ are the direction vectors of direct and reflected signals, respectively,

$$\mathbf{a}_0 = col\left\{1, e^{-j\omega\frac{d}{c}\sin\theta_0}, \dots, e^{-j\omega(M-1)\frac{d}{c}\sin\theta_0}\right\}$$

$$\mathbf{a}_1 = col\left\{1, e^{-j\omega\frac{d}{c}\sin\theta_1}, \dots, e^{-j\omega(M-1)\frac{d}{c}\sin\theta_1}\right\}$$

τ_1 is the delay of the reflected signal relative to the direct arrival and w_1 stands for reflection coefficient. The spectral matrix of the array output is easily derived from (5.80)

$$\mathbf{S}_f(\omega) = [\mathbf{a}_0, \mathbf{a}_1][1, w_1 e^{-j\omega\tau_1}]^T [1, w_1 e^{-j\omega\tau_1}][\mathbf{a}_0, \mathbf{a}_1]^H S_0(\omega) + \sigma_\eta^2 \mathbf{I} \tag{5.81}$$

Define $\tilde{\mathbf{a}} = [\mathbf{a}_0, \mathbf{a}_1][1, w_1 e^{-j\omega\tau_1}]^T = \mathbf{a}_0 + w_1 e^{-j\omega\tau_1}\mathbf{a}_1$ and rewrite (5.81) as

$$\mathbf{S}_f(\omega) = \tilde{\mathbf{a}} S_0(\omega)\tilde{\mathbf{a}}^H + \sigma_\eta^2 \mathbf{I} \tag{5.82}$$

Note that (5.82) is of the same form as (4.12b). Naturally, it is possible to derive a subspace algorithm to estimate the parameters, θ_0 and θ_1; alternatively, range and height of the source above the reflecting surface, in terms of which θ_0 and θ_1, can be expressed. The Music spectrum is now a function of two parameters, range and height, instead of frequency as in the conventional Music spectrum. For this reason it may be worthwhile to call it a parametric spectrum. A numerical example of parametric spectrum for a source situated above a reflecting surface and an eight sensor vertical ULA is given in fig. 5.12. The peak of the parametric spectrum is correctly located but this good result has been achieved because we have used the exact value of the reflection coefficients. Even a small error, on the order of 1%, appear to completely alter
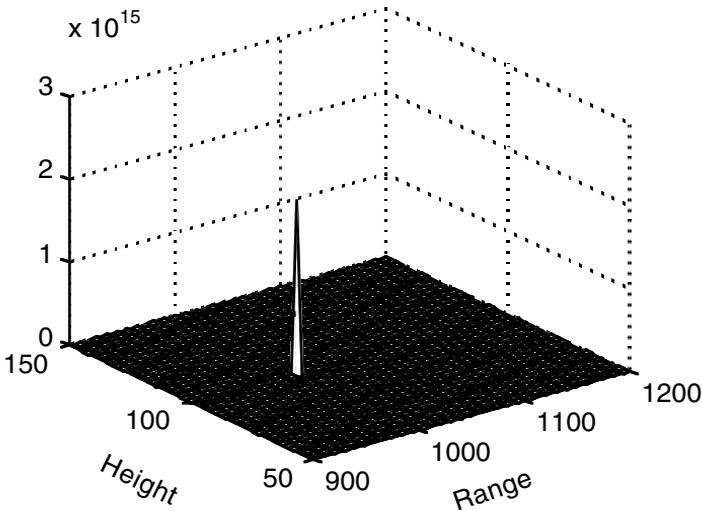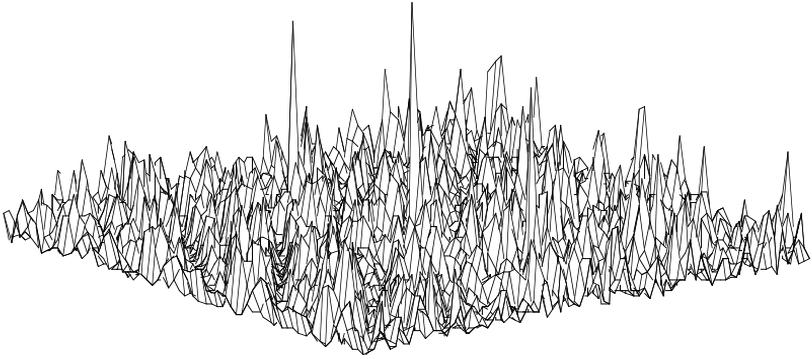
Figure 5.12: Parametric spectrum for a source situated above a reflecting surface ($100\,\lambda$) and $1000\,\lambda$ away from an eight element vertical ULA. The array center is at $16\,\lambda$ above the reflecting surface.

the picture; in particular, the range estimation becomes very difficult.

5.5.2 <u>Two Reflecting Surfaces</u>: To keep the analysis reasonably simple we shall confine ourselves to two parallel reflecting surfaces and a vertical ULA. A broadband signal is assumed, as simple frequency averaging seems to yield good results (see fig. 5.13). Consider a uniform channel of depth H meters where a vertical array of equispaced (spacing $d_0 < \lambda/2$) sensors is located at depth z_R (depth is measured from the surface of the channel to the top sensor of the array) and a source of radiation at a depth z_s and is horizontally separated by a distance R_0 from the array. We have noted in chapter 1 (page 26) that depending upon the range and the channel characteristics the array would be able to see a certain number of significant images, say P. The waveform received from P images at m^{th} sensor (m=0 is the top sensor) is given by

$$f_m(t) = \sum_{p=0}^{P-1} \frac{\alpha_p}{R_p} f_0(t - \tau_{pm}) + \eta_m(t) \qquad (5.83)$$

where $f_0(t)$ is the signal emitted by the source, presumed to be a stationary stochastic process, τ_{pm} is time delay from p^{th} image to m^{th} sensor, R_p is the

(a)
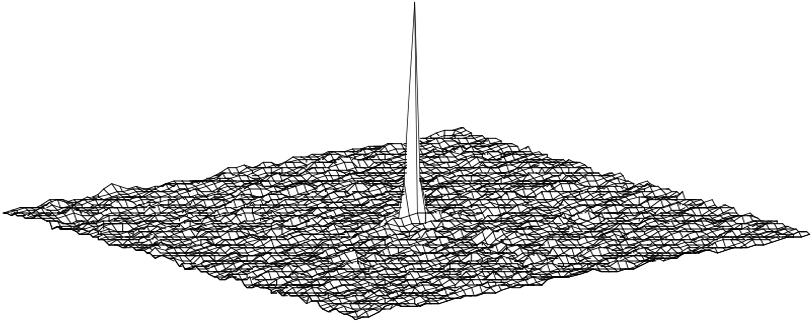
(b)

Figure 5.13: (a) Parametric spectrum at 1600 Hz is shown as a function of range and depth. A broadband acoustic source is assumed at 2000m (range) and 25m (depth) in a channel of depth 100m. Twenty frequency snapshots were used to compute the spectral matrix. Next, the parametric spectra computed at 50 frequencies, equispaced in the band 1000Hz to 2000Hz, were averaged. The averaged spectrum is shown in (b). (From [47]. © 1999, With permission from Elsevier Science)

distance to the pth image and α_p is the reflection coefficient for the ray arriving from the pth image. $\eta_m(t)$ is noise received by the mth sensor. It is easy to show that

$$\tau_{pm} = \tau_p + (m-1)\frac{d}{c}\sin\theta_p \qquad (5.84)$$

where τ_p is travel time from pth image to the top sensor and θ_p is azimuth angle (with respect to the horizontal plane) of the wave vector from pth image and c is sound speed in sea water. This angle can be computed from the geometry of the image structure as shown in fig.1.14. Using the spectral representation of the stationary stochastic process (5.83) may be written in the frequency domain as

$$dF_m(\omega) = \sum_{p=0}^{P-1} w_p dF_0(\omega) e^{j(m-1)2\pi \frac{d}{\lambda} \sin \theta_p} + dN_m(\omega) \qquad (5.85)$$

where $w_p = \dfrac{\alpha_p}{R_p} e^{j\omega\tau_p}$ [44]. Let us express (5.85) in a compact matrix notation by defining

$$\mathbf{A} = \begin{bmatrix} 1 & \cdots & 1 \\ e^{j\omega\frac{d}{c}\sin\theta_0} & \cdots & e^{j\omega\frac{d}{c}\sin\theta_{P-1}} \\ & \cdots & \\ e^{j\omega\frac{d}{c}(M-1)\sin\theta_0} & \cdots & e^{j\omega\frac{d}{c}(M-1)\sin\theta_{P-1}} \end{bmatrix}$$

and $\mathbf{w} = col[w_0, w_1, \ldots, w_{P-1}]$. Equation (5.85) may be written in a matrix form,

$$d\mathbf{F}(\omega) = \mathbf{A}\mathbf{w}dF_0(\omega) + d\mathbf{N}(\omega) \qquad (5.86)$$

The spectral matrix is obtained from (5.86) by squaring and averaging

$$\mathbf{S}_f(\omega) = \mathbf{A}\mathbf{w}\mathbf{w}^H\mathbf{A}^H S_{f_0}(\omega) + \sigma_\eta^2 \mathbf{I} \qquad (5.87)$$

where $S_{f_0}(\omega)$ (scalar) is the spectrum of the source radiation, and σ_η^2 is the variance of the background noise. Define a vector $\tilde{\mathbf{A}} = \mathbf{A}\mathbf{w}$ and rewrite (5.87) as

$$\mathbf{S}_f(\omega) = \tilde{\mathbf{A}} S_{f_0}(\omega)\tilde{\mathbf{A}}^H + \sigma_\eta^2 \mathbf{I} \qquad (5.88)$$

Note that the structure of spectral matrix in (5.88) is the same as in (4.12b). This enables us to develop a subspace algorithm to estimate unknown parameters, range and depth of the source (also azimuth, if a horizontal array is used) [44].

Source localization in a bounded space is prone to errors in the assumed model as well as the estimation errors due to finite available data. The geometry of the channel needs` to be specified with an accuracy of a fraction of wavelength and the wavespeed must be known accurately. In practice these requirements are often difficult to satisfy. However, it is possible, at the cost of increased array length, to achieve good results when the channel is only partially known [46]. Good results have also been obtained with limited data but using a broadband signal as demonstrated in fig. 5.13 [47].

§5.6 **Exercises:**

1. The line represented by (5.79) intersects the circle at two points p_1 and p_2 (see fig. 5.10). Let two adjacent sensors be on the x-axis. Show that one of the points will be at the intersection of the x-axis and circle. Where will be the second point?

2. In chapter 4 (4.13a) it was shown that $\mathbf{A}^H\mathbf{Q} = \mathbf{0}$ where $\mathbf{Q}$ was defined in terms of partitions of the spectral matrix. Is $\mathbf{Q}$ itself the noise subspace? Remember that in obtaining $\mathbf{Q}$ no eigendecomposition was required.

3. Show that, taking into account the variation in the sensitivity of the sensors, equation (5.72b) takes the form

$$\tilde{\mathbf{c}}_f^\alpha(\tau) = c_f^\alpha(\tau)\alpha_2(\theta_0)\mathbf{a}(\alpha,\theta_0)$$

where

$$\alpha_2(\theta_0) = diag\{|\alpha_0(\theta_0)|^2, |\alpha_1(\theta_0)|^2, ..., |\alpha_{M-1}(\theta_0)|^2\}.$$

Interestingly, phase variations themselves do not affect the relation given in (5.71b).

4. How would you cancel the interference coming from the users of no interest in equations (5.60d) and (5.69)? Can you also cancel the term due to the noise?

5. In subsection 5.1.2 we have shown how to restore the rank of a rank deficient spectral matrix by smoothing the spectral matrices of subarrays. Show that this can also be achieved by smoothing of the signal subspace eigenvectors of the spectral matrix of the full array. This approach is suggested in [48] but on a covariance matrix.

6. Let $\mathbf{J}$ be an exchange matrix which collects all odd and even elements of a vector; for example, a 4x4 exchange matrix is

$$\mathbf{J} = \begin{bmatrix} 1,0,0,0 \\ 0,0,1,0 \\ 0,1,0,0 \\ 0,0,0,1 \end{bmatrix}$$

Consider a ULA with M sensors and P uncorrelated wavefronts that are incident on the array. Let $\mathbf{V}_s$ be a matrix whose columns are the signal eigenvectors of the array spectral matrix. Show that $\mathbf{J}\mathbf{v}_s = \begin{bmatrix} \mathbf{A} \\ \mathbf{A}\Gamma \end{bmatrix} \mathbf{G}^{-1}$ where matrices $\mathbf{A}$ ($\frac{M}{2} \times P$) and $\mathbf{G}$ ($P \times P$) are as defined in (5.6c) and the Γ ($P \times P$) is as defined in (2.57). This result provides an alternate approach to the ESPRIT algorithm described in §5.1. It can be used to extend the concept of subspace rotation to multiple subarrays as described in [49].

References

1. R. O. Schmidt, Multiple emitter location and signal parameter estimation, in Proc. of RADC Spectral Estimation Workshop, Griffiths AFB, New York, pp. 243-258, Reprinted in IEEE Trans. AP-34, pp. 276-280, 1986.
2. G. Bienvenu and L. Kopp, Principe de la goniometri passive adaptive, Proc. Colloque, CRESTI, Nice (France), pp. 106/1-10, 1979.
3. V. F. Pisarenko, The retrieval of harmonics from covariance function, Geophy. J. R. Astron. Soc., vol. 33, pp. 347-366, 1973.
4. C. U. Padmini and P. S. Naidu, Circular array and estimation of direction of arrival of a broadband source, Signal Proc., vol. 37, pp. 243-254, 1994.
5. A. B. Gershman and V. T. Ermolaev, Electronics Letters, vol. 28, pp. 1114-1115, 1992.
6. S. U. Pillai, *Array Signal Processing*, Springer-Verlag, New York, 1989.
7. A. B. Gershman and V. T. Ermolaev, Optimal subarray size for spatial smoothing, IEEE SP Letters, vol.2, pp. 28-30, 1995.
8. A. J. Barabell, Improving the resolution performance of eigenstructure based direction finding algorithm, IEEE Proc ICASSP, pp. 336-339, 1983.
9. B. D. Rao, Performance analysis of root Music, IEEE Trans., ASSP-37, pp. 1939-1049, 1989.
10. R. Roy, A. Paulraj, and T. Kailath, Direction of arrival estimation by subspace rotation method - ESPRIT, IEEE. Proc ICASSP, Tokyo, pp. 2495-2498, 1986.
11. D. Slepian and H. D. Pollack, and H. J. Landau, Prolate spheroidal wave function, Fourier analysis and uncertainty, Bell Systems Tech Jour., vol. 40, pp. 43-84, 1961.

12. U. Reddy, B. Egart, and T. Kailath, Least squares type algorithm for adaptive implementation of Pisarenko's harmonic retrieval method, IEEE Trans. ASSP-30, pp. 339-405, 1982.

13. P. Yang and M. Kaveh, Adaptive eigensubspace algorithm for direction or frequency estimation and tracking, IEEE Trans, ASSP-36, pp. 241-251, 1988.

14. J. Karhunen, Adaptive algorithms for estimation of eigenvectors of correlation type matrices, IEEE, Proc. ICASSP-84, pp. 14.6.1-4, 1984.

15. G. Stewart, *Introduction to Matrix Computations*, Academic Press, New York, 1973.

16. H. Wang and M. Kaveh, Coherent signal subspace processing for the detection and estimation of multiple wideband sources, IEEE, ASSP-30, pp. 823-831, 1985.

17. H. Wang and M Kaveh, Focussing matrices for coherent signal subspace processing, IEEE ASSP-36, pp. 1272-1281, 1988.

18. J. Krolick and D. N. Swinger, Focussed wideband array processing via spatial sampling, IEEE, ASSP-38, pp. 356-360, 1990.

19. K. M. Buckly and L. Griffiths, Broadband signal subspace spatial spectrum (BASS-ALE) estimation, IEEE ASSP-36, pp. 953-964, 1988.

20. D. N. Swinger and J. Krolik, Source location bias in the coherently focused high resolution broadband beamformation, IEEE ASSP-37, pp. 143-144, 1989.

21. J. Krolik and D. Swingler, Multiple broadband source location using steered covariance matrices, IEEE Trans. ASSP-37, pp. 1481-1494, 1989.

22. S. Valae and P. Kabal, Wideband array processing using a two sided correlation transformation, IEEE Trans SP-43, pp. 160-172, 1995.

23. K. M. Buckley and L. Griffiths, Eigenstructure based broadband source localization, IEEE, ICASSP-86, pp. 1867-1872, 1986.

24. C. Usha Padmini, Wideband source localization, ME dissertation, ECE Dept, IISc, Bangalore, India, 1991.

25. P. S. Naidu, *Modern Spectrum Analysis of Time Series*, CRC Press, Boca Raton,FL.,1996.26. J. W. Mauchley, Significance test for sphericity of a normal n-variate distribution, Ann. Math. Stat., vol.11. pp. 204-209, 1940.

27. W. S. Ligget, Jr., Passive sonar: fitting models to multiple time series, in Signal Processing: Proc. NATO ASI Signal Processing, J. W. R. Griffiths and P.L. Stocklin, Eds., Academic, New York, pp. 327-345, 1973.

28. D. B. Williams and D. H. Johnson, Using the sphericity test for source detection with narrow band passive array, IEEE Trans., ASSP-38, pp. 2008-2014, 1990.

29. M. Wax and T. Kailath, Detection of signals by information theoretic criteria, IEEE Trans., ASSP-33, pp. 387-392, 1985.

30. Qi-Tu Zhang, Kon Max Wong, P. C. Yip, and J. P. Reilly, Statistical analysis of the performance of information theoretic criteria in the detection of the number of signals in array processing, IEEE Trans., ASSP-39, pp. 1557-1567,1989.

31. J. P. Henrioux and G. Jourdain, Use of large bandwidth-duration binary phase shift keying signals for bearing measurements in active sonar classification, JASA, vol. 97, pp. 1737-1746, 1995.

32. R. L. Pickholz, D. L. Schilling, and L. B. Milstein, Theory of spread spectrum communications-A Tutorial, IEEE Trans COM-30, pp. 855-884, 1982.

33. R. L. Pickholz, L. B. Milstein, and D. L. Schilling, Spread spectrum for mobile communications, IEEE Trans.Vehicular Technology, vol. 40, pp. 313-322, 1991.

34. A. F. Naguib, A. Paulraj, and T. Kailath, Capacity improvement with base station antenna array in cellular CDMA, IEEE Trans.,vol. 43, pp. 691-698, 1994.

35. W. A. Gardner, *Introduction to Random Processes with Applicarion to Signals and Systems*, New York, MacMillan, 1985.

36. W. A Gardner, Exploitation of spectral redundancy in cyclostationarity signals, IEEE., Signal Proc. Mag., pp. 14-36, April 1991.

37. S. V. Schell, R. A. Calabretta, W. A. Gardner and B. G. Agee, Cyclic MUSIC algorithm for signal selective direction finding, Proc. ICASSP89, vol. 4, pp. 2278-2281, 1989.

38. G. Xu and T. Kailath, Direction of arrival estimation via exploitation of cyclostationarity- A combination of temporal and spatial processing, IEEE Trans., SP-40, pp. 1775-1786, 1992.

39.R. T. Compton, Pointing accuracy and dynamic range in steered beam adaptive array, IEEE Trans., AES-16, pp. 280-287, 1980.

40. V. C. Soon, L. Tong, Y. F. Huang, and R. Liu, A subspace method for estimating sensor gains and phases, IEEE Trans., SP-42, pp. 973-976, 1994.

41. J. J. Smith, Y. H. Leung, and A. Cantoni, The partitioned eigenvector method for towed array shape estimation, IEEE Trans SP-44, pp. 2273-2283, 1996.

42. D. A. Gray, W. O. Wolfe and J. L. Riley, An eigenvector method for estimating the positions of the elements of an array of receivers, Proc. ASSPA 89, Signal Processing theories, implementation, and applications, R. F. Barrett (Ed.), Adelaide Australia, pp. 391-393, 1989.

43. S. B. Kesler, J. Kesler and G. Levita, Resolution of coherent target in the near field of an array antenna, IEEE, Proc. ICASSP-86, pp. 1937-1940, 1986.

44. P. S. Naidu, A new method of acoustic source localization in shallow sea. Proc. Institute of Acoustics, vol. 18, part 5, pp. 39-48, 1996.

45. H. F. Silverman, Some analysis of microphone arrays for speech data acquisition, IEEE Trans., ASSP-35, pp. 1699-1712, 1987.

46. P. S. Naidu and T. Ganesan, Source localization in a partially known shallow water channel, Journ Acoust Soc of Am, vol. 98, pp. 2554-2559, 1995.

47. P. S. Naidu and H. Uday Shankar, Broadband source localization in shallow water, Signal Processing, vol.72, pp. 107-116, 1999.

48. S. S. Reddi and A.B. Gershman, An alternative approach to coherent source location proble, IEEE Trans. SP-59, pp. 221-233, 1997.

49. A. L. Swindlehurst, B. Ottersten, R. Roy, and T. Kailath, Multiple invariance ESPRIT, IEEE Trans. SP-40, pp. 867-881, 1992.

50. A. Papoulis, *Probability, Random Variables and Stochastic Processes*, Third edition, McGraw-Hill Book Co., New York, 1991.

Source Estimation

The temporal signals radiated by sources are separated on the principle of nonoverlapping or partially overlapping temporal spectral characteristics of the signals. A perfect separation is possible only when the signal spectra are nonoverlapping. The spatio-temporal signals possess an additional degree of variability, namely, the spatial spectrum. The differences in the spatial spectra can be used, in addition to the differences in the temporal spectra, for the purpose of signal separation. The signals coming from widely different directions will have nonoverlapping spatial spectra and therefore they can be separated using an array of sensors. However, when the signal sources are quite close, perfect separation is not possible. There will be some cross talk. We shall evaluate the Wiener filter, which was derived in chapter 3, from the point of cross talk power in relation to the total signal power. Suppression of unwanted signal or interference is achieved by placing a null or a collection of nulls in the spatial frequency band occupied by the interference. The effectiveness of nulls is enhanced when additional constraints are placed on the filter; for example, the filter response is unity in the direction of useful signal. This leads to the well known Capon's filter which is also known as the minimum variance filter. It is found that the Capon's filter is quite effective when the signal and the interference sources are highly directional. The filter will automatically place a null wherever there is a strong interference. Finally, when the direction of interference is known a priori, it is possible to devise a filter which will place a sharp null at the spatial frequency corresponding to the direction of arrival of the interference. The null can be steered to any desired position, depending upon how the interference is changing its direction. Thus, null steering can be effectively used to suppress slowly varying interference.

§6.1 **Wiener Filters**:
Previously, in chapter 3 we have derived Wiener filters for extraction of a wavefront incident at M sensor ULA and UCA. It may be recalled that the Wiener filter minimizes the mean square error between the filter output and the desired signal. The signal and noise are assumed to be stationary stochastic processes. The basic equation (3.38) is reproduced here for convenience,

$$\mathbf{S}_f(\omega)\mathbf{H}(\omega) = \mathbf{S}_0(\omega) \qquad\qquad (3.38a)$$

where $\mathbf{S}_f(\omega)$ is the spectral matrix of the array output,

$$\mathbf{S}_f(\omega) = \begin{bmatrix} S_{0,0}(\omega) & S_{0,1}(\omega) & \dots & S_{0,M-1}(\omega) \\ S_{1,0}(\omega) & S_{1,1}(\omega) & \dots & S_{1,M-1}(\omega) \\ & & \dots & \\ & & \dots & \\ & & \dots & \\ S_{M-1,0}(\omega) & S_{M-1,1}(\omega) & \dots & S_{M-1,M-1}(\omega) \end{bmatrix}$$

and $\mathbf{S}_0(\omega)$ is the cross-spectral vector between the desired output and the array output. We had derived a specific result for a single wavefront with white background noise. It was shown that the filter function is given by

$$\mathbf{H}_W(\omega,\theta_0) = Q\,\mathbf{a}_0(\omega,\theta_0) \qquad\qquad (6.1a)$$

where
$$Q = \dfrac{\dfrac{S_0(\omega)}{\sigma_\eta^2}}{1 + M\dfrac{S_0(\omega)}{\sigma_\eta^2}}$$

$\mathbf{a}_0(\omega,\theta_0)$ is a steering vector and $S_0(\omega)$ is the spectrum of the incident signal.

6.1.1 <u>Filter Output</u>: We shall now use this filter on the Fourier transformed array output, $\hat{\mathbf{F}}(\omega)$

$$\begin{aligned} \hat{\mathbf{F}}(\omega) &= \mathbf{H}_W^H(\omega,\theta_0)\mathbf{F}(\omega) \\ &= Q\mathbf{a}_0^H(\omega,\theta_0)\big[\mathbf{a}_0(\omega,\theta_0)\Xi_0(\omega) + \mathbf{N}(\omega)\big] \qquad (6.1b) \\ &= \Xi_0(\omega) + Q\mathbf{a}_0^H(\omega,\theta_0)\mathbf{N}(\omega) \end{aligned}$$

Note that the signal component in the output remains undisturbed and the noise variance is reduced by a factor of $\dfrac{1}{M}$. The filter response to an incident plane wave coming from a different direction with a direction vector $\mathbf{a}(\omega,\theta)$ is given by $Q\mathbf{a}_0^H(\omega,\theta_0)\mathbf{a}(\omega,\theta)$. The response of the Wiener filter will be compared later with that of the Capon filter in fig. 6.4.

<u>Array Gain</u>: Array gain is defined as a ratio of the output signal-to-noise ratio (snr) to the input signal-to-noise ratio. For the purpose of illustration let us

consider a single source in presence of white noise. The spectral matrix of the array output (4.12b) which we shall write in a slightly different form showing the signal power explicitly

$$\mathbf{S}_f(\omega) = \sigma_s^2 \mathbf{a}_0(\omega)\tilde{S}_0(\omega)\mathbf{a}_0^{H}(\omega) + \sigma_\eta^2 \mathbf{I}$$

where $\tilde{S}_0(\omega)$ is a normalized spectrum. The input signal-to-noise ratio is evidently equal to $\dfrac{\sigma_s^2}{\sigma_\eta^2}$. We assume that the Wiener filter (6.1a) is used to extract the signal whose DOA is known a priori. The output power is given by

$$\mathbf{H}_W^{H}(\omega,\theta_0)\mathbf{S}_f(\omega)\mathbf{H}_W(\omega,\theta_0) =$$

$$\sigma_s^2 Q^2 |\mathbf{a}_0(\omega)|^2 \tilde{S}_0(\omega)|\mathbf{a}_0(\omega)|^2 + \sigma_\eta^2 Q^2 |\mathbf{a}_0(\omega)|^2$$

It is easy to show the output snr as $\dfrac{\sigma_s^2}{\sigma_\eta^2}|\mathbf{a}_0(\omega)|^2$. By definition, the array gain is equal to $|\mathbf{a}_0(\omega)|^2$ which, for ideal sensor array with omnidirectional and unit response sensors, is equal to M (the number of sensors).

6.1.2 <u>Two Source Case</u>: We shall now consider two wavefronts incident on an array (ULA) in presence of white background noise. The sources are assumed to be uncorrelated between themselves as well as with the background noise. The aim is to extract a first source signal while we suppress the signal from the second source as well as the background noise. Corresponding to this model the spectral matrix and the cross-spectral vector in the Wiener filter equation take the following form,

$$\mathbf{S}(\omega) \;=\; \begin{cases} \mathbf{a}_0(\omega,\theta_0)\mathbf{a}_0^{H}(\omega,\theta_0)S_0(\omega) \\ +\mathbf{a}_1(\omega,\theta_1)\mathbf{a}_1^{H}(\omega,\theta_1)S_1(\omega) + \sigma_\eta^2 \mathbf{I} \end{cases} \qquad (6.2)$$

$$\mathbf{S}_0(\omega) = \mathbf{a}_0(\omega,\theta_0)S_0(\omega)$$

where

$$\mathbf{a}_0(\omega,\theta_0) = col[1,\, e^{-j\frac{\omega}{c}d\sin\theta_0},\, e^{-j\frac{\omega}{c}2d\sin\theta_0},\, \dots e^{-j\frac{\omega}{c}(M-1)d\sin\theta_0}\,]$$

and

$$\mathbf{a}_1(\omega,\theta_1) = col[1, e^{-j\frac{\omega}{c}d\sin\theta_1}, e^{-j\frac{\omega}{c}2d\sin\theta_1}, \dots e^{-j\frac{\omega}{c}(M-1)d\sin\theta_1}]$$

where θ_0 and θ_1 are DOA angles of the first and the second wavefront, respectively. The filter response for extraction of the first source signal is given by

$$\mathbf{H}_W(\omega,\theta_0) = \begin{bmatrix} \mathbf{a}_0(\omega,\theta_0)\mathbf{a}_0^H(\omega,\theta_0) + \\ \mathbf{a}_1(\omega,\theta_1)\mathbf{a}_1^H(\omega,\theta_1)\dfrac{S_1(\omega)}{S_0(\omega)} + \dfrac{\sigma_\eta^2}{S_0(\omega)}\mathbf{I} \end{bmatrix}^{-1} \mathbf{a}_0(\omega,\theta_0) \quad (6.3)$$

The inverse of the quantity inside the square brackets in (6.3) has been derived in chapter 4 (p. 234). Using that result we obtain

$$\mathbf{H}_W(\omega,\theta_0) = \begin{bmatrix} \mathbf{V}^{-1} - \dfrac{S_1(\omega)}{S_0(\omega)} \times \\ \dfrac{\mathbf{V}^{-1}\mathbf{a}_1(\omega,\theta_1)\mathbf{a}_1^H(\omega,\theta_1))\mathbf{V}^{-1}}{1 + \dfrac{S_1(\omega)}{S_0(\omega)}\mathbf{a}_1^H(\omega,\theta_1)\mathbf{V}^{-1}\mathbf{a}_1(\omega)} \end{bmatrix} \mathbf{a}_0(\omega,\theta_0) \quad (6.4a)$$

where

$$\mathbf{V}^{-1} = \dfrac{S_0(\omega)}{\sigma_\eta^2}\left[\mathbf{I} - Q\mathbf{a}_0(\omega,\theta_0)\mathbf{a}_0^H(\omega,\theta_0)\right] \quad (6.4b)$$

where

$$Q = \dfrac{\dfrac{S_0(\omega)}{\sigma_\eta^2}}{1 + M\dfrac{S_0(\omega)}{\sigma_\eta^2}}$$

Though the filter is tuned to receive the first wavefront some amount of energy from the second wavefront will leak into the filter output. This is known as cross talk. Ideally, the cross talk should be zero. Let $\mathbf{a}_1(\omega)\Xi_1(\omega)$ be the Fourier transform of the signal emitted by the second source. The output of the filter tuned to the first source is given by

$$output = \begin{cases} \mathbf{H}_W^H(\omega,\theta_0)\mathbf{a}_0(\omega,\theta_0)\Xi_0(\omega) + \\ \mathbf{H}_W^H(\omega,\theta_0)\mathbf{a}_1(\omega,\theta_1)\Xi_1(\omega) + \mathbf{H}_W^H(\omega,\theta_0)N(\omega) \end{cases} \quad (6.5a)$$

Further, to evaluate different terms in (6.5a) we need the following results:

$$\mathbf{a}_0^H(\omega,\theta_0)\mathbf{V}^{-1}\mathbf{a}_1(\omega,\theta_1) = \frac{S_0(\omega)}{\sigma_\eta^2}\frac{\mathbf{a}_0^H(\omega,\theta_0)\mathbf{a}_1(\omega,\theta_1)}{1+M\dfrac{S_0(\omega)}{\sigma_\eta^2}} \quad (6.5b)$$

and

$$\mathbf{a}_1^H(\omega,\theta_1)\mathbf{V}^{-1}\mathbf{a}_1(\omega,\theta_1) =$$
$$\frac{S_0(\omega)}{\sigma_\eta^2}\left[M - Q\left|\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)\right|^2\right] \quad (6.5c)$$

Using (6.5b) and (6.5c) in (6.5a) we obtain

$$\mathbf{H}_W^H(\omega,\theta_0)\mathbf{a}_0(\omega,\theta_0)\Xi_0(\omega) =$$

$$MQ\Xi_0(\omega) - \frac{S_1(\omega)}{S_0(\omega)}\frac{Q^2\left|\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)\right|^2\Xi_0(\omega)}{1+\dfrac{S_1(\omega)}{\sigma_\eta^2}\left[M - Q\left|\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)\right|^2\right]} \quad (6.6a)$$

$$\mathbf{H}_W^H(\omega,\theta_0)\mathbf{a}_1(\omega,\theta_1)\Xi_1(\omega)$$

$$(6.6b)$$

$$= Q\frac{\mathbf{a}_0^H(\omega,\theta_0)\mathbf{a}_1(\omega,\theta_1)\Xi_1(\omega)}{1+\dfrac{S_1(\omega)}{\sigma_\eta^2}\left[M - Q\left|\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)\right|^2\right]}$$

We shall assume that the array signal-to-noise ratio (asnr) is much greater than one, that is, $\dfrac{MS_0(\omega)}{\sigma_\eta^2} \gg 1$. Using this approximation in (6.6) the array output power may be approximated as

$$Output\ power \approx S_0(\omega) +$$

$$
\left\{
\begin{array}{l}
\left(\dfrac{S_1(\omega)}{S_0(\omega)}\right)^2 \times \dfrac{\dfrac{\left|\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)\right|^4}{M^4}S_0(\omega)}{\left[1+\dfrac{S_1(\omega)}{\sigma_\eta^2}M\left[1-\dfrac{\left|\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)\right|^2}{M^2}\right]\right]^2} \\[3em]
-2\,\dfrac{\dfrac{\left|\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)\right|^2}{M^2}S_1(\omega)}{1+\dfrac{S_1(\omega)}{\sigma_\eta^2}M\left[1-\dfrac{\left|\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)\right|^2}{M^2}\right]} \\[3em]
+\dfrac{\dfrac{\left|\mathbf{a}_0^H(\omega,\theta_0)\mathbf{a}_1(\omega,\theta_1)\right|^2}{M^2}S_1(\omega)}{\left[1+\dfrac{S_1(\omega)}{\sigma_\eta^2}M\left[1-\dfrac{\left|\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)\right|^2}{M^2}\right]\right]^2} \\[3em]
+\mathbf{H}_W^H(\omega,\theta_0)\mathbf{H}_W(\omega,\theta_0)\sigma_\eta^2(\omega)
\end{array}
\right\}
\qquad (6.7)
$$

The first term is the desired signal power. The remaining terms represent the interference. Of these three, the magnitude of the first is much lower than that of the second and third terms. The magnitude of the first is proportional to $\dfrac{\left|\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)\right|^4}{M^4}$ while the magnitude of the second and third terms is proportional to $\dfrac{\left|\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)\right|^2}{M^2}$. Hence, we drop the first term from the interference expression. The cross talk, defined as a ratio of the power leaked from the second source and the actual power in the second source, is given by

$$
cross\ talk \approx \dfrac{\left|\mathbf{a}_0^H(\omega,\theta_0)\mathbf{a}_1(\omega,\theta_1)\right|^2}{M^2}\times
$$

$$\left\{ 1 - \frac{\left[\dfrac{S_1(\omega)}{\sigma_\eta^2} M\left(1 - \dfrac{\left|\mathbf{a}_0^H(\omega,\theta_0)\mathbf{a}_1(\omega,\theta_1)\right|^2}{M^2}\right) \right]^2}{\left[1 + \dfrac{S_1(\omega)}{\sigma_\eta^2} M\left(1 - \dfrac{\left|\mathbf{a}_0^H(\omega,\theta_0)\mathbf{a}_1(\omega,\theta_1)\right|^2}{M^2}\right) \right]^2} \right\} \qquad (6.8)$$

Note that $\dfrac{\left|\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)\right|^2}{M^2}$ represents the square of cosine of the angle

between two direction vectors $\mathbf{a}_0(\omega,\theta_0)$ and $\mathbf{a}_1(\omega,\theta_1)$. When

$\mathbf{a}_0(\omega,\theta_0) = \mathbf{a}_1(\omega,\theta_1)$, $\dfrac{\left|\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)\right|^2}{M^2} = 1$ the cross talk $= 1$, as

expected. But when $\mathbf{a}_0(\omega,\theta_0) \perp \mathbf{a}_1(\omega,\theta_1))$, $\dfrac{\left|\mathbf{a}_0^H(\omega,\theta_0)\mathbf{a}_1(\omega,\theta_1)\right|^2}{M^2} = 0$ and

cross talk is zero. Aside from these two extreme situations the cross talk may

be reduced to zero if $\dfrac{MS_1(\omega)}{\sigma_\eta^2}$ >>1, that is, the array signal-to-noise ratio

(asnr) for the second source is very large. The cross-talk is shown in
6.1.3 Linear Least Squares Estimate (LLSE): As previously stated the signal
model is

$$\mathbf{f}(t) = [\mathbf{a}_0, \mathbf{a}_1, \ldots \mathbf{a}_{P-1}] \begin{bmatrix} \xi_0(t) \\ \xi_1(t) \\ \\ \xi_{P-1}(t) \end{bmatrix} + \eta(t) \qquad (6.9)$$

$$= \mathbf{A}\xi(t) + \eta(t), \qquad t = 0, 1, \ldots, N$$

where $\mathbf{a}_0, \mathbf{a}_1, \ldots \mathbf{a}_{P-1}$ are steering vectors of P sources. It is assumed that $\mathbf{A}$ is

known and $\xi(t)$ is unknown but deterministic. The least squares estimate is

obtained by minimizing with respect to $\xi(t)$

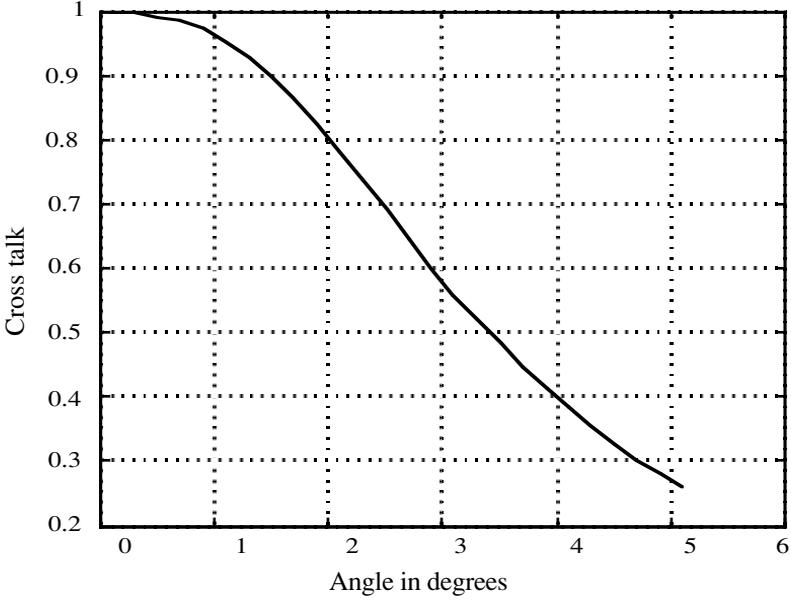$$\left\| \mathbf{f}(t) - \mathbf{A}\xi(t) \right\|^2 = \min \qquad (6.10)$$

Figure 6.1: Wiener Filter: Cross talk as a function of angular distance between two sources. The first source is on broadside (DOA=0$^{\text{o}}$). Array signal-to-noise ratio (asnr)=10 and eight sensor ULA is assumed.

Differentiating (6.10) with respect to $\xi(t)$ and setting the derivative to zero we obtain

$$
\begin{aligned}
\hat{\xi}(t) &= \left[\mathbf{A}^H\mathbf{A}\right]^{-1}\mathbf{A}^H\mathbf{f}(t) \\
&= \xi(t) + \left[\mathbf{A}^H\mathbf{A}\right]^{-1}\mathbf{A}^H\eta(t)
\end{aligned}
\tag{6.11}
$$

We observe that the signal term is extracted without any distortion but the noise term, given by $\left[\mathbf{A}^H\mathbf{A}\right]^{-1}\mathbf{A}^H\eta(t)$, behaves differently; for example, the noise becomes correlated. The output noise covariance matrix is given by

$$
\mathbf{C}_{\hat{\eta}} = \left[\mathbf{A}^H\mathbf{A}\right]^{-1}\sigma_{\eta}^2
\tag{6.12}
$$

When $\mathbf{A}^H\mathbf{A}$ is singular or close to being singular, that is, with a large eigenvalue spread, the noise in the output may get amplified. Consider a case of
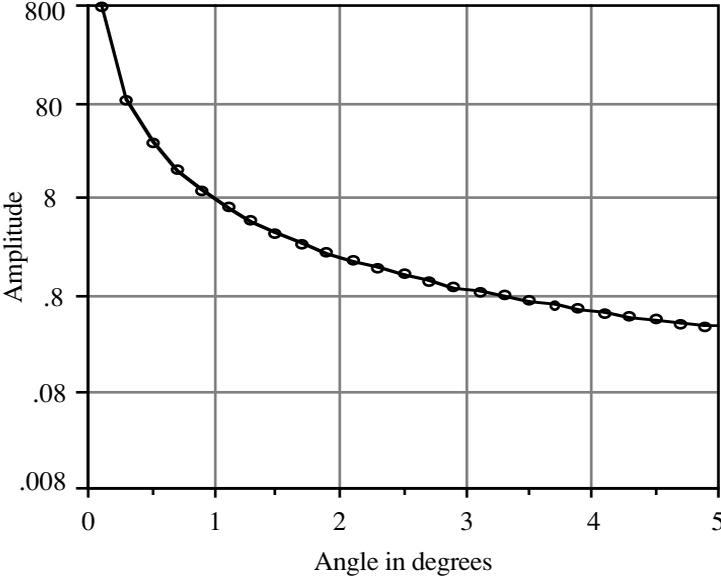
Figure 6.2: The noise amplification factor as a function of DOA of source #2. Source #1 is held fixed at 0° and source #2 is moved. Array size M=8.

two sources with direction vectors $\mathbf{a}_0$ and $\mathbf{a}_1$ and corresponding to directions of arrival θ_0 and θ_1, respectively. It is easy to show that

$$\det\left[\mathbf{A}^H\mathbf{A}\right] = M^2\left(1 - \left|\frac{\mathbf{a}_0^H\mathbf{a}_1}{M}\right|^2\right) \tag{6.13}$$

Using (6.13) in (6.12) we obtain the variance of the noise in all array outputs

$$\sigma_{\hat{\eta}}^2 = \frac{1}{M\left(1 - \left|\frac{\mathbf{a}_0^H\mathbf{a}_1}{M}\right|^2\right)}\sigma_{\eta}^2 \tag{6.14}$$

We have plotted in fig. 6.2 the noise amlpification factor, the factor multiplying σ_{η}^2, as a function of the angular separation between two sources.

From the plot we note that the noise is amplified only when the angular separation is a fraction of a degree.

In the stochastic model, the signal waveform is (stationary) stochastic and the LLSE turns out to be the same as the Wiener filter, which is easily derived as (Exercise #6.1),

$$\mathbf{H}_W(\omega) = \mathbf{S}_f^{-1}(\omega)\mathbf{A}\mathbf{S}_0(\omega)$$

where $\mathbf{A} = \left[\mathbf{a}_0, \mathbf{a}_1, ..., \mathbf{a}_{P-1}\right]$. In the Wiener filter we require spectral matrix and cross-spectral vectors; both of these are obtained by statistical averaging. Such a filter will naturally be applicable to an ensemble of time series having the same second order statistical properties. In the deterministic approach the LLSE filter is adapted to a particular data set. It is claimed in [1] that the stochastic LLSE results in a lower mean square error than the deterministic LLSE.

Circular Array: The result on cross talk (6.8) is of a general nature, valid for any array geometry. The direction vector needs to be appropriately defined. We now consider a circular array (UCA) for which the direction vector is given by (2.51) where we let $\theta = 90^o$ so that the source is in the same plane as the circular array,

$$\mathbf{a}(\varphi) = col\left[e^{-j\frac{\omega a}{c}\cos(\varphi)}, e^{-j\frac{\omega a}{c}\cos(\frac{2\pi}{M}-\varphi)}, ..., e^{-j\frac{\omega a}{c}\cos(\frac{2\pi(M-1)}{M}-\varphi)} \right] \quad (6.15)$$

The interest is to find out how a circular array fares in comparison with a linear array with respect to cross talk capability. For a given number of sensors (say, M) the maximum aperture of an ULA is fixed at $\dfrac{(M-1)\lambda}{2}$ but the aperture of a circular array can be very large, at least in principle. Since the array aperture is the main factor deciding the cross talk it is expected that the circular array ought to perform better in terms of lower cross talk. Using the same number of sensors, cross talk may be reduced when arranged over a large circle. This is shown in fig. 6.3 for an eight sensor UCA. The cross talk has been reduced considerably when the array radius is increased from four to thirty-two wavelengths. A linear array of eight sensors will have aperture of 3.5λ. An UCA having an aperture of 3.5λ was found to show much higher cross talk than that of the ULA shown in fig. 6.1. Thus, performance of an UCA is significantly better only when the radius is increased considerably.

6.1.4 Effects of Errors in DOA : Waveform estimation requires a knowledge of the direction of arrival of wavefronts. But the DOA estimates are subject to
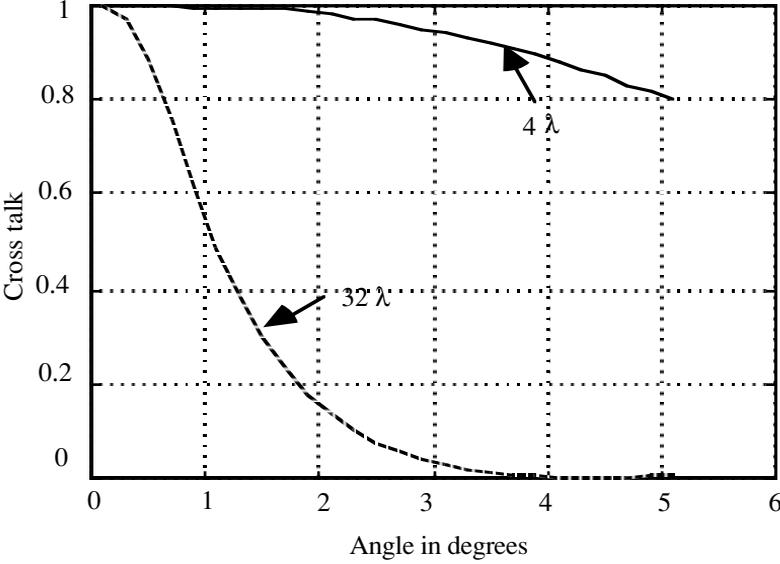
Figure 6.3: Cross talk as a function of angular separation of two sources. Eight sensors are uniformly spaced over a circle of radius 4λ (solid curve) and 32λ (dashed curve).

errors largely on account of finite data length used in their estimate. In this section we investigate the effects of errors in DOA estimates on interference from signals of no interest. A comprehensive analysis of the effects of model errors including the errors in DOA may be found in [2]. Let $\hat{\theta}_p = \theta_p + \Delta\theta_p$, $p = 0, 1, \ldots, P-1$ be the estimated DOAs, where θ_p is correct DOA and $\Delta\theta_p$ is estimation error. We shall rewrite (6.11) showing explicitly the dependence on the estimated DOAs

$$\hat{\xi}(t) = \left[\mathbf{A}^H(\hat{\theta})\mathbf{A}(\hat{\theta})\right]^{-1}\mathbf{A}^H(\hat{\theta})\mathbf{f}(t) \qquad (6.16)$$

Let us express $\mathbf{A}(\hat{\theta}) = \begin{bmatrix} \mathbf{B} & \mathbf{c} \end{bmatrix}$ where $\mathbf{B}$ is a matrix whose columns are the direction vectors of all interfering sources (signals of no interest) and $\mathbf{c}$ is a vector representing the direction vector of the signal of interest (soi). $\left[\mathbf{A}^H(\hat{\theta})\mathbf{A}(\hat{\theta})\right]^{-1}$ may be simplified [2] as

$$\left[\mathbf{A}^H(\hat{\theta})\mathbf{A}(\hat{\theta})\right]^{-1} = \left[[\mathbf{B}\ \mathbf{c}]^H[\mathbf{B}\ \mathbf{c}]\right]^{-1}$$

$$= \begin{bmatrix} \mathbf{B}^H\mathbf{B} & \mathbf{B}^H\mathbf{c} \\ \mathbf{c}^H\mathbf{B} & \mathbf{c}^H\mathbf{c} \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \left[\mathbf{B}^H\mathbf{B} - \dfrac{\mathbf{B}^H\mathbf{c}\mathbf{c}^H\mathbf{B}}{\|\mathbf{c}\|^2}\right]^{-1} & -\left[\mathbf{B}^H\mathbf{B} - \dfrac{\mathbf{B}^H\mathbf{c}\mathbf{c}^H\mathbf{B}}{\|\mathbf{c}\|^2}\right]^{-1}\dfrac{\mathbf{B}^H\mathbf{c}}{\|\mathbf{c}\|^2} \\ -\left[\mathbf{c}^H\mathbf{c} - \mathbf{c}^H\mathbf{P}\mathbf{c}\right]^{-1}\mathbf{c}^H\mathbf{B}_1 & \left[\mathbf{c}^H\mathbf{c} - \mathbf{c}^H\mathbf{P}\mathbf{c}\right]^{-1} \end{bmatrix} \tag{6.17}$$

$$= \dfrac{1}{\mathbf{c}^H(\mathbf{I} - \mathbf{P})\mathbf{c}} \begin{bmatrix} (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{c}^H(\mathbf{I} - \mathbf{P})\mathbf{c} + \mathbf{B}_1\mathbf{c}\mathbf{c}^H\mathbf{B}_1^H & -\mathbf{B}_1\mathbf{c} \\ -\mathbf{c}^H\mathbf{B}_1^H & 1 \end{bmatrix}$$

where $\mathbf{P} = \mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H$ and $\mathbf{B}_1 = (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H$. Using (6.17) in (6.16) and simplifying we obtain

$$\left[\mathbf{A}^H(\hat{\theta})\mathbf{A}(\hat{\theta})\right]^{-1}\mathbf{A}^H(\hat{\theta})$$

$$= \dfrac{1}{\mathbf{c}^H(\mathbf{I} - \mathbf{P})\mathbf{c}} \begin{bmatrix} (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{c}^H(\mathbf{I} - \mathbf{P})\mathbf{c} + \mathbf{B}_1\mathbf{c}\mathbf{c}^H\mathbf{B}_1^H & -\mathbf{B}_1\mathbf{c} \\ -\mathbf{c}^H\mathbf{B}_1^H & 1 \end{bmatrix} \begin{bmatrix} \mathbf{B}^H \\ \mathbf{c}^H \end{bmatrix} \tag{6.18}$$

$$= \dfrac{1}{\mathbf{c}^H(\mathbf{I} - \mathbf{P})\mathbf{c}} \begin{bmatrix} \mathbf{B}_1\left[\mathbf{c}^H(\mathbf{I} - \mathbf{P})\mathbf{c}\mathbf{I} - \mathbf{c}\mathbf{c}^H(\mathbf{I} - \mathbf{P})\right] \\ \mathbf{c}^H(\mathbf{I} - \mathbf{P}) \end{bmatrix}$$

Using (6.18) in (6.16) an estimate of the signal of interest (soi) is obtained

$$\hat{\xi}_{soi}(t) = \dfrac{\mathbf{c}^H(\mathbf{I} - \mathbf{P})}{\mathbf{c}^H(\mathbf{I} - \mathbf{P})\mathbf{c}}\mathbf{f}(t) = \mathbf{d}^H(\hat{\theta})\mathbf{f}(t) \tag{6.19}$$

We now expand $\mathbf{d}^H(\hat{\theta})$ in a Taylor's series and retain the first derivative term only, $\mathbf{d}(\hat{\theta}) \approx \mathbf{d}(\theta) + \mathbf{d}_1(\theta)\Delta\theta$, where $\mathbf{d}_1(\theta)$ is the first derivative of $\mathbf{d}(\theta)$ with respect to θ. The estimated signal of interest reduces to

$$\hat{\xi}_{soi}(t) = \mathbf{d}^H(\theta)\mathbf{A}\xi(t) + \Delta\theta^H\mathbf{d}_1^H(\theta)\mathbf{A}\xi(t) + \mathbf{d}(\hat{\theta})\eta(t)$$
$$= \xi_{soi}(t) + \Delta\theta^H\mathbf{d}_1^H(\theta)\mathbf{A}\xi(t) + \mathbf{d}(\hat{\theta})\eta(t) \qquad (6.20)$$

The interference term in (6.20) is $\Delta\theta^H\mathbf{d}_1^H(\theta)\mathbf{A}\xi(t)$ which we shall rewrite showing the contribution of each DOA

$$\Delta\xi_{soi}(t) = \sum_{p=0}^{P-1}\left[\beta^H\xi(t)\right]_p \Delta\theta_p \qquad (6.21)$$

where $\Delta\xi_{soi}(t)$ stands for the error in $\hat{\xi}_{soi}(t)$, $\beta = \left[\mathbf{A}^H\mathbf{d}_1(\theta)\right]$ and $[.]_p$ stands for the pth element of a column vector.

The errors in the calibration of an array can cause serious errors in the waveform estimation [3]. It is, therefore, imperative to carefully calibrate the array response, particularly when subspace methods are used for localization or waveform estimation. One simple solution to this problem of high sensitivity to calibration errors is to reduce the degrees of freedom by using a low rank approximation of the spectral matrix [4, 5]. In this approximation a spectral matrix is simply approximated by its signal subspace, that is,

$$\mathbf{S}_f(\omega) = \mathbf{v}_s\alpha_s\mathbf{v}_s^H + \mathbf{v}_\eta\alpha_\eta\mathbf{v}_\eta^H$$
$$\approx \mathbf{v}_s\alpha_s\mathbf{v}_s^H \qquad (6.22)$$

where $\alpha_s = diag\{\alpha_0,\alpha_1,...,\alpha_{P-1}\}$ and $\alpha_\eta = diag\{\alpha_P,\alpha_{P+1},...,\alpha_{M-1}\}$.

§6.2 **Minimum Variance** (Capon Method):
In §4.3 we have devised a filter (in the frequency domain) which minimizes the output power while maintaining unit response within an angular sector (extended source) whose axis is directed in the desired direction. This leads to the following equation,

$$\frac{1}{\lambda}\mathbf{H} = (\mathbf{S}_f^{-1}\Gamma)\mathbf{H} \qquad (6.23a)$$

where

$$\Gamma = \frac{1}{\Delta\theta}\int_{\theta-\frac{\Delta\theta}{2}}^{\theta+\frac{\Delta\theta}{2}}\mathbf{a}(\omega,\phi)\mathbf{a}^H(\omega,\phi)d\phi \qquad (6.23b)$$

$\mathbf{a}(\omega,\phi)$ is the steering vector and $\lambda = \mathbf{H}^H \mathbf{S}_f \mathbf{H}$ is output power. To further minimize the output power we must select a filter vector as the eigenvector of $(\mathbf{S}_f^{-1}\Gamma)$ corresponding to the largest eigenvalue; in which case the output power will be equal to the inverse of the largest eigenvalue of $(\mathbf{S}_f^{-1}\Gamma)$. A closed form solution of (6.23) can be obtained for $\Delta\theta = 0$ (point source). Equation (6.23) simplifies to

$$\frac{\mathbf{H}}{\mathbf{H}^H \mathbf{S}_f \mathbf{H}} = (\mathbf{S}_f^{-1}\mathbf{a}(\omega,\theta)\mathbf{a}^H(\omega,\theta)\mathbf{H} \qquad (6.24)$$

Multiply both sides of (6.24) by $\mathbf{a}^H(\omega,\theta)$ and simplify the resulting expression,

$$\mathbf{H}^H \mathbf{S}_f \mathbf{H} = \frac{1}{\mathbf{a}^H(\omega,\theta)\mathbf{S}_f^{-1}\mathbf{a}(\omega,\theta)} \qquad (6.25)$$

By direct substitution in (6.24) it may be verified that the solution is given by

$$\mathbf{H}_{cap} = \frac{\mathbf{S}_f^{-1}\mathbf{a}(\omega,\theta)}{\mathbf{a}^H(\omega,\theta)\mathbf{S}_f^{-1}\mathbf{a}(\omega,\theta)} \qquad (6.26)$$

and the output power is equal to $\dfrac{1}{\mathbf{a}^H(\omega,\theta)\mathbf{S}_f^{-1}\mathbf{a}(\omega,\theta)}$ The optimum filter given in (6.26) is also known as the Applebaum filter [6].

6.2.1 <u>Extended Source</u>: For ULA the Γ matrix takes the following form

$$[\Gamma]_{m,n} = \sum_{k=0,2,4...} \varepsilon_k J_k((m-n)\frac{\omega d}{c})\cos(k\theta)\frac{\sin(k\frac{\Delta\theta}{2})}{k\frac{\Delta\theta}{2}} +$$

$$j \sum_{k=1,3,5...} 2J_k((m-n)\frac{\omega d}{c})\sin(k\theta)\frac{\sin(k\frac{\Delta\theta}{2})}{k\frac{\Delta\theta}{2}}$$

where $\varepsilon_0 = 1$ and $\varepsilon_n = 2$ for all k. To evaluate the integral in (6.23b) we have used the result derived in (1.53b). For UCA the Γ matrix takes the following form:

$$\Gamma \;=\; \mathbf{W} \quad \Theta \quad \mathbf{W}^H$$

$$(M \times M) \quad (M \times \alpha) \quad (\alpha \times \alpha) \quad (\alpha \times M) \quad \alpha \to \infty$$

where

$$\{\mathbf{W}\}_{km} = e^{j\frac{2\pi km}{M}} \quad \left\{ \begin{matrix} m = 0,1,\dots M-1 \\ k = 0,\pm 1,\dots \infty \end{matrix} \right\}$$

$$[\Theta]_{k,l} = J_k(\frac{\omega a}{c})J_l(\frac{\omega a}{c})e^{-j(\frac{\pi}{2}+\varphi_0)(k-l)}\sin c((k-l)\frac{\Delta\theta}{2})$$

valid for all values of $\Delta\theta$. We have used a series expansion of the steering vector for a circular array (2.51).

6.2.2 <u>Single Point Source Case</u>: Let there be a single source emitting a plane wave with a direction vector $\mathbf{a}_0(\omega,\theta_0)$ in presence of white noise. The spectral matrix is given by $\mathbf{S}_f = \mathbf{a}_0(\omega,\theta_0)\mathbf{a}_0^H(\omega,\theta_0)\mathbf{S}_\xi(\omega)+\sigma_\eta^2\mathbf{I}$, and its inverse is given by

$$\mathbf{S}_f^{-1} = \frac{1}{\sigma_\eta^2}\left[\mathbf{I} - Q\mathbf{a}_0(\omega,\theta_0)\mathbf{a}_0^H(\omega,\theta_0)\right] \tag{6.27}$$

where

$$Q = \frac{\dfrac{\mathbf{S}_\xi(\omega)}{\sigma_\eta^2}}{1 + M\dfrac{\mathbf{S}_\xi(\omega)}{\sigma_\eta^2}}$$

For large array signal-to-noise ratio, $\dfrac{\mathbf{S}_\xi(\omega)}{\sigma_\eta^2} \gg 1$, we may approximate

$Q \approx \dfrac{1}{M}$. Using (6.27) in (6.26) we obtain

$$\mathbf{H}_{cap} = \frac{1}{M}\left[\frac{\mathbf{a}(\omega,\theta)-Q\left[\mathbf{a}_0^H(\omega,\theta_0)\mathbf{a}(\omega,\theta)\right]\mathbf{a}_0(\omega,\theta_0)}{1-Q\dfrac{\left|\mathbf{a}^H(\omega,\theta)\mathbf{a}_0(\omega\theta_0)\right|^2}{M}}\right] \tag{6.28}$$

Let us obtain the filter output for the signal model of a single source in the presence of white noise

$$\hat{F}(\omega) = \mathbf{H}_{cap}^{H} \left[\mathbf{a}_0(\omega, \theta_0) \Xi_0(\omega) + \mathrm{N}(\omega) \right] \qquad (6.29a)$$

The signal term turns out to be

$$\mathbf{H}_{cap}^{H} \mathbf{a}_0(\omega, \theta_0) \Xi_0(\omega)$$

$$= \frac{\mathbf{a}^{H}(\omega, \theta) \mathbf{a}_0(\omega, \theta_0)}{M} \left[\frac{1 - QM}{1 - QM \dfrac{\left| \mathbf{a}^{H}(\omega, \theta) \mathbf{a}_0(\omega\theta_0) \right|^2}{M^2}} \right] \Xi_0(\omega) \qquad (6.29b)$$

$$= \Xi_0(\omega) \qquad \text{when } \mathbf{a}^{H}(\omega, \theta) = \mathbf{a}_0(\omega, \theta_0)$$

and the noise term

$$\mathbf{H}_{cap}^{H} \mathrm{N}(\omega)$$

$$= \frac{1}{M} \left[\frac{\mathbf{a}(\omega, \theta) - Q \left[\mathbf{a}_0^{H}(\omega, \theta_0) \mathbf{a}(\omega, \theta) \right] \mathbf{a}_0(\omega, \theta_0)}{1 - QM \dfrac{\left| \mathbf{a}^{H}(\omega, \theta) \mathbf{a}_0(\omega\theta_0) \right|^2}{M^2}} \right] \mathrm{N}(\omega) \qquad (6.29c)$$

$$= \frac{\mathbf{a}_0^{H}(\omega, \theta_0)}{M} \mathrm{N}(\omega) \qquad \text{when } \mathbf{a}^{H}(\omega, \theta) = \mathbf{a}_0(\omega, \theta_0)$$

The variance of the noise in the filter output (6.29c), when $\mathbf{a}^{H}(\omega, \theta) = \mathbf{a}_0(\omega, \theta_0)$, turns out to be $\dfrac{\sigma_N^2}{M}$. The response of the filter was computed from (6.29a) for different steering directions. The DOA of the incident wavefront was assumed to be at 0°. An ULA with eight sensors was considered. The response of the filter depends upon the signal-to-noise ratio (snr). For snr=10 (or asnr=80) the response becomes extremely sharp as demonstrated in . For comparison we have also plotted a Wiener filter response (dashed curve) where we have assumed the signal-to-noise ratio equal to ten while other parameters remain the same as for the Capon filter.
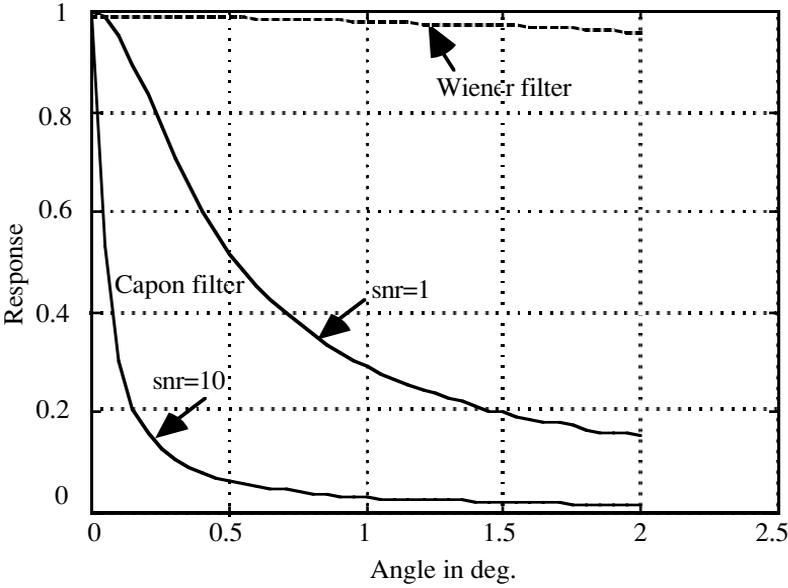
Figure 6.4: Response of Capon filter as a function of the angle of the steering vector. An ULA of 8 elements was assumed. Equivalent Wiener filter response is also shown by the dashed curve.

The response of the Capon filter is much superior to that of an equivalent Wiener filter. The shape of the Capon filter is strongly dependent on signal-to-noise ratio; on the other hand the dependence of the Wiener filter on the signal-to-noise ratio is merely to the extent of changing the scale factor. Even at a very low signal-to-noise ratio the Capon filter seems to outperform the Wiener filter.

6.2.3 <u>Two Source Case</u>: As in the case of the Wiener filter, we shall now consider the Capon filter specifically for two uncorrelated wavefronts in the presence of white noise. The spectral matrix is given by (6.2) and its inverse is computed in the same manner as shown there. Using the expression for inverse of the spectral matrix in (6.26) we obtain an expression for the Capon filter,

$$\mathbf{H}_{Cap}(\omega,\theta_0) = \frac{num}{den} \text{ where}$$

$$num = \left[\mathbf{V}^{-1} - \frac{S_1(\omega)}{S_0(\omega)} \frac{\mathbf{V}^{-1}\mathbf{a}_1(\omega,\theta_1)\mathbf{a}_1^H(\omega,\theta_1)\mathbf{V}^{-1}}{1 + \frac{S_1(\omega)}{S_0(\omega)}\mathbf{a}_1^H(\omega,\theta_1)\mathbf{V}^{-1}\mathbf{a}_1(\omega,\theta_1)} \right] \mathbf{a}_0(\omega,\theta_0)$$

(6.30a)

$den = \mathbf{a}_0^H(\omega,\theta_0) \cdot num$ and $\mathbf{V}$ is defined in (6.4b). The denominator may be further reduced to

$$den = \frac{S_0(\omega)}{\sigma_\eta^2}(1-QM)\frac{M(1+\dfrac{S_1(\omega)}{\sigma_\eta^2}\left[1-\dfrac{\left|\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)\right|^2}{M^2}\right])}{(1+\dfrac{S_1(\omega)}{\sigma_\eta^2}\left[1-QM\dfrac{\left|\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)\right|^2}{M^2}\right])}$$

(6.30b)

The filter output may be shown to be

$$output = \left\{\begin{array}{l}\Xi_0(\omega)\\ +\mathbf{H}_{Cap}^H(\omega,\theta_0)\mathbf{a}_1(\omega,\theta_1)\Xi_1(\omega) + \mathbf{H}_{Cap}^H(\omega,\theta_0)N(\omega)\end{array}\right\} \quad (6.31)$$

Notice that the signal term is equal to the actual signal. But this was not so in the case of the Wiener filter where we had to assume a large array snr in order to arrive at this result.

The contribution of the second source is represented by the second term in (6.31). The cross talk is then given by

$$Cross\ talk = \mathbf{a}_1^H(\omega,\theta_1)\mathbf{H}_{Cap}(\omega,\theta_0)\mathbf{H}_{Cap}^H(\omega,\theta_0)\mathbf{a}_1(\omega,\theta_1)$$

$$= \frac{\left|\dfrac{\mathbf{a}_0^H(\omega,\theta_1)\mathbf{a}_1(\omega,\theta_0)}{M}\right|^2}{\left\{1+\dfrac{S_1(\omega)}{\sigma_\eta^2}\left[1-\dfrac{\left|\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)\right|^2}{M^2}\right]\right\}^2} \quad (6.32)$$

The cross talk is plotted in fig. 6.5 for an ULA and in fig. 6.6 for an UCA. Compare these two figures with figures 6.1 and 6.3 where we have plotted the cross talk for the Wiener filter. Clearly, the Capon filter performs better for both types of array geometries. An expression for snr which agrees with (6.32) is derived in [7].

Finally, we shall evaluate the noise term, that is, the leftover noise in the array output,
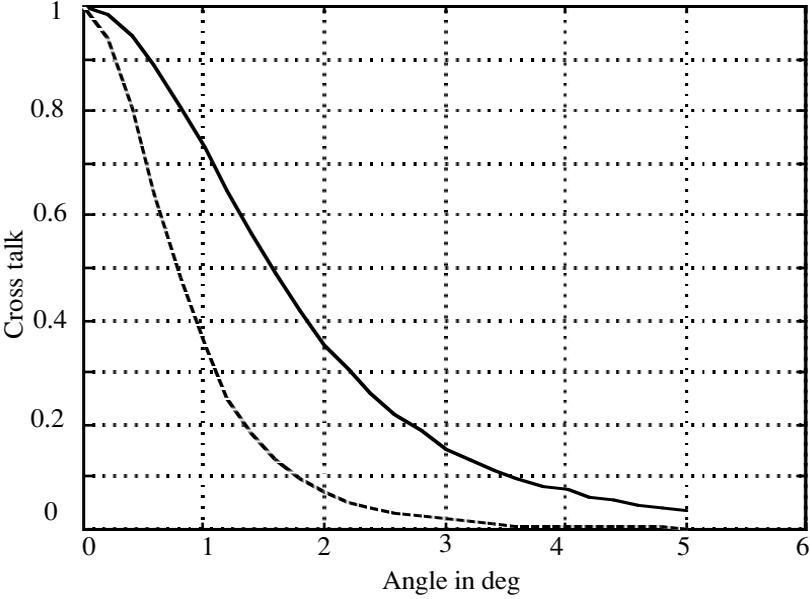
Figure 6.5: Capon Filter: Cross talk as a function of angular distance between two sources. The first source is on broadside (DOA=0°). Array signal-to-noise ratio (asnr)=10. Continuous curve: 8 sensor ULA and Dashed curve: 16 sensor ULA.

$$\mathbf{H}_{Cap}^{H}(\omega,\theta_0)N(\omega)$$

$$= \frac{\mathbf{a}_0^{H}(\omega,\theta_0)\left[\mathbf{V}^{-1} - \frac{S_1(\omega)}{S_0(\omega)}\frac{\mathbf{V}^{-1}\mathbf{a}_1(\omega,\theta_1)\mathbf{a}_1^{H}(\omega,\theta_1)\mathbf{V}^{-1}}{1 + \frac{S_1(\omega)}{S_0(\omega)}\mathbf{a}_1^{H}(\omega,\theta_1)\mathbf{V}^{-1}\mathbf{a}_1(\omega,\theta_1)}\right]N}{den} \quad (6.33)$$

$$= \frac{\mathbf{a}_0^{H}(\omega,\theta_0)\mathbf{V}^{-1}N - \psi\mathbf{a}_1^{H}(\omega,\theta_1)\mathbf{V}^{-1}N}{den}$$

where the denominator term, den, is given in (6.30b) and

$$\psi = \frac{\frac{S_1(\omega)}{S_0(\omega)}\mathbf{a}_0^{H}(\omega,\theta_0)\mathbf{V}^{-1}\mathbf{a}_1(\omega,\theta_1)}{1 + \frac{S_1(\omega)}{S_0(\omega)}\mathbf{a}_1^{H}(\omega,\theta_1)\mathbf{V}^{-1}\mathbf{a}_1(\omega,\theta_1)}$$
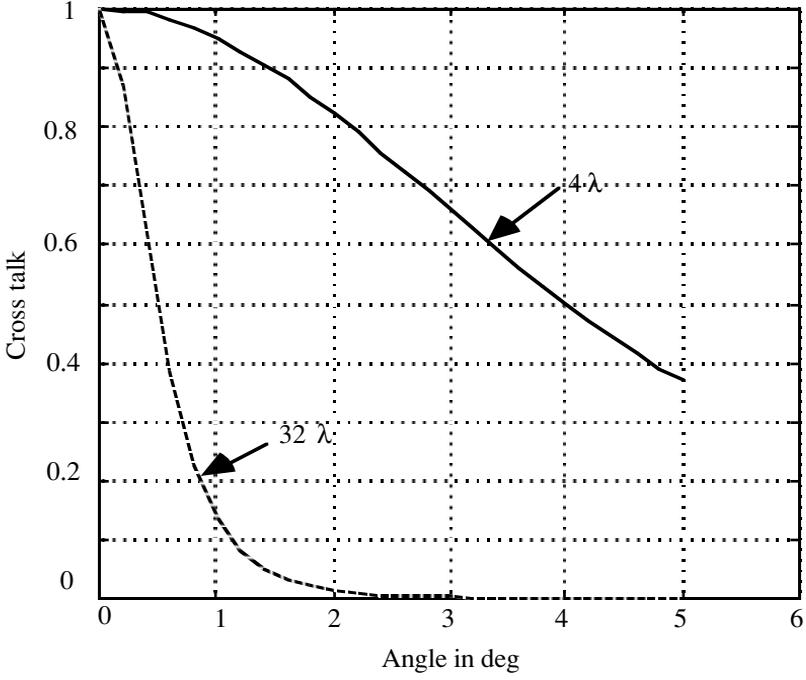
Figure 6.6: Capon Filter: Cross talk as a function of angular separation of two sources. Eight sensors are uniformly spaced over a circle of radius 4λ (solid curve) and 32λ (dashed curve).

The variance of the leftover noise may be computed from (6.33). We obtain after simplification

$$
Var\left\{\mathbf{H}_{Cap}^{H}(\omega,\theta_0)N(\omega)\right\}
$$

$$
= \frac{\sigma_\eta^2}{M}\left[1 - \psi\frac{\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)}{M}\right]
$$

$$
= \frac{\sigma_\eta^2}{M}\left[\frac{1 + \dfrac{S_1(\omega)}{\sigma_\eta^2}M(1 - \left|\dfrac{\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)}{M}\right|^2)}{1 + \dfrac{S_1(\omega)}{\sigma_\eta^2}M(1 - QM\left|\dfrac{\mathbf{a}_1^H(\omega,\theta_1)\mathbf{a}_0(\omega,\theta_0)}{M}\right|^2)}\right] \qquad (6.34a)
$$

When the array snr is large, then $QM \approx 1$ and (6.34) reduces to a simple form

$$Var\left\{\mathbf{H}_{Cap}^{H}(\omega,\theta_0)N(\omega)\right\} \approx \frac{\sigma_\eta^2}{M} \qquad (6.34b)$$

§6.3 Adaptive Beamformation:

In adaptive beamformation the array processor is so designed that it receives a signal coming from a desired direction and it *automatically* suppresses signals (that is, interference) coming from all other directions. This is achieved by means of a filter which adapts itself to the incoming signal and interference. Suppression of interference is achieved through predictive cancellation; an example of this approach was described in chapter 2 in connection with noise cancellation. The weight coefficients are upgraded through an algorithm, such as LMS or one of its kind. There are excellent texts on the topic of adaptive signal processing which includes adaptive beamformation [8,9]. We do not intend to cover this topic in any detail. The aim here is to briefly discuss a few selected topics, namely, null steering, adaptive interference canceller and adaptive Capon filter.

6.3.1 <u>Null Steering</u>: A basic step in beamformation is the weighted and delayed sum of the array outputs

$$\hat{f}(t) = \sum_{m=0}^{M-1} a_m f_m(t - m\tau) \qquad (6.35a)$$

where τ is delay per sensor and a_m is a real weight coefficient. In temporal frequency domain (6.35a) may be expressed as

$$\hat{F}(\omega) = \sum_{m=0}^{P-1} (a_m e^{-jm\tau\omega}) F(md,\omega)$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} A(u) F^*(u,\omega) du \qquad (6.35b)$$

where

$$A(u) = \sum_{m=0}^{P-1} \tilde{a}_m e^{-jmu}$$
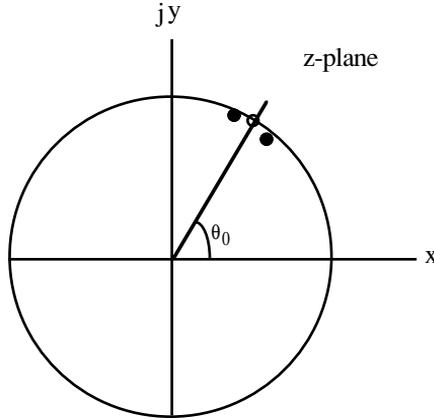
$$\tilde{a}_m = a_m e^{-jm\tau\omega}$$

Figure 6.7: Position of a null and two poles in its immediate neighborhood but slightly inside the unit circle.

$$F(u,\omega) = \sum_{m=0}^{P-1} F(md,\omega)e^{-jmu}$$

Through a process of selecting the complex weight coefficients we can assign the desired property to the response function of the weight function; for example, $A(u)$ may be required to have a deep null in some desired direction and unit response in all directions. Let us express $A(u)$ as a z-transform of $\tilde{a}_m$,, $A(z) = \sum_{m=0}^{P-1} \tilde{a}_m z^{-1}$ where $z = e^{ju}$. We can now design a notch filter in the z-plane. A simple approach is to place a zero on the unit circle and two poles, one on either side of the zero. Let the zero be at $z_0 = e^{j\pi \sin\theta_0}$ and the poles at $z_+ = re^{j\pi \sin(\theta_0 + \frac{\Delta\theta}{2})}$ and $z_- = re^{j\pi \sin(\theta_0 - \frac{\Delta\theta}{2})}$ (see fig. 6.7) where $0.9 \le r \le 0.99$. We have assumed that $d = 0.5\lambda$.

$$A(z) = \frac{(z - z_0)}{(z - z_+)(z - z_-)} \tag{6.36}$$

The response of the notch filter given by (6.36) is shown in fig. 6.8. See exercise 3 for another type of null steering filter independent of frequency.

6.3.2 <u>Beam Steered Adaptive Array</u>: In Capon filter (6.26) let us replace the inverse of the spectral matrix by its eigenvalue eigenvector representation,
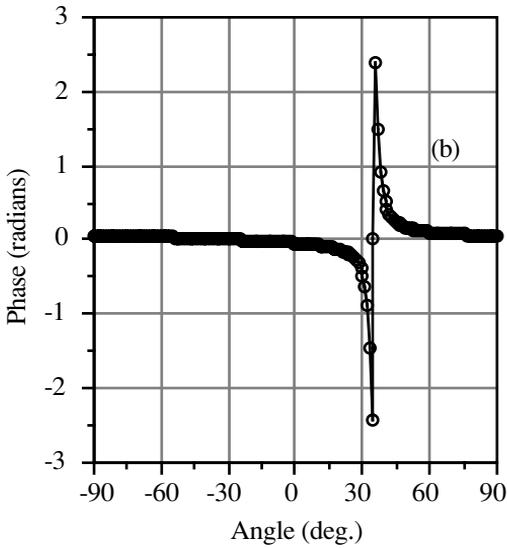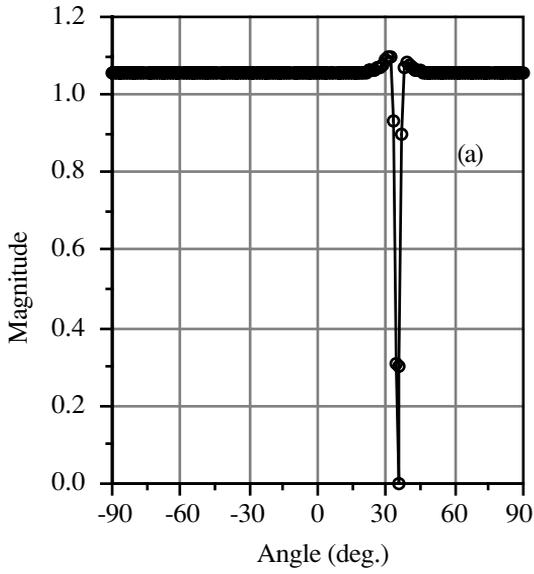
Figure 6.8: Wavenumber response of the notch filter (a) amplitude response and (b) phase response.

$$\mathbf{H}_{cap} = \frac{\mathbf{S}_f^{-1}\mathbf{a}(\omega,\theta)}{\mathbf{a}^H(\omega,\theta)\mathbf{S}_f^{-1}\mathbf{a}(\omega,\theta)}$$

$$= \frac{\left\{\mathbf{v}_s\lambda_s^{-1}\mathbf{v}_s^H + \mathbf{v}_\eta\lambda_\eta^{-1}\mathbf{v}_\eta^H\right\}\mathbf{a}(\omega,\theta)}{\mathbf{a}^H(\omega,\theta)\left\{\mathbf{v}_s\lambda_s^{-1}\mathbf{v}_s^H + \mathbf{v}_\eta\lambda_\eta^{-1}\mathbf{v}_\eta^H\right\}\mathbf{a}(\omega,\theta)} \tag{6.37a}$$

We assume that the look direction vector $\mathbf{a}(\omega,\theta)$ lies in the signal subspace; hence $\mathbf{a}(\omega,\theta)\perp\mathbf{v}_\eta$. Hence (6.37a) simplifies to

$$\mathbf{H}_{cap} = \frac{\left\{\mathbf{v}_s\lambda_s^{-1}\mathbf{v}_s^H\right\}\mathbf{a}(\omega,\theta)}{\mathbf{a}^H(\omega,\theta)\left\{\mathbf{v}_s\lambda_s^{-1}\mathbf{v}_s^H\right\}\mathbf{a}(\omega,\theta)} \tag{6.37b}$$

The Capon filter has an interesting property, that is, when the look direction coincides with one of the signal directions, for example, $\mathbf{a}(\omega,\theta) = \mathbf{a}(\omega,\theta_0)$ it is approximately orthogonal to all direction vectors, $\mathbf{a}(\omega,\theta_m), m = 1,2,...,P-1$. To show this, consider

$$\mathbf{H}_{cap}^H\mathbf{a}(\omega,\theta_m) = \frac{\mathbf{a}^H(\omega,\theta_0)\left\{\mathbf{v}_s\lambda_s^{-1}\mathbf{v}_s^H\right\}\mathbf{a}(\omega,\theta_m)}{\mathbf{a}^H(\omega,\theta_0)\left\{\mathbf{v}_s\lambda_s^{-1}\mathbf{v}_s^H\right\}\mathbf{a}(\omega,\theta_0)}, \quad m \neq 0 \tag{6.38a}$$

In (6.38a) the inverse of the signal eigenvalues may be approximated, for large snr or large array size, by $\lambda_m^{-1} \approx \alpha_m^{-1} - \dfrac{\sigma_\eta^2}{\alpha_m^2}$. Therefore, (6.38a) may be expressed as

$$\mathbf{H}_{cap}^H\mathbf{a}(\omega,\theta_m) \approx \left\{ \begin{array}{l} \dfrac{\mathbf{a}^H(\omega,\theta_0)\left\{\mathbf{v}_s\alpha_s^{-1}\mathbf{v}_s^H\right\}\mathbf{a}(\omega,\theta_m)}{\mathbf{a}^H(\omega,\theta_0)\left\{\mathbf{v}_s\lambda_s^{-1}\mathbf{v}_s^H\right\}\mathbf{a}(\omega,\theta_0)} \\[1.5em] -\dfrac{\mathbf{a}^H(\omega,\theta_0)\sum\limits_{m=0}^{P-1}\dfrac{\sigma_\eta^2}{\alpha_m^2}\mathbf{v}_m\mathbf{v}_m^H\mathbf{a}(\omega,\theta_m)}{\mathbf{a}^H(\omega,\theta_0)\left\{\mathbf{v}_s\lambda_s^{-1}\mathbf{v}_s^H\right\}\mathbf{a}(\omega,\theta_0)} \end{array} \right\}, \quad m \neq 0 \tag{6.38b}$$

We shall show that the first term inside the brackets is indeed zero. Let us consider the signal term alone from (4.12b) and compute the pseudoinverse (denoted by #) on both sides of the equation

$$\mathbf{S}_f^{\#} = \mathbf{A}^{\#}\mathbf{S}_0^{-1}\mathbf{A}^{\#H} \tag{6.39}$$

On premultiplying by $\mathbf{A}^H$ and postmultiplying by $\mathbf{A}$ both sides of (6.39) we obtain

$$\mathbf{A}^H\mathbf{S}_f^{\#}\mathbf{A} = \mathbf{A}^H\mathbf{A}^{\#}\mathbf{S}_0^{-1}\mathbf{A}^{\#H}\mathbf{A} = \mathbf{S}_0^{-1} \tag{6.40}$$

When the sources are *uncorrelated*, $\mathbf{S}_0$ as well as $\mathbf{S}_0^{-1}$ will be diagonal matrices. It, therefore, follows that all cross product terms in $\mathbf{A}^H\mathbf{S}_f^{\#}\mathbf{A}$ must be equal to zero. The numerator of the first term is one such cross product term. The second term in (6.38b) will be small when the asnr (array snr) is high Thus, we obtain the following approximate result

$$\mathbf{H}_{cap}^H\mathbf{a}(\omega,\theta_m) \approx 0 \quad m = 1, 2, ..., P-1 \tag{6.41}$$

A modified Capon filter is defined as

$$\tilde{\mathbf{H}}_{cap} = \frac{\mathbf{S}_s^{\#}(\omega)\mathbf{a}(\omega,\theta)}{\mathbf{a}^H(\omega,\theta)\mathbf{S}_s^{\#}(\omega)\mathbf{a}(\omega,\theta)} \tag{6.42}$$

where $\mathbf{S}_s^{\#}(\omega)$ is the pseudoinverse of the signal-only spectral matrix. It is shown in [10] that the modified Capon filter is robust against look direction errors. It was shown in chapter 5 (page 274) that the generalized eigenvector corresponding to a given direction of arrival (DOA) is orthogonal to all other direction vectors.

6.3.3 <u>Adaptive Capon filter</u>: In chapter 3 we have derived two different filters to extract a wavefront coming from a specified direction. It may be recalled that the Wiener filter minimizes the mean square error between the filter output and the desired signal and the Capon filter minimizes the output power while maintaining unit response in the desired direction. The Wiener filter is given by $\mathbf{H}_W(\omega) = \mathbf{S}_f^{-1}(\omega)\mathbf{S}_0(\omega)$ and the Capon filter is given by $\mathbf{H}_{Cap}(\omega) = \dfrac{\mathbf{S}_f^{-1}(\omega)\mathbf{a}(\omega,\theta)}{\mathbf{a}^H(\omega,\theta)\mathbf{S}_f^{-1}(\omega)\mathbf{a}(\omega,\theta)}$ where $\mathbf{a}(\omega,\theta)$ is the direction vector in the desired direction. Note that $\mathbf{S}_0(\omega) = \mathbf{a}(\omega,\theta)S_0(\omega)$ where $S_0(\omega)$ is power from the desired direction. In the Wiener filter this power is assumed to be a constant but in the Capon filter the power is estimated

$(= \dfrac{1}{\mathbf{a}^H(\omega,\theta)\mathbf{S}_f^{-1}(\omega)\mathbf{a}(\omega,\theta)})$ and it is a function of the desired direction. This subtle difference is probably responsible for the improved performance of the Capon filter.

To implement the Wiener or Capon filter we need the inverse of the spectral matrix $\mathbf{S}_f^{-1}(\omega)$. In practical terms with the arrival of new data in the form of new time or frequency snapshot we should be able to improve upon the available estimate of the spectral matrix and its inverse. In chapter 4 we have shown how to recursively estimate these quantities. We shall rewrite equations (4.62a and 4.62c) in a more general fashion

$$\hat{\mathbf{S}}_f^{N+1}(\omega) = \mu\hat{\mathbf{S}}_f^{N}(\omega) + (1-\mu)\mathbf{F}_{N+1}(\omega)\mathbf{F}_{N+1}^{H}(\omega) \qquad (6.43a)$$

$$\left[\hat{\mathbf{S}}_f^{N+1}(\omega)\right]^{-1} =$$
$$\frac{1}{\mu}\left[\hat{\mathbf{S}}_f^{N}(\omega)\right]^{-1} - \frac{(1-\mu)\mathbf{z}\mathbf{z}^{H}}{\mu\left((1-\mu)\mathbf{F}_{N+1}^{H}(\omega)\left[\hat{\mathbf{S}}_f^{N}(\omega)\right]^{-1}\mathbf{F}_{N+1}(\omega) + \mu\right)} \qquad (6.43b)$$

where $\mathbf{z} = \left[\hat{\mathbf{S}}_f^{N}(\omega)\right]^{-1}\mathbf{F}_{N+1}(\omega)$ and $\mu \in (0,1)$, which is a free parameter to be chosen depending upon how fast the estimated spectral matrix changes from snapshot to snapshot. For stationary process where the change is small, $\mu \approx 1$. Using (6.43a) and (6.43b) we can recursively compute the Wiener and Capon filters. We will do this for the Capon filter. Multiply both sides of (6.43b) by $\dfrac{\mathbf{a}(\omega,\theta)}{\mathbf{a}^H(\omega,\theta)\left[\hat{\mathbf{S}}_f^{N}(\omega)\right]^{-1}\mathbf{a}(\omega,\theta)}$. Noting the definition of the Capon filter and assuming that the power from the desired direction does not change much from snapshot to snapshot we obtain

$$\mathbf{H}_{Cap}^{N+1} = \frac{1}{\mu}\mathbf{H}_{Cap}^{N} - \beta\mathbf{z}\mathbf{z}^{H}\frac{\mathbf{a}(\omega,\theta)}{\mathbf{a}^H(\omega,\theta)\left[\hat{\mathbf{S}}_f^{N}(\omega)\right]^{-1}\mathbf{a}(\omega,\theta)} \qquad (6.43c)$$

where

$$\beta = \frac{(1-\mu)}{\mu\left((1-\mu)\mathbf{F}_{N+1}^{H}(\omega)\left[\hat{\mathbf{S}}_f^{N}(\omega)\right]^{-1}\mathbf{F}_{N+1}(\omega) + \mu\right)}$$

But, we can show that

$$\mathbf{z}^H \frac{\mathbf{a}(\omega,\theta)}{\mathbf{a}^H(\omega,\theta)\left[\hat{\mathbf{S}}_f^N(\omega)\right]^{-1}\mathbf{a}(\omega,\theta)} = \mathbf{F}_{N+1}^H(\omega)\mathbf{H}_{Cap}^N$$

and reduce (6.43c) to

$$\mathbf{H}_{Cap}^{N+1} = \frac{1}{\mu}\mathbf{H}_{Cap}^N - \beta\mathbf{z}\mathbf{F}_{N+1}^H(\omega)\mathbf{H}_{Cap}^N \qquad (6.43d)$$

It is shown in [11] that (6.43d) yields a stable estimate of the Capon filter.

§6.4 Beamformation with Coded Signals

In modern communication systems coded signals are used for transmitting information, which consists of a bit stream of ones and zeros. The bits are suitably coded into identifiable waveforms; for example '1' may be coded into a sinusoid of frequency f_1 and '0' is coded into another sinusoid of frequency f_0 $(f_1 \neq f_0)$ as in FSK (frequency shift keying) modulation. The bits thus coded after mixing with a carrier are sequentially transmitted. Since the physical channel is a shielded cable (including optical cable) there is much less cross channel interference. But, in the radio communication scenario, both transmitter and receiver are in the open space. Naturally a sensor will receive signals from more than one source. It is therefore of great interest to minimize this co-channel interference, a problem unique to radio communication. To overcome the problem of co-channel interference modern cellular radio communication has been devised. A user needs to communicate to the nearest base station which in turn is connected to a central exchange. Thus, it is possible to communicate with a distant user without having to radiate a lot of power causing a drain on the battery (in case of a mobile transmitter) and creating interference to other users. Even this system seems to fail when many users in the same cell are trying to reach the base station. To further mitigate the problem of co-channel interference it is proposed to use a highly directional antenna at the base station. This would enable the base station to separate the users having different bearings and reduce the co-channel interference among them. In the urban environment there is the additional problem of multipath propagation due to scattering. Since the multipaths are likely to arrive from different directions, the use of a directional array, it is hoped, will help to alleviate the problem of fading, loss of bits, etc. However, as this is still a research problem we shall not discuss this aspect.

The essential step in beamformation is estimation of the direction vector of each source. Given the direction vector a beam may be formed in that

direction using one of the beamforming methods described in §6.1 and §6.2. When there are a large number of sources (users), ordinarily we need to have a large sensor array (more sensors than the number of users) for the estimation of direction vectors. The advantage of using coded signals is that this limitation no longer exists. We can have more number of users than the number of sensors.

6.4.1 <u>Co-channel Interference</u>: We have derived in §5.2 a postcorrelation covariance matrix (see (5.68)) and expressed the same in a form (5.67) suitable for application of the subspace algorithm. The direction vectors to all users may be estimated provided the array has more sensors than the number of users. When this is not satisfied, the presence of users in excess of the number of sensors will only introduce interference, known as co-channel interference. We shall show how by employing extra information, which is available but not used, we can overcome the problem of co-channel interference.

At the base station we will assume an array of sensors (EM dipoles). The preferred array shape is a circular array (UCA) with uniform response in all directions as the users are likely to be all around the base station. A linear array (ULA) may also be used particularly when most users are on the broadside where the array has the best possible response.

The postcorrelation covariance matrix is reproduced here for convenience.

$$\mathbf{C}_{z_0 z_0} = p_0^2 \mathbf{a}(\theta_0)\mathbf{a}(\theta_0)^H + \frac{2}{3L}\sum_{k=1}^{Q} p_k^2 \mathbf{a}(\theta_k)\mathbf{a}(\theta_k)^H + \frac{\sigma_\eta^2}{L}\mathbf{I} \quad (5.68)$$

The first term on the right hand side is of interest as we would like to estimate the direction vector of the user of interest. The second term represents the co-channel interference from all other sources. Notice that this term will be small for large L.

6.4.2 <u>Estimation of All Direction Vectors</u>: We like to estimate the direction vectors of all users in the same cell. For this we shall compute the postcorrelation covariance matrices for all users. Thus, we will have Q equations of the type given by (5.66)

$$\mathbf{C}_{g_k g_k} = p_k^2 \mathbf{a}(\theta_k)\mathbf{a}(\theta_k)^H + \frac{2}{3L}\sum_{\substack{i=0 \\ k \neq i}}^{Q} p_i^2 \mathbf{a}(\theta_i)\mathbf{a}(\theta_i)^H + \frac{\sigma_\eta^2}{L}\mathbf{I} \quad (6.44)$$

$$k = 0, 1, \ldots, Q-1$$

We have in (6.44) Q matrix equations and Q matrix unknowns $\mathbf{a}(\theta_k)\mathbf{a}(\theta_k)^H$, k=0, ... Q-1. We shall express (6.44) in a matrix form

$$
\begin{bmatrix} \mathbf{C}_{z_0 z_0} \\ \mathbf{C}_{z_1 z_1} \\ . \\ . \\ . \\ \mathbf{C}_{z_{Q-1} z_{Q-1}} \end{bmatrix} = \begin{bmatrix} 1 & \frac{2}{3}L & \cdots & \frac{2}{3}L \\ \frac{2}{3}L & 1 & \cdots & \frac{2}{3}L \\ . & & \cdots & . \\ . & & \cdots & . \\ . & & \cdots & . \\ \frac{2}{3}L & \frac{2}{3}L & \cdots & 1 \end{bmatrix} \begin{bmatrix} p_0^2 \mathbf{a}(\theta_0)\mathbf{a}(\theta_0)^H \\ p_1^2 \mathbf{a}(\theta_1)\mathbf{a}(\theta_1)^H \\ \cdots \\ \cdots \\ \cdots \\ p_{Q-1}^2 \mathbf{a}(\theta_{Q-1})\mathbf{a}(\theta_{Q-1})^H \end{bmatrix} + \frac{\sigma_\eta^2}{T_s} \begin{bmatrix} \mathbf{I} \\ \mathbf{I} \\ . \\ . \\ . \\ \mathbf{I} \end{bmatrix}
$$

$$\qquad Q\times 1 \qquad\qquad\qquad Q\times Q \qquad\qquad\qquad\qquad Q\times 1 \qquad\qquad\qquad\qquad Q\times 1$$

$$(6.45\text{a})$$

In a compact form

$$\mathbf{C} = \pounds\Theta + \frac{\sigma_\eta^2}{T_s}\mathbf{II} \tag{6.45b}$$

Multiplying by $\pounds^{-1}$ on both sides of (6.45b) we can express it as

$$\Theta = \pounds^{-1}\mathbf{C} + \frac{\sigma_\eta^2}{T_s}\pounds^{-1}\mathbf{II} \tag{6.46}$$

The error term in (6.46) may be expressed as a product of a diagonal matrix and a column unit matrix. The elements of the diagonal matrix are equal to row sums of $\pounds^{-1}$. Thus, the noise covariance matrix in the estimated direction matrix remains diagonal. The variance of the noise may be estimated from the eigenvalues of the direction matrix. The power, transmitted by each user, that is, p_i^2, can also be estimated from the largest eigenvalue of the direction matrix.

6.4.3 <u>Simulation Results</u>: The estimated direction vector of a user is compared with the known direction vector. A dot product between the two vectors is computed as a measure of similarity,

$$\varepsilon_l = \frac{\hat{\mathbf{a}}^H(\theta_l)\mathbf{a}(\theta_l)}{\left|\hat{\mathbf{a}}^H(\theta_l)\right\|\mathbf{a}(\theta_l)\right|}$$

Note that $0 \le |\varepsilon_l| \le 1$, the lower limit represents the worst estimate and the upper limit represents the best estimate. We have computed the mean and the

snr	Mean	Variance
-20 dB	0.6029	0.1059
-10 dB	0.9746	3.8480e-04
-5 dB	0.9945	8.8470e-06
0 dB	0.9979	8.6691e-07
No noise	0.9983	1.0934e-06

Table 6.1: Error in the estimation of direction vector for different snrs. 10 sensor ULA, randomly distributed 10 users, 100 snapshots, code length = 63 chips.

Number of users	Mean	Variance
10	0.9983	1.0934e-06
20	0.9949	6.7374e-06
40	0.9930	1.5846e-05
50	0.9909	2.3899e-05
60	0.9878	3.6955e-05

Table 6.2: Error in the estimation of direction vector with increasing number of users. 10 sensor ULA, randomly distributed users, 100 snapshots, no noise, and code length = 63 chips.

variance of ε_l as a measure of quality of estimate. The results are shown in table 6.1. The postcorrelation matrix approach for the estimation of direction vectors as described here is not limited by the requirement that the number of sensors must be greater than the number of users as in the approach described in [12] using both pre- and postcorrelation matrices. In fact to verify this claim, above simulation was repeated with no noise for a different number of users. The results are shown in table 6.2. There is, however, a slight decrease in the quality of estimate.

6.4.4 <u>Beamforming with Cyclostationary Signals</u>: We consider P sources emitting cyclostationary signals with different but known cyclic frequencies. We like to find a set of weight coefficients which forms a beam in the direction of a source having a specified cyclic frequency. The array output $\mathbf{f}(t)$ is governed by the signal model given in (2.18). The noise is stationary but not necessarily white. We use the frequency shifted version of $\mathbf{f}(t)$ defined in (5.70). Let $\mathbf{w}_{+}$ and $\mathbf{w}_{-}$ be the beamforming weight coefficient vectors for $\mathbf{f}_{+}(t)$ and $\mathbf{f}_{-}(t)$, respectively. The cross-correlation of the outputs is given by

$$\rho = E\left\{\mathbf{w}_-^{\,H}\mathbf{f}_-(t)\mathbf{f}_+^{\,H}(t)\mathbf{w}_+\right\}$$
$$= \mathbf{w}_-^{\,H}\mathbf{c}_f^{\alpha}\mathbf{w}_+ \tag{6.47}$$

where $\mathbf{c}_f^{\alpha}$ is the cyclic covariance matrix defined in (5.71a) for zero lag. The filter coefficients are chosen to maximize the cross-correlation (6.47) or its magnitude square. Further, we require that $\mathbf{w}_+$ and $\mathbf{w}_-$ are unit norm vectors.

$$\left|\mathbf{w}_-^{\,H}\mathbf{c}_f^{\alpha}(\tau)\mathbf{w}_+\right|^2 = \max, \quad \mathbf{w}_+^{\,H}\mathbf{w}_+ = 1, \quad \mathbf{w}_-^{\,H}\mathbf{w}_- = 1 \tag{6.48a}$$

The solution of (6.48a) is given by left and right singular vectors corresponding to the largest singular value of $\mathbf{c}_f^{\alpha}(\tau)$. Also $\mathbf{w}_-$ and $\mathbf{w}_+$ are, respectively, the eigenvectors corresponding to the largest eigenvalues of $\mathbf{c}_f^{\alpha}(\tau)\mathbf{c}_f^{\alpha}(\tau)^H$ and $\mathbf{c}_f^{\alpha}(\tau)^H\mathbf{c}_f^{\alpha}(\tau)$ [5].The cyclic covariance matrix appearing in (6.48) is in practice replaced by the cyclic autocorrelation function defined in terms of time average,

$$\hat{\mathbf{c}}_f^{\alpha} = \frac{1}{T}\sum_{-\frac{T}{2}}^{\frac{T}{2}}\mathbf{f}_-(t)\mathbf{f}_+^{\,H}(t) \tag{6.48b}$$

When the carrier frequencies of different sources are sufficiently apart and the signal duration T is large the cyclic autocorrelation matrix given by (6.48) approaches the cyclic covariance matrix of a single source which is a rank one matrix as shown in (5.71b) whose left and right singular vectors corresponding to the largest singular value are equal to $\dfrac{1}{\sqrt{M}}\mathbf{a}_0$ (for ULA). Thus, we have an interesting result

$$\mathbf{w}_+ = \mathbf{w}_- = \frac{1}{\sqrt{M}}\mathbf{a}_0 \tag{6.48c}$$

§6.5 Multipath Channel:

In many real life situations a signal (also interference) may reach an array of sensors via more than one path. The resulting signal will be an overlap of many copies of the actual signal emitted by the source. The different copies of the signal may differ in arrival time and amplitude but remain highly correlated unless they have traveled along widely different paths through a random
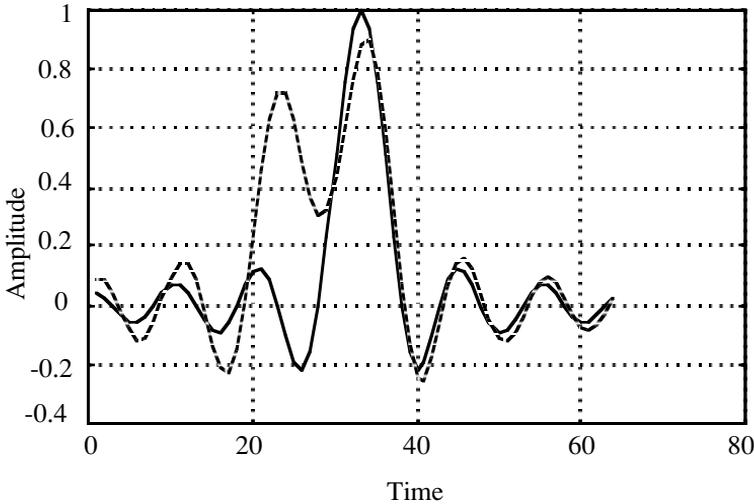
Figure 6.9: A source emits a waveform (sinc function) shown by solid curve. The signal reaches a sensor via four paths as in equation (6.49). The sensor output is shown by dashed curve.

medium. For the time being we shall not consider such a situation. We shall assume that all copies of the signal remain correlated. Even in this simplified model of multipath propagation the waveform received by a sensor may have no resemblance with the signal emitted by the source. To appreciate the kind of deterioration a signal may suffer let us consider a source emitting a signal described by sinc function ($= \dfrac{\sin(x)}{x}$) and a multipath channel having four paths. The relative amplitudes and delays are assumed to be random numbers as given below:

$$f(t) = \begin{cases} f_0(t) - 0.0562 f_0(t - 7.0119) \\ +0.5135 f_0(t + 9.1032) + 0.3967 f_0(t + 2.6245) \end{cases} \quad (6.49)$$

where the signal emitted by the source is a sinc function, $f_0(t) = \sin c(0.2\pi(t - 32))$. The sum waveform is shown in fig. 6.9. Generally, the broadband signals are more distorted than the narrowband ones.

In chapter 4, in the context of source localization in a bounded medium, we have modeled the array signal as

$$\mathbf{F}(\omega) = \mathbf{A}\mathbf{w}F_0(\omega) + \mathbf{N}(\omega) \quad (6.50)$$

where $\mathbf{A} = [\mathbf{a}_0, \mathbf{a}_1, \dots \mathbf{a}_{q-1}]$ and $\mathbf{a}_0, \mathbf{a}_1, \dots \mathbf{a}_{q-1}$ are direction vectors pointing to q paths ($\mathbf{a}_0$ is assumed to point to the direct path), and $\mathbf{w}$ stands for complex weight vector, $\mathbf{w} = [w_0, w_1, \dots, w_{q-1}]$ where $w_0, w_1, \dots, w_{q-1}$ are complex weighting coefficients applied to the temporal Fourier coefficients of the array signal. They represent attenuation and phase change due to propagation delays. If the channel is well characterized both $\mathbf{A}$ and $\mathbf{w}$ can be estimated as a part of the source localization step, which, we will assume, has been carried out prior to waveform estimation. Thus, as a first step, assuming that $\mathbf{A}\mathbf{w}$ is known let us explore how well $f_0(t)$ can be recovered in the presence of background noise and interference from other sources.

6.5.1 <u>Known Channel</u>: A few examples of simple channels, whose characteristics are known or can be modeled reasonably well, are sketched in fig. 6.10. The channel shown in fig. 6.10a is a good approximation to surface reflected radar signal, surface reflected seismic signal, etc. The reflection coefficient of the surface is the unknown quantity but the direction of arrival of the direct wavefront as well as the reflected wavefront can be easily computed given the source location. For example, let the source be located at (l, h_s) where l is range and h_s is depth from the surface. The direction of the arrival of the direct wavefront is given by

$$\tan\theta_0 = \frac{h_s - h_a}{l} \qquad \text{and} \qquad \tan\theta_1 = -\frac{h_s + h_a}{l} \qquad (6.51)$$

and h_a is the depth of the array (midpoint) from the surface. In deriving (6.51) we have used the method of images as outlined in chapter 1. For this channel the direction vectors and weight vector, respectively, are $\mathbf{A} = [\mathbf{a}_0, \mathbf{a}_1]$ where

$$\mathbf{a}_0 = col\left\{1, e^{-j\frac{2\pi}{\lambda}d\sin(\theta_0)}, e^{-j\frac{2\pi}{\lambda}2d\sin(\theta_0)}, \dots, e^{-j\frac{2\pi}{\lambda}(M-1)d\sin(\theta_0)}\right\}$$

$$\mathbf{a}_1 = col\left\{1, e^{-j\frac{2\pi}{\lambda}d\sin(\theta_1)}, e^{-j\frac{2\pi}{\lambda}2d\sin(\theta_1)}, \dots, e^{-j\frac{2\pi}{\lambda}(M-1)d\sin(\theta_1)}\right\}$$

$$\mathbf{w} = col\{1, r\}$$

where r is the coefficient of reflection which we shall assume for simplicity independent of the angle of incidence. We can now express $\mathbf{A}\mathbf{w}$ vector in terms
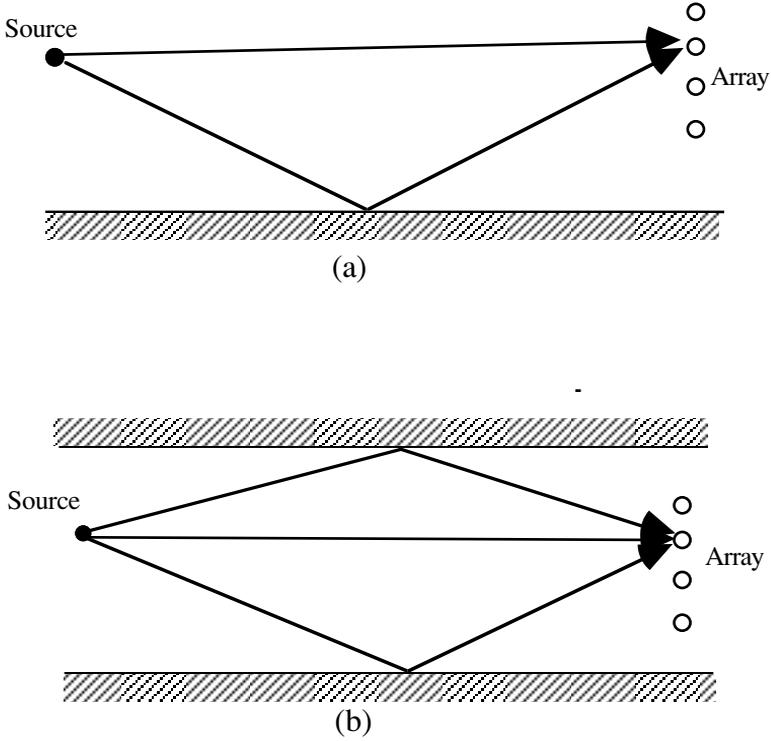
Figure 6.10: Two types of simple channels for which the $\mathbf{Aw}$ vector may be estimated from a knowledge of source location and channel characteristics.

of the direction vectors of the direct wavefront from the source and the one from the image of the source.

$$\mathbf{Aw} = \begin{bmatrix} \mathbf{a}_0 + r\mathbf{a}_1 \end{bmatrix} \tag{6.52}$$

First, let us try the linear least squares estimate (LLSE) of $F_0(\omega)$. We shall transform (6.11) into frequency domain, leading to

$$\begin{aligned}
\hat{F}_0(\omega) &= \left[\mathbf{w}^H \mathbf{A}^H \mathbf{Aw} \right]^{-1} \mathbf{w}^H \mathbf{A}^H \mathbf{F}(\omega) \\
&= F_0(\omega) + \left[\mathbf{w}^H \mathbf{A}^H \mathbf{Aw} \right]^{-1} \mathbf{w}^H \mathbf{A}^H \mathbf{N}(\omega)
\end{aligned} \tag{6.53a}$$

The quantity inside the square brackets in (6.53a) may be evaluated using (6.52). It is given by

$$\mathbf{w}^H \mathbf{A}^H \mathbf{A} \mathbf{w} = \mathbf{a}_0^H \mathbf{a}_0 + r^H \mathbf{a}_1^H \mathbf{a}_0 + r \mathbf{a}_0^H \mathbf{a}_1 + |r|^2 \mathbf{a}_1^H \mathbf{a}_1$$

$$= 2M \left(1 + \frac{r^H \mathbf{a}_1^H \mathbf{a}_0 + r \mathbf{a}_0^H \mathbf{a}_1}{2M} \right) \qquad (6.53b)$$

The variance of the noise power in the filtered array output is given by

$$\sigma_{\hat{\eta}}^2 = \frac{\sigma_\eta^2}{\left[\mathbf{w}^H \mathbf{A}^H \mathbf{A} \mathbf{w} \right]}$$

$$= \frac{\sigma_\eta^2}{M \left(1 + |r|^2 + \dfrac{r^H \mathbf{a}_1^H \mathbf{a}_0 + r \mathbf{a}_0^H \mathbf{a}_1}{M} \right)} \qquad (6.54)$$

Compare (6.54) with (6.14), which was derived for two uncorrelated sources. In the present case both sources are correlated (the second source is an image of the primary source). The variance of the noise is reduced by a factor four, when r=1 and $\mathbf{a}_1 = \mathbf{a}_0$. The multipath propagation has indeed helped to improve signal estimation.

Next, we shall try the Capon filter to estimate the waveform in the presence of interference, another source at known location, and the usual background white noise. The position of the sources and the receiving array are shown in . We shall assume that the sources are uncorrelated. The directions of arrival of the direct and the reflected wavefronts are given by

$$\tan \theta_{00} = \frac{h_s - h_a}{l}, \qquad \tan \theta_{10} = -\frac{h_s + h_a}{l}$$

$$\tan \theta_{01} = \frac{h_s + \Delta h - h_a}{l}, \qquad \tan \theta_{11} = \frac{h_s + \Delta h + h_a}{l}$$

For simplicity we assume that the coefficient of reflection r is the same for both sources. The direction vectors are given by

$$\mathbf{A}_0 \mathbf{w} = \left[\mathbf{a}_0 + r \mathbf{a}_1 \right]$$

$$\mathbf{A}_1 \mathbf{w} = \left[\mathbf{a}_{01} + r \mathbf{a}_{11} \right]$$
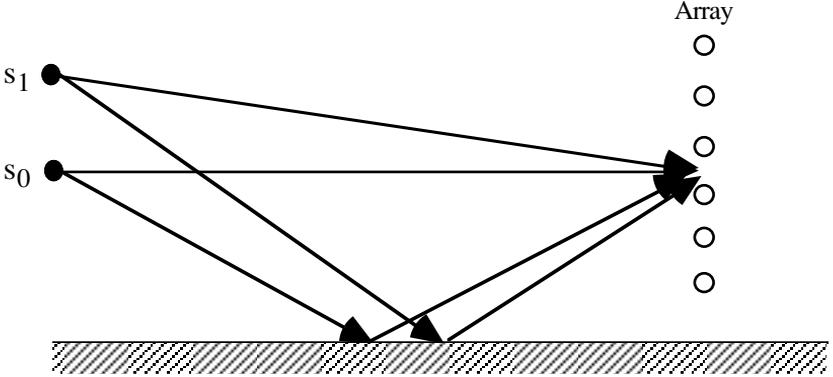
Figure 6.11: Two sources at the same range but at different heights. A vertical array of sensors is assumed.

The array output is modeled as

$$\mathbf{F}(\omega) = \mathbf{A}_0 \mathbf{w} F_0(\omega) + \mathbf{A}_1 \mathbf{w} F_1(\omega) + \mathbf{N}(\omega) \qquad (6.55)$$

To compute the Capon filter given by (6.26) we need the spectral matrix of the array output. Since the sources are assumed uncorrelated and the background noise is white, the spectral matrix may expressed as

$$\mathbf{S}(\omega) = \mathbf{A}_0 \mathbf{w}\mathbf{w}^H \mathbf{A}_0^H S_0(\omega) + \mathbf{A}_1 \mathbf{w}\mathbf{w}^H \mathbf{A}_1^H S_1(\omega) + \sigma_\eta^2 \mathbf{I} \qquad (6.56)$$

The Capon filter is specified by (6.26). To estimate the waveform emitted by the zeroth source the required filter is given by

$$\mathbf{H}_{cap} = \frac{\mathbf{S}_f^{-1}\mathbf{A}_0\mathbf{w}}{\mathbf{w}^H \mathbf{A}_0^H(\omega,\theta)\mathbf{S}_f^{-1}\mathbf{A}_0\mathbf{w}} \qquad (6.57)$$

Applying the filter given in (6.57) to the array output we obtain

$$\mathbf{H}_{cap}^H \mathbf{F}(\omega) = F_0(\omega) + \frac{\mathbf{w}^H \mathbf{A}_0^H \mathbf{S}_f^{-1}\mathbf{A}_1\mathbf{w}}{\mathbf{w}^H \mathbf{A}_0^H \mathbf{S}_f^{-1}\mathbf{A}_0\mathbf{w}} F_1(\omega) + \frac{\mathbf{w}^H \mathbf{A}_0^H \mathbf{S}_f^{-1}\mathbf{N}(\omega)}{\mathbf{w}^H \mathbf{A}_0^H \mathbf{S}_f^{-1}\mathbf{A}_0\mathbf{w}} \qquad (6.58)$$
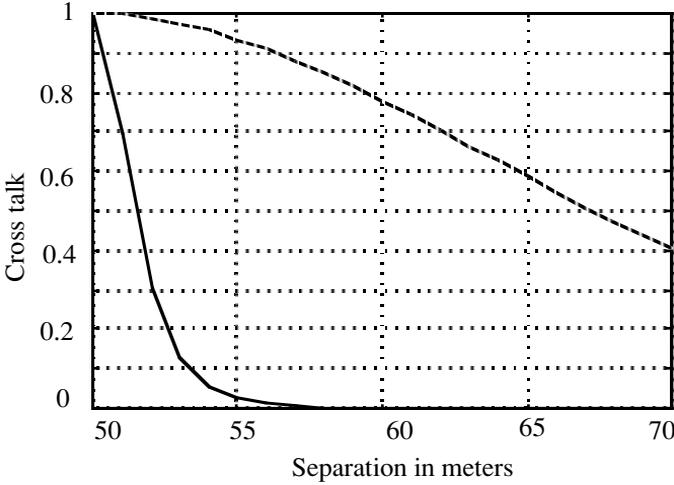
Figure 6.12: Cross talk as a function of separation between two sources. The range is the same for both sources. Solid curve: range=1000m and dashed curve=5000m. The first source is 50 meters above the surface.

While the signal from the zeroth source has been fully extracted there are two terms in (6.58) interfering with the desired signal. Here we shall consider the term representing the interference from the source #1. The second term, that is, the noise term, will not be considered as it follows the same approach used previously in connection with the single source case. The interference due to second source will be measured in terms of cross talk as defined previously. In the present case the cross talk is given by

$$cross\ talk = \frac{(\mathbf{w}^H \mathbf{A}_1^H \mathbf{S}_f^{-1} \mathbf{A}_0 \mathbf{w})(\mathbf{w}^H \mathbf{A}_0^H \mathbf{S}_f^{-1} \mathbf{A}_1 \mathbf{w})}{\left| \mathbf{w}^H \mathbf{A}_0^H \mathbf{S}_f^{-1} \mathbf{A}_0 \mathbf{w} \right|^2} \qquad (6.59)$$

The cross talk as a function of source separation for two uncorrelated sources is plotted in fig. 6.12. A 16 sensor vertical ULA was assumed. Notice that for a short range the second source has little influence on the first source. But this influence grows rapidly as the range increases.

6.5.2 Partially Known Channel: In many real life problems the channel characteristics are never fully known as it is impossible to measure the micro level variations causing path length variations on the order of a fraction of wavelength. Such variations are known to affect the performance of source localization algorithms, particularly those belonging to a high resolution class [13]. On the other hand we may have fairly good knowledge about the general

features of a channel but not enough to characterize at micro level. One such example is a shallow water sea where the sea bottom variability is high. While the general geometry of the shallow water channel, broad undulations of the sea bottom, water temperature variations, etc. are available from the actual measurements, the factors which affect the performance of waveform estimation such as the details of sea bottom, particularly sediment distribution, and sea surface undulation are unknown. In chapter 5 we have shown how source localization can be carried out with a full knowledge of a shallow water channel. Here we shall describe an approach capable of estimating the source location as well as signal waveform emitted by the source given only partial knowledge.

Consider a single source in a shallow water channel and a vertical array of sensors. We have shown in chapter 5 that the eigenvector corresponding to the largest eigenvalue is related to $\mathbf{Aw}$,

$$\mathbf{E}_s = \frac{\mathbf{Aw}}{|\mathbf{Aw}|}, \quad |\mathbf{Aw}| \neq 0 \qquad (6.60)$$

In (6.60) the weight vector $\mathbf{w}$ is dependent on the channel parameters and the columns of $\mathbf{A}$ matrix on the source position. Notice that the weight vector $\mathbf{w}$ in (6.60) occurs linearly while the source parameters in $\mathbf{A}$ occur nonlinearly. A least squares method of estimating the nonlinear parameters by first eliminating the linear parameters followed by minimizing the norm of the error vector was first suggested by Guttman et al.[14] and applied to a signal processing problem by Tuft and Kumaresan [15]. We shall exploit here this approach. In (6.60) we assume that the source position is approximately known and write the equation in terms of the unknown $\mathbf{w}$ vector,

$$\mathbf{E}_s = \mathbf{A}\frac{\mathbf{w}}{|\mathbf{Aw}|} = \mathbf{A}\tilde{\mathbf{w}} \qquad (6.61)$$

Let $\mathbf{A}^{\#}$ be the pseudoinverse of $\mathbf{A}$. The least squares estimate of $\tilde{\mathbf{w}}$ will be given by

$$\hat{\tilde{\mathbf{w}}} = \mathbf{A}^{\#}\mathbf{E}_s$$

Substitute back into (6.60) and obtain an estimate of $\mathbf{E}_s$. The mean square error is given by

$$\|error\|_2 = \left\|(\mathbf{I} - \mathbf{A}\mathbf{A}^{\#})\mathbf{E}_s\right\|_2 = \left\|\mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{E}_s\right\|_2$$

where $\mathbf{P}_{\mathbf{A}}^{\perp}$ is the orthogonal projection complement of matrix $\mathbf{A}$. The mean square error is now minimized with respect to the source location parameters. We define a parametric spectrum as,

$$S(R_0, Z_s) = \frac{1}{\mathbf{E}_s^{H} \mathbf{P}_{\mathbf{A}}^{\perp} \mathbf{E}_S} \tag{6.62}$$

When the source location is exactly known, $\mathbf{P}_A^{\perp} = (\mathbf{I} - \mathbf{A}\mathbf{A}^{\#}) = \mathbf{E}_\eta \mathbf{E}_\eta^{\#}$, the parametric spectrum defined in (6.62) turns out to be same as in (4).

In practical implementation the parametric spectrum is computed over a dense grid in the range-depth space. At each grid point the projection matrix $\mathbf{P}_{\mathbf{A}}^{\perp}$ is computed and (6.62) is evaluated. Thus, in the proposed method, the subspace spanned by the columns of the $\mathbf{A}$ matrix (range space) is steered until $\mathbf{P}_{\mathbf{A}}^{\perp}$ coincides with $\mathbf{E}_\eta \mathbf{E}_\eta^{H}$, that is, when the parametric spectrum becomes very large (ideally infinite). After obtaining the source location parameters we use them to estimate the weight vector,

$$\hat{\tilde{\mathbf{w}}} = \mathbf{A}_{max}^{\#} \mathbf{E}_s \tag{6.63}$$

where $\mathbf{A}_{max}^{\#}$ is $\mathbf{A}^{\#}$ evaluated where the parametric spectrum is maximum. Thus, $\hat{\tilde{\mathbf{w}}}$ is the least mean square estimate of $\tilde{\mathbf{w}}$. An example of parametric spectrum is shown in fig. 6.13. A low power source (-10dB) is assumed at range 4600m and depth 50m in a Pekeris channel of depth 200 meters. A vertical ULA is placed at a depth of 70m. For the same channel the reflection coefficients were computed (6.63) from the eigenvector corresponding to the largest eigenvalue. The reflection coefficients are normalized with respect to $|\mathbf{Aw}|$, which may be obtained from the fact that the weighting coefficient corresponding to the direct path is by definition equal to one; hence its actual observed value must be equal to $|\mathbf{Aw}|$. In table 6.3 the estimated and the actual reflection coefficients for the first eight images out of twenty multipaths used in computation are listed for two different array lengths.

In computer simulation we have found that, for good results, the number of sensors has got to be many more than the lower limit given in [16]. The least mean square error in the estimated reflection coefficients for different number of sensors is shown in fig. 6.14. Here the number of significant images
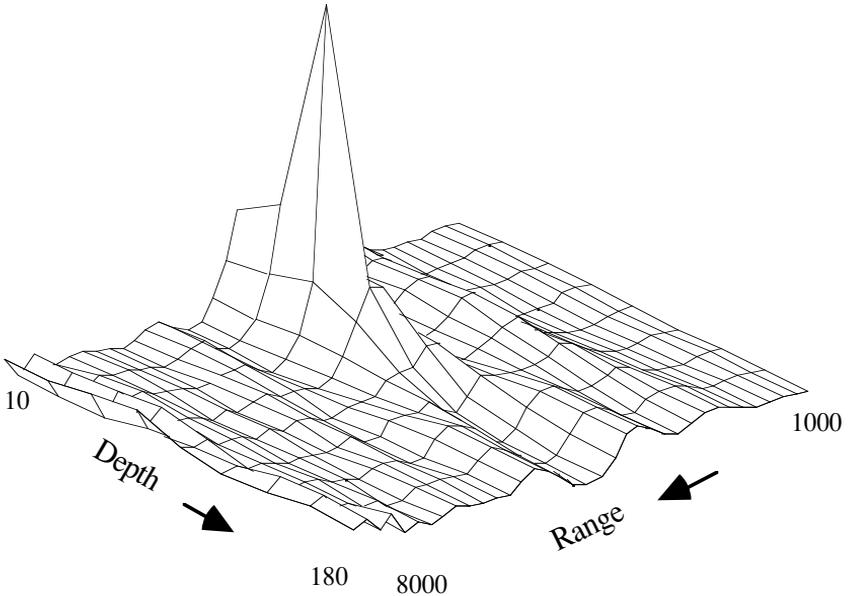
Figure 6.13: The parametric spectrum for a single source located at range 4600 m and depth 50 m. The source power is -10dB (relative to the background noise). Under water channel is 200m deep Pekeris channel (soft bottom, speed: 1600m/s and relative density: 2.0). A vertical ULA consisting of 40 sensors is placed at a depth of 70m from the surface. Range scan is from 1000 m to 8000 m in steps of 200 m. Depth scan is from 10 m to 180 m in steps of 20 m.

Image #	True reflection coefficients		Estimated reflection coefficients			
			M=40		M=60	
	Real	Imag	Real	Imag	Real	Imag
1	1.0	0.0	1.0	0.0	1.0	0.0
2	-.861	-.506	-.847	-.718	-.870	-.490
3	.467	.882	.612	.829	.416	.886
4	.995	-.064	1.075	-.260	.998	-.091
5	.981	.167	1.395	-.053	.925	.205
6	.905	.407	1.040	.251	.858	.459
7	.626	-.768	.705	-.783	.641	-.752
8	.107	.981	-.013	.964	.107	.986

Table 6.3: A comparison of estimated reflection coefficients (first eight coefficients) with true reflection coefficients, computed for a channel described in fig. 6.13. The results for M=40 and 60.
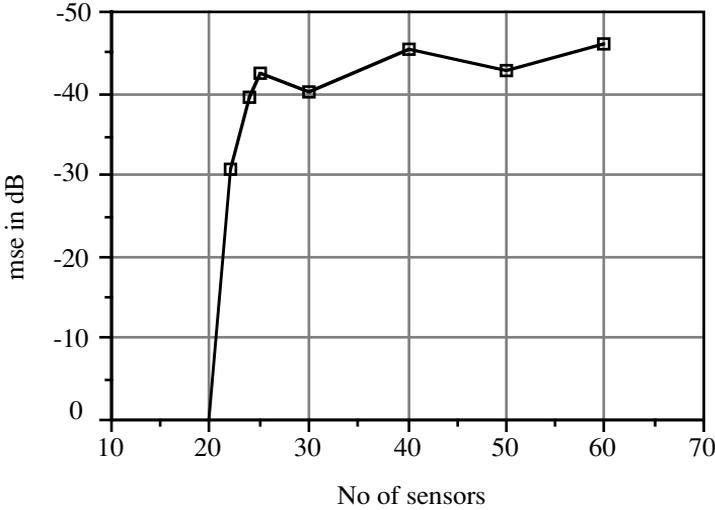
Figure 6.14: The role of array size on the mean square error in the estimated reflection coefficients is shown above. Twenty multipaths were assumed.

is twenty (P=20) and hence, according to the lower limit $(> 2P+ 2)$ the minimum number of sensors ought to be more than 42. We observe that an array of sixty sensors appears to be optimum.

In addition to unknown reflection coefficients we have background noise which is likely to be both spatially and temporally correlated. But, since it is uncorrelated with the signal, it occurs in a linear combination; as such it may be estimated using the approach used for the estimation of the reflection coefficients. Indeed, such an approach has been used by Boehme [17] for the estimation of the noise spectrum first using an approximate knowledge of the channel. In the next step, the previously estimated noise spectrum is used in the expression for the likelihood ratio which is then maximized with respect to the unknown channel parameters and the source position. The maximization of the likelihood ratio is, however, highly computation intensive [18].

§6.6 **Exercises**:
1. Consider P uncorrelated wavefronts incident on an M sensor ULA. Let the background noise be uncorrelated with all signals. Show that the Wiener filter in the frequency domain is given by

$$\mathbf{H}_W(\omega) = \mathbf{S}_f^{-1}(\omega)\mathbf{A}\mathbf{S}_0(\omega)$$

where $\mathbf{A}$ is a matrix whose columns are direction vectors of the wavefronts which are incident on the array.

2. Show that the Wiener filter (also Capon filter) can be written as

$$\mathbf{H}_W = \mathbf{v}_s \alpha_s^{-1} \mathbf{v}_s^H \mathbf{a}_0 S_0$$

where $\mathbf{a}_0$ is the direction vector of the wavefront from the first source and S_0 is signal power, (Hint: Use (3.38b) and the property given in (4.14b).) Such a filter is robust against calibration error (see p.342).

3. In §6.1 and §6.2 we have seen that in estimating a waveform in the presence of interference there is always some cross talk, that is, leakage of power from the interfering signal to the desired signal. It is possible to devise a set of filters, $H_m(\omega)$, $m = 0,1,\ldots,M-1$, which will null the interference without distorting, that is, $\sum_{m=0}^{M-1} H_m(\omega) = 1$, the desired signal (but with no noise). The array has been steered to receive the desired signal. The DOA of the desired signal and that of the interference are known. Show that

$$H_m(\omega) = \frac{M - e^{j\omega\tau_m} \sum_{i=0}^{M-1} e^{-j\omega\tau_i}}{M^2 - \left| \sum_{i=0}^{M-1} e^{-j\omega\tau_i} \right|^2}$$

where τ_i is the time delay of the interference at i^{th} sensor. The interference is nulled except when $\omega\tau_i$ is equal to an integer multiple of 2π [4].

References
1. B. Otterstein, R. Roy, and T. Kailath, Signal waveform estimation in sensor array processing, Asilomar Conf, pp. 787-791, 1989.
2. B. Friedlander and A.J. Weiss, Effects of model errors on waveform estimation using the MUSIC algorithm, IEEE Trans., SP-42, pp. 147-155, 1994.
3. R. T. Compton, Pointing accuracy and dynamic range in steered beam adaptive array, IEEE Trans., AES-16, pp. 280-287, 1980.
4. M.T. Hanna and M. Simaan, Absolutely optimum array filters for sensor array, IEEE Trans., ASSP-33, pp. 1380-1386, 1985.
5. Q. Wu and K. Wong, Blind adaptive beamforming for cyclostationary signals, IEEE Trans., SP-44, pp. 2757-2767, 1996.

6. S.P. Applebaum, Adaptive array, IEEE Trans., AP-24, pp. 585-598, 1976.

7. L. C. Godara, Error analysis of the optimal antennal array processors, IEEE Trans, AES-22, pp. 395-409, 1986.

8. B. Widrow and S. D. Stearms, *Adaptive Signal Processing*, Prentice-Hall, Inc, Englewood Cliffs, NJ, 1985.

9. J. R. Treichler, C. R. Johnson, Jr., and M. G. Larimore, Theory and design of adaptive filters, Wiley Interscience, New York, 1987.

10. J. W. Kim and C. K. Un, A robust adaptive array based on signal subspace approach, IEEE Trans., SP-41, pp. 3166-3171, 1993.

11. R. Schreiber, Implementation of adaptive array algorithms, IEEE Trans., ASSP-34, pp. 1038-1045, 1986.

12. A. F. Naguib, A. Paulraj, and T. Kailath, Capacity improvement with base-station antenna array in cellular CDMA, IEEE Trans. on Vehicular Technology, vol. 43, pp. 691-698, 1994.

13. J. R. Daugherty and J. F. Lynch, Surface wave, internal wave, and source motion effects on matched field processing in a shallow water waveguide, J. Acoust. Soc. Am., vol. 87, pp. 2503-2526, 1990.

14. I. Guttman, V. Peereyra, and H. D. Scholnik, Least squares estimation for a class of nonlinear models, Technometrics, vol. 15, pp. 209-218,1973.

15.D. W. Tuft and R. Kumaresan, Improved spectral resolution II, Proc. ICASSP, Denver, CO, 1980, pp. 392-397.

16. P. S. Naidu and T. Ganesan, Source localization in a partially known shallow water, J. Acoust. Soc. of Am., vol 98, Pt. 1, pp. 2554-2559, 1995.

17. J. F. Boehme, Array Processing *in* Advances in Spectrum Analysis and Array Processing, S. Haykin (Ed), pp. 1-63, Prentice-Hall, Englewood Cliffs, NJ, 1991.

18. C. F. Mecklenbraeuker, A. B. Gershman, and J. F. Boehme, ML estimation of environmental parameters in shallow ocean using unknown broadband sources, Proc. IEEE ICNNSP-95, Nanjing, China, 1995.

Chapter Seven

Tomographic Imaging

As a wavefield propagates through a medium it is subjected to time delays and loss of power. The wavefield is reflected from interfaces separating media of different impedances, and is scattered by inhomogeneities present in the medium. By observing these effects it is possible to study the characteristics of the medium through which the field has propagated. Seismic exploration, on which depends the future discoveries of the petroleum deposits, exploits these effects of propagation to produce a detailed image of the subsurface geologic structure which may be conducive to accumulation of the hydrocarbon deposits. Likewise the ultrasonic imaging used in medical diagnoses and in nondestructive testing also exploits the propagation effects of the wavefield. In this chapter we shall study these effects of propagation for the purpose of constructing a three dimensional image of the medium. Tomography refers to cross-sectional imaging of objects from either transmitted, reflected or diffracted wavefields. Accordingly, there are three different types of tomographic imaging methods. One or more effects of propagation such as accumulated attenuation, travel time, wavefield produced by diffraction or scattering are observed in all directions (360° for 3D imaging). The observations such as travel time delays or accumulated attenuation are inverted by solving a system of linear equations. Where the medium is a diffracting type, that is, the size of inhomogeneities is comparable to the wave length of the illuminating wavefield, the preferred approach is Fourier inversion. The subject of tomographic imaging is covered in the next four sections. In the last section we investigate how to estimate the shape of an object from its scattered field.

§7.1 Nondiffracting Radiation:

When the wavelength of illuminating radiation (e.g., x-rays, ultrasound) is much smaller than the dimensions of the inhomogeneities in the propagating medium the concept of ray propagation becomes useful. The rays may travel in straight line or along a curved line depending upon the average wave speed: a straight line when the average speed is constant or curved path when the average speed is spatially variable as in a layered medium. The local speed variation is assumed to have a negligible effect on the ray paths. The propagation has two effects on the wave, namely, wave attenuation and delay, both of which are of great significance from the point of tomographic imaging. Typically in x-ray tomography, wave attenuation is used and in ultrasound tomography and also in seismic tomography total time delay is used.

7.1.1 <u>Absorption</u>: Consider for example an object cross section represented by a function f(x,y). A straight line ray intersects the object and suffers a certain amount of attenuation depending upon the length of the ray path lying inside
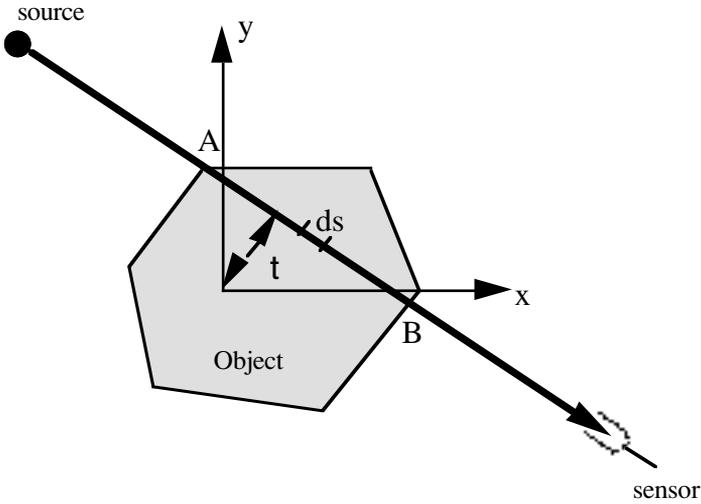
Figure 7:1. A ray passing through an absorbing object suffers an attenuation proportional to the integral over the path lying within the object.

the object (see fig. 7.1). Let N_{in} be the number of photons incident on the object at point A and N_d be the number of photons coming out at point B within the time interval of measurement. N_{in} and N_d are related as below
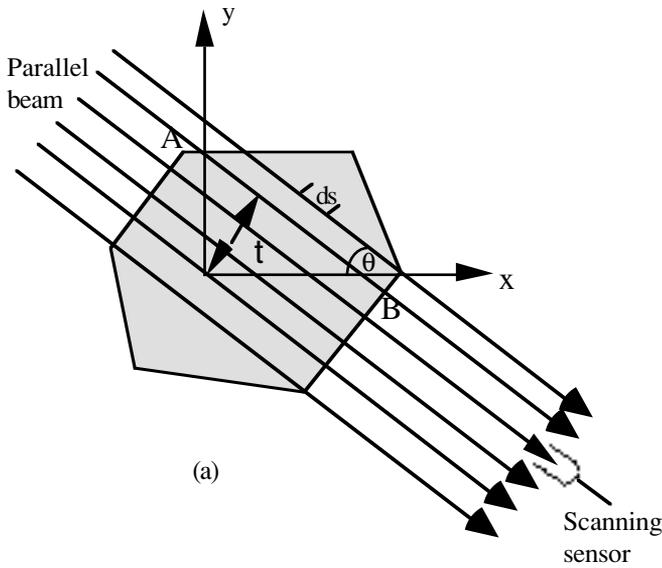
$$N_d = N_{in} \exp[-\int_A^B f(x,y)ds] \qquad (7.1a)$$

Define attenuation as negative of $\log \dfrac{N_d}{N_{in}}$, which turns out to be equal to the integral of $f(x,y)$ along path AB,

$$p = -\log_e \frac{N_d}{N_{in}} = \int_A^B f(x,y)ds \qquad (7.1b)$$

where p is also known as projection. The equation of line AB may be expressed as

$$x\cos(\theta) + y\sin(\theta) = t$$
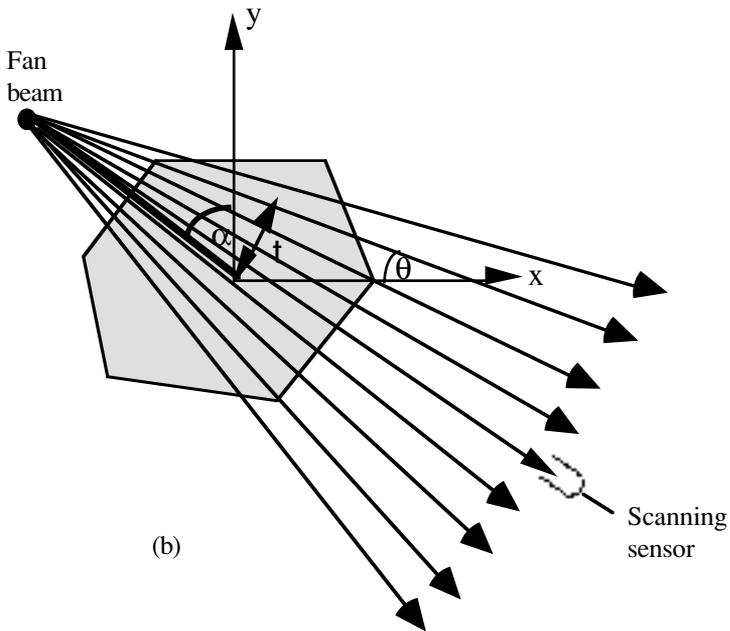
Figure 7.2: (a) A parallel projection is obtained by illuminating an object with a parallel beam. (b) A fan projection is obtained by illuminating an object with a fan beam generated by a point source at a radial distance d and angular distance $(90+\alpha)^{\circ}$.

where θ is slope of the line and t stands for perpendicular distance to the line from the origin (see fig. 7.1). Equation (7.1b) may be expressed as follows:

$$p_\theta(t) = -\log_e \frac{N_d}{N_{in}} = \int_A^B f(x,y)ds$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\delta(x\cos(\theta) + y\sin(\theta) - t)dxdy$$

(7.2)

$p_\theta(t)$ is known as Radon transform (see page 21). For a fixed θ and variable t, we obtain $p_\theta(t)$, a continuous function of t, known as parallel projection which may be generated by illuminating an object with a parallel beam and scanning the output with a receiver (see fig. 7.2a)

Taking Fourier transform on both sides of (7.2) we obtain

$$P_\theta(\omega)$$

$$= \int_{-\infty}^{\infty} p_\theta(t)\exp(-j\omega t)dt$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)dxdy \int_{-\infty}^{\infty} \delta(x\cos(\theta) + y\sin(\theta) - t)\exp(-j\omega t)dt$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\exp(j\omega(x\cos(\theta) + y\sin(\theta)))dxdy$$

(7.3)

$$= F(\omega\cos(\theta), \omega\sin(\theta))$$

where $F(\omega\cos(\theta), \omega\sin(\theta))$ is the 2D Fourier transform of the object function f(x,y) evaluated on $u = \omega\cos(\theta)$ and $v = \omega\sin(\theta)$. The Fourier transform of a projection of an object function taken at an angle θ is equal to the slice of 2D Fourier transform of the object function evaluated on a radial line at an angle θ as shown in fig. 7.3. This is known as Fourier slice theorem [1, p. 372]. By changing θ continuously the 2D Fourier transform of the object function is evaluated over a series of radial lines as shown in fig. 7.3.

A fan beam is more appropriate, as a point source at finite distance emits a spherical beam which, when used to illuminate a finite target, may be considered as a conical beam in three dimensions or a fan beam in two dimensions. Both slope and the perpendicular distance of a ray depend upon the ray angle measured with respect to radius vector of the source. An exact

reconstruction algorithm for a fan beam has been worked out after some tedious geometrical simplifications [1], but a simpler approach, where a series of fan beams covering 360° may be regrouped into a series of parallel beams, is of greater interest from the array signal processing point of view. This approach is outlined below.

7.1.2 <u>Filtered Backprojection Algorithm</u>: The reconstruction algorithm consists of inverse Fourier transformation of a radially sampled object Fourier transform which in turn is obtained from parallel projection data. Let us express the 2D inverse Fourier transform in polar coordinates.

$$f(x,y) = \frac{1}{4\pi^2} \int\limits_{-\infty}^{+\infty}\!\!\int F(u,v)e^{j(ux+vy)}\,dudv$$

$$= \frac{1}{4\pi^2} \int\limits_{0}^{\infty} sds \int\limits_{0}^{2\pi} F(s,\theta)e^{js(x\cos\theta+y\sin\theta)}\,d\theta$$

(7.4)

The inner integral may be expressed as a sum of two integrals,

$$\int\limits_{0}^{2\pi} F(s,\theta)e^{js(x\cos\theta+y\sin\theta)}\,d\theta$$

$$= \int\limits_{0}^{\pi} F(s,\theta)e^{js(x\cos\theta+y\sin\theta)}\,d\theta + \int\limits_{\pi}^{2\pi} F(s,\theta)e^{js(x\cos\theta+y\sin\theta)}\,d\theta$$

$$= \int\limits_{0}^{\pi} F(s,\theta)e^{js(x\cos\theta+y\sin\theta)}\,d\theta$$

$$+ \int\limits_{0}^{\pi} F(s,\pi+\theta)e^{js(x\cos(\pi+\theta)+y\sin(\pi+\theta))}\,d\theta$$

Using the mapping

$$F(s,\pi+\theta) = F(-s,\theta)$$

equation (7.4) may be written as

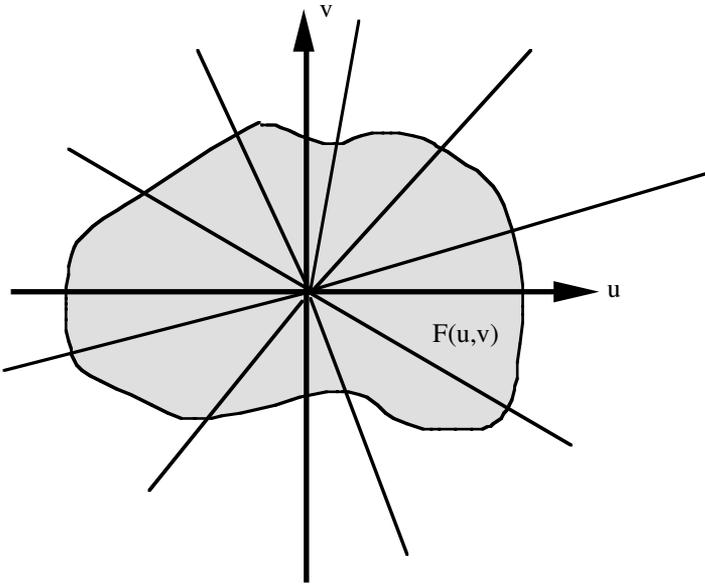Figure 7.3: The Fourier transform of the object function is scanned along a series of radial lines, one for each projection.

$$f(x,y) = \frac{1}{2\pi}\int\limits_0^\pi \left\{ \frac{1}{2\pi}\int\limits_{-\infty}^\infty F(s,\theta)|s|e^{js(x\cos\theta+y\sin\theta)}\,ds \right\}d\theta$$

$$= \frac{1}{2\pi}\int\limits_0^\pi \left\{ \frac{1}{2\pi}\int\limits_{-\infty}^\infty F(s,\theta)|s|e^{jst}\,ds \right\}d\theta \tag{7.5}$$

Note that, for a fixed θ, $F(s,\theta)$ is equal to $P_\theta(\omega)$ given in (7.3). Thus, the quantity inside the curly brackets in (7.5) may be obtained from the parallel projection by simply filtering it with a filter having a transfer function, $H(s) = |s|$.

$$f(x,y) = \frac{1}{2\pi}\int\limits_0^\pi \tilde{f}(x\cos\theta + y\sin\theta)d\theta \tag{7.6}$$

where

$$\tilde{f}(x\cos\theta + y\sin\theta) = \left\{ \frac{1}{2\pi}\int\limits_{-\infty}^\infty P_\theta(\omega)|\omega|e^{j\omega(x\cos\theta+y\sin\theta)}\,d\omega \right\} \tag{7.7}$$

is often known as a filtered projection. In (7.7) the projection data is projected back onto a section $t = x\cos\theta + y\sin\theta$ for a fixed θ and hence the reconstruction procedure is called the

back projection algorithm. The process of reconstruction consists of filtering each projection with a filter whose transfer function is $H(\omega) = |\omega|$ and then backprojecting according to (7.6).

7.1.3 <u>Algebraic Reconstruction</u>: The back projection algorithm requires a ULA capable of going round the target object. Such an idealistic experimental setup cannot be achieved atleast in one important area of application, namely, exploration for earth resources. The sensor array tends to be nonuniform and generally distributed over a large area. Furthermore, since experimental observations are necessarily confined to the earth's surface or a few deep borewells, it is practically impossible to go round the target object to get a 4π solid angle coverage. The back projection algorithm cannot be applied in most real situations except perhaps in seismic exploration for oil where a near ideal experimental setup can be realized. It is, therefore, necessary to devise an alternate approach, albeit less accurate and of lower resolution. The basic idea in this alternate approach is to divide the target object into as many homogeneous cells as possible. The cell size is small enough to allow the assumption of no variation in physical parameters (e.g., wave speed) over the size of a cell and large enough to allow the validity of ray propagation. As a ray propagates through the target object, it passes through many cells lying in its path. The ray as it reaches a sensor carries the cumulative effect of all cells, for example, sum of all delays introduced by all those cells lying in its path. It is further assumed that a ray does not suffer any refraction (or reflection) on the boundary of a cell and hence it travels in a straight line joining source and sensor. For this assumption to hold good the change in the acoustic impedance must be small (on the order of 10%). It is, however, possible to relax this assumption but only at the cost of increased computation as the ray paths have to be traced numerically. The reconstruction process becomes iterative where, starting from an initial estimate of the speed variation, successive corrections are introduced consistent with the available time delay information. The ray paths will have to be traced afresh at the beginning of each iteration. Each cell would introduce a delay proportional to the path length inside the cell and the unknown wave speed in the cell. Let d represent the length of each side of a cell and c_m be the wave speed in m^{th} cell. The maximum delay introduced by m^{th} cell will be

$$\tau_m = \frac{\sqrt{2}\delta}{c_m}.$$ The contribution of a cell toward the total delay observed at a

sensor would depend upon the sensor position. It is possible that some cells do not contribute at all. Indeed in any typical sensor array distribution there are many cells through which a ray joining the source to sensor will not pass at all; hence there can be no contribution from such cells. This may be easily seen in fig. 7.4 where out of 56 cells a ray passes through less than 10 cells. Let
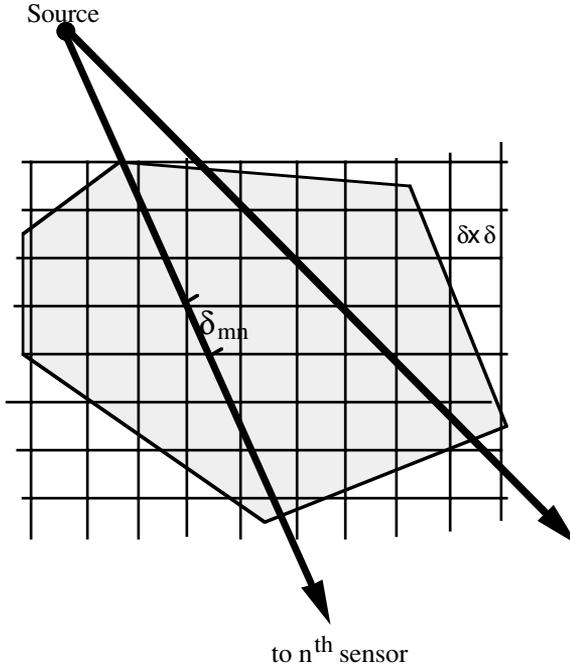
Figure 7.4: A target is divided into many square cells (or cubes in 3D). The wave speed is assumed to be constant in a cell. The path delay introduced by m^{th} cell is given by $\dfrac{\delta_{mn}}{c_m}$ where δ_{mn} is path length in m^{th} cell of a ray going to n^{th} sensor.

$w_{m,n}$ represent a weighting coefficient which when used along with the maximum delay gives a delay contributed by m^{th} cell to n^{th} sensor. In terms of path length of a ray in the mth cell the weighting coefficient is given by $w_{m,n} = \dfrac{\delta_{mn}}{\sqrt{2}\delta}$ (see fig 7.4). Thus, the total delay observed at the nth sensor is given by

$$t_n = \sum_{m=0}^{M-1} w_{m,n}\tau_m \qquad \text{n=0, 1,...,N-1} \qquad (7.8a)$$

where we have assumed that there are M cells and N sensors. In matrix form (1) may be expressed as

$$\mathbf{t} = \mathbf{w}\,\tau \qquad (7.8b)$$

where $\mathbf{t} = col\{t_0, t_1, \dots, t_N\}$, $\tau = col\{\tau_0, \tau_1, \dots, \tau_M\}$ and $\mathbf{w}$ is a $N \times M$ matrix of weight coefficients. When $N \geq M$ a unique solution of (7.8b) may be given by

$$\tau = (\mathbf{w}^T \mathbf{w})^{-1} \mathbf{w}^T \mathbf{t} \tag{7.9}$$

provided $\mathbf{w}^T \mathbf{w}$ is nonsingular. The question of the rank of the weight matrix $\mathbf{w}$ has no quantitative answer but we can give some qualitative guidelines:

      i) The ray path lengths in different cells must be quite different so that there is correlation among weight coefficients. In fig.7.4, this is more likely to happen with a fan beam than with a parallel beam.

      ii) The weight matrix is fully determined by the sensor array geometry. For example, if sensors are too closely spaced all rays will travel through the same group of cells and each cell will contribute roughly the same delay. The weight matrix will then tend to be more singular.

      iii) The sensor and source arrays must be so designed that the rays pass through different cells in different directions. More on this possibility will be discussed in the next subsection on borehole tomography.

It may be noted that since the weight matrix is generally a large sparse matrix, efficient techniques have been designed for fast and economical (in terms of memory requirements) inversion of the weight matrix. This is, however, beyond the scope of this book. The interested reader may like to review an article by S. Ivansson [2].

<u>Borehole Tomography</u>: The use of a source array often improves the rank condition of the weight matrix. Consider a P source array arranged in some unspecified form. Equation (7.8b) may be used to express the output as

$$\mathbf{t}_p = \mathbf{w}_p \tau, \qquad p = 0, 1, \dots P-1 \tag{7.10}$$

where $\mathbf{t}_p$ is the array output due to p^{th} source and $\mathbf{w}_p$ is the weight matrix corresponding to the position of pth source. Next, we stack up all array vectors into a single vector. Note that τ is independent of the source position. Equation (7.10) reduces to

$$\tilde{\mathbf{t}} = \tilde{\mathbf{w}} \tau \tag{7.11}$$

where $\tilde{\mathbf{t}} = \{\mathbf{t}_0, \mathbf{t}_1, \dots, \mathbf{t}_{P-1}\}$ and $\tilde{\mathbf{w}} = \{\mathbf{w}_0, \mathbf{w}_1, \dots, \mathbf{w}_{P-1}\}$. The solution of (7.11) may be expressed as
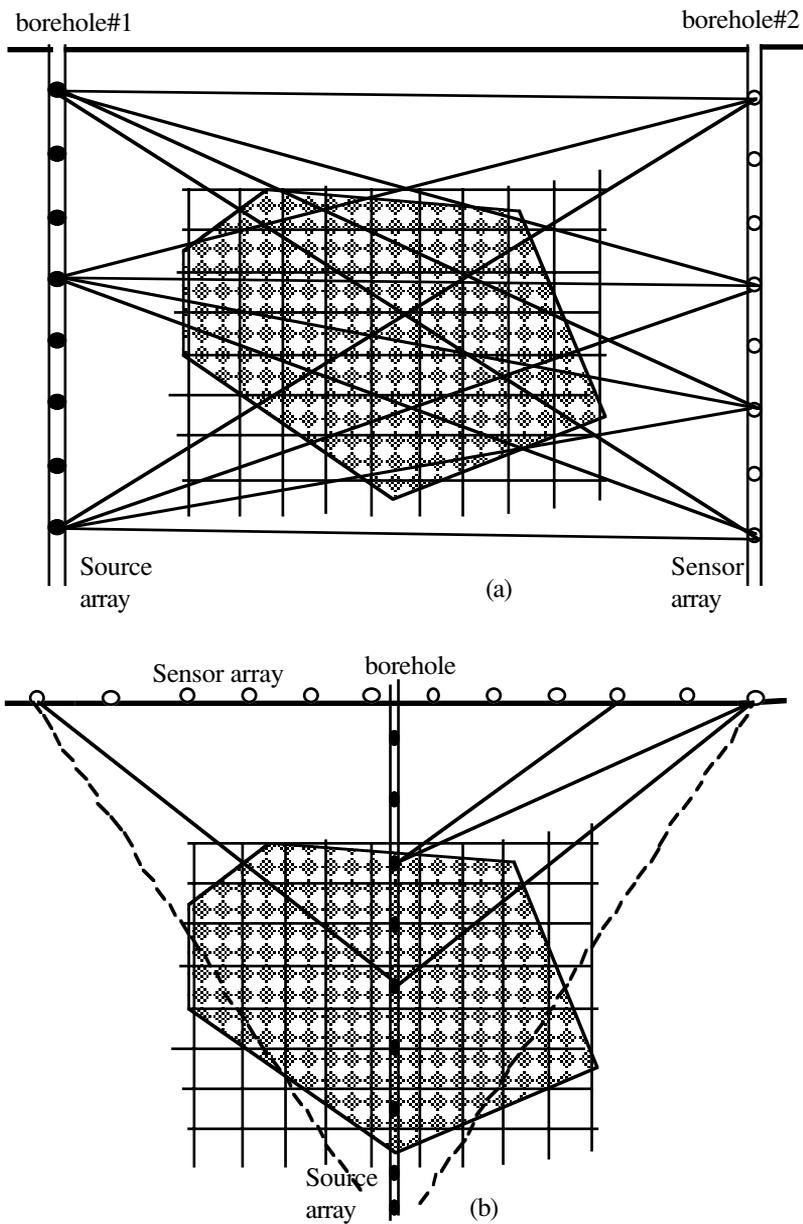
Figure 7.5: Source and sensor arrays in borehole tomography. (a) A target lies between two boreholes. In borehole# 1 sources are arranged as a ULA of sources and in borehole# 2 sensors are arranged as ULA of sensors. (b) In another arrangement the sensor array is on the surface and the source array is in the borehole.
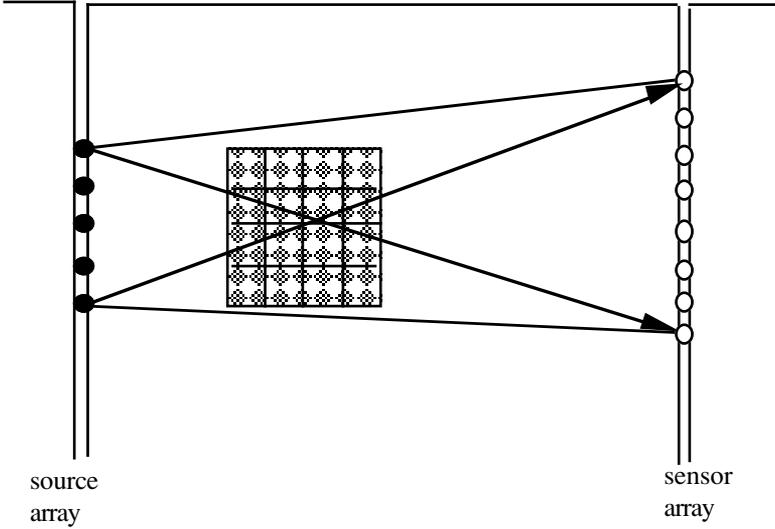
Figure 7.6: A target consisting of 16 cells (unit cells) lies between two boreholes separated by 16 units. An eight sensor array is located in the right hole and the source array (up to five sources) is in the left hole. The center of sensor array and source array is at the average depth of the target.

$$\tau = (\tilde{\mathbf{w}}^T \tilde{\mathbf{w}})^{-1} \tilde{\mathbf{w}}^T \tilde{\mathbf{t}} \tag{7.12}$$

We shall now consider a specific example of source and sensor arrays used in borehole tomography in geophysical exploration [3]. A typical arrangement in borehole tomography is shown in fig. 7.5. The sources are fired sequentially and the transmitted signals are recorded for later processing.

Source/Sensor Array Design: For successful reconstruction of wave speed variations the primary requirement is that $\tilde{\mathbf{w}}^T \tilde{\mathbf{w}}$ in (7.12) must be invertible. Since $\tilde{\mathbf{w}}$ is entirely determined by the source and sensor array geometry, it is possible to come up with a proper design for the source and sensor arrays which would make the rank of $\tilde{\mathbf{w}}$ equal to the number of cells. A simple numerical example is worked out to show how the rank of $\tilde{\mathbf{w}}^T \tilde{\mathbf{w}}$ depends on the number of sources, source spacing, and sensor spacing. The interesting outcome of this exercise is the fact that the number of sensors need not be greater than the number of cells. One may achieve by using multiple sources what could be achieved by using more sensors.

The source and sensor array geometry along with the target location are shown in fig. 7.6. The sensor spacing (d units) and source spacing (s units) are the variable parameters and the rest of the geometry remains fixed. From each

Number of sources	d=1			d=1.5	
	s=0.5	=1.0	=1.5	s=1.0	s=1.5
1	6			7	
3	13	14	15	16 (1350)	16 (1850)
5	15	15	16 (2200)	16 (2200)	16 (835)

Table 7.1: Rank of $\tilde{\mathbf{w}}^T\tilde{\mathbf{w}}$ as a function of source and sensor spacing. The bracketed quantity represents the eigenvalue spread.

source eight rays (straight lines) were drawn toward eight sensors. The line intercept in each cell was found and the weight coefficient was computed as described on page 370 (also see fig. 7.4). The weight matrix $\tilde{\mathbf{w}}$ is first computed and then the rank of $\tilde{\mathbf{w}}^T\tilde{\mathbf{w}}$, whose inverse is used in (7.12). The results are shown in table 7.1. To achieve the full rank property for $\tilde{\mathbf{w}}^T\tilde{\mathbf{w}}$ we must have three to five sources and the sensor spacing should be around 1.5 units. Note that we have considered only eight sensors which is half the number of cells. We have compensated for this deficiency by using three to five sources. When the source array is close to the target the angular width of the illuminating beam becomes large, which in turn requires a large aperture sensor array to capture the wide illuminating beam. However, indefinite increase of the sensor spacing will not help. There exists a range of sensor separation over which not only $\tilde{\mathbf{w}}^T\tilde{\mathbf{w}}$ is full rank but it is also stable as shown in fig 7.7 whereas for sensor separation between 1.5 and 3.0 units the eigenvalue spread of $\tilde{\mathbf{w}}^T\tilde{\mathbf{w}}$ is low and the matrix becomes singular outside the range 1.0 to 4.0. The above findings are specific to the geometry of source and sensor arrays and the target; nevertheless similar behavior is expected in other situations.

§7.2 **Diffracting Radiation:**
Diffraction becomes important whenever the inhomogeneities in an object are comparable in size to the wavelength of the wavefield used for imaging. In tomographic imaging, an object is illuminated from many different directions, either sequentially or simultaneously, and the image is reconstructed from the scattered wave field collected by an array of sensors, usually a linear array. Early workers who attacked the problem of deriving the inversion algorithm for tomography with diffracting wavefields were Iwata and Nagata [4] and Mueller
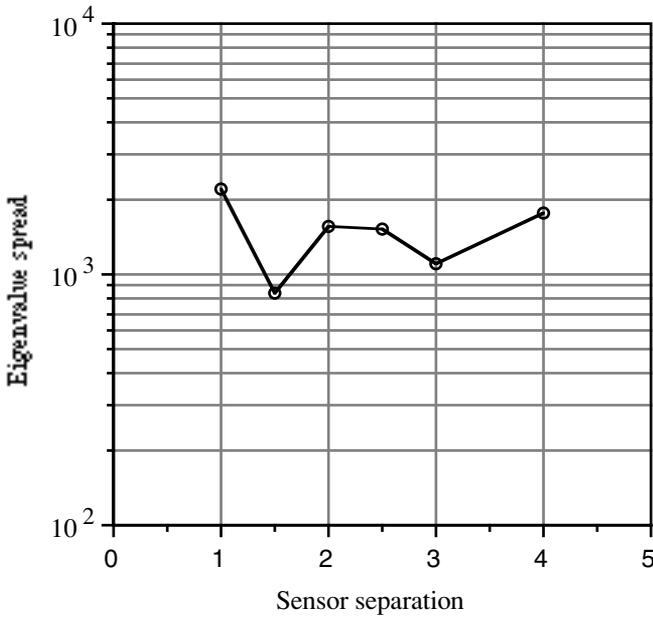
Figure 7.7: Eigenvalue spread as a function of sensor separation. Five sources are spaced at interval 1.5 units. The target used is same as in fig. 7.6.

et al. [5] who based their work on Wolf's work [6] on the inverse scattering problem assuming the first order Born approximation. A good review of diffraction tomography may be found in [1, 7].

Linear Array: An object is illuminated from various directions with a diffracting source of radiation such as acoustic waves whose wavelength is comparable to the scale of inhomogeneities. The incident wave energy is scattered in all directions by the diffraction process within the object. A long linear array of sensors facing the incident wave field is used to record the forward scatter (see fig. 7.8). In §1.6 we derived an expression for the scattered field in the x-y plane due to a plane wave traveling in z direction and illuminating a three dimensional object (spherical). A similar result for a two-dimensional object (cylindrical) was also given. For the sake of simplicity we shall talk about tomographic imaging of a two-dimensional object. Consider an arrangement wherein a cylindrical object is illuminated with a plane wave traveling at right angle to the axis and a linear array of sensors located on the opposite side as shown in fig. 7.8. The Fourier transform of the scattered field, which is measured at a set of discrete points by the sensor array, may be obtained from (1.82) where set $u_0=0$ and $v_0=k_0$,
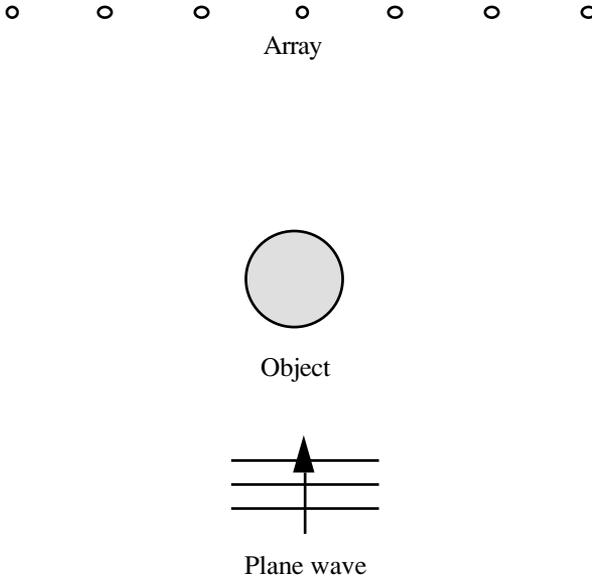
Array

Object

Plane wave

Figure 7.8: A basic experimental setup for tomographic imaging. A linear array, ideally of infinite aperture, is used to receive the forward scatter from the object.

$$P_0(u) = \frac{jk_0^2}{2} \frac{e^{j\sqrt{k_0^2 - u^2}\, l}}{\sqrt{k_0^2 - u^2}} \Delta\tilde{c}(u, \sqrt{k_0^2 - u^2} - k_0)$$

$$|u| \le k_0$$

(7.13a)

where $\Delta\tilde{c}(u,v)$ is a 2D Fourier transform of $\delta c(x,y)$. As u varies from $-k_0$ to $+k_0$ $P_0(u)$ traces a cross section of $\Delta\tilde{c}(u,v)$ along a semicircular arc as shown in fig. 7.9. The circle is centered at $(0, -k_0)$ and the radius is equal to k_0. The circular arc is described by an equation $v = \sqrt{k_0^2 - u^2} - k_0$, $|u| \le k_0$. The entire $\Delta\tilde{c}(u,v)$ may be sampled over a series of arcs either by rotating the object but keeping the direction of illumination and the position of array fixed or vice versa. We consider the first case. When the object is rotated so is its Fourier transform through the same angle. Hence, for an object, which is rotated through an angle φ, the scattered field, analogous to (7.13a), is given by the following equation:
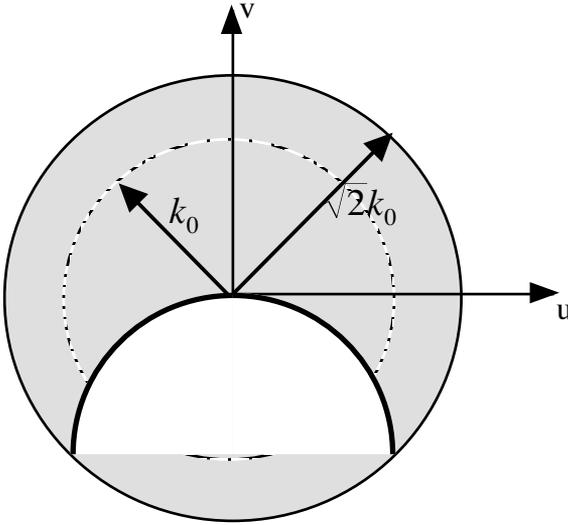
Figure 7.9: As u varies from - k_0 to +k_0 $P_0(u)$ traces a cross section of $\Delta\tilde{c}(u,v)$ along a semicircular arc (thick curve).

$$P_\varphi(u) = \frac{jk_0^2}{2} \frac{e^{j\sqrt{k_0^2 - u^2}\, l}}{\sqrt{k_0^2 - u^2}} \Delta\tilde{c}\left(\begin{array}{l} u\cos\varphi + (\sqrt{k_0^2 - u^2} - k_0)\sin\varphi, \\ -u\sin\varphi + (\sqrt{k_0^2 - u^2} - k_0)\cos\varphi \end{array} \right) \quad (7.13b)$$

$$|u| \le k_0 \ and \ -\frac{\pi}{2} \le \varphi \le \frac{\pi}{2}$$

The center of the circular arcs will all lie on a circle of radius k_0 centered at $(-k_0\sin\varphi, -k_0\cos\varphi)$ (see figure 7.10).

7.2.1 <u>Filtered Backpropagation Algorithm</u>: This is an adaptation of the filtered backprojection algorithm developed for nondiffracting radiation to diffracting radiation. The essential difference is that the sampling paths are now arcs of a circle instead of radial lines in nondiffracting radiation. We start by expressing the Fourier integral in polar coordinates (see (7.5-7.7))

$$f(x,y) = \frac{1}{2\pi}\int_0^\pi \left\{ \frac{1}{2\pi}\int_{-\infty}^\infty F(s,\varphi)|s|e^{js(x\cos\varphi + y\sin\varphi)}\, ds \right\} d\varphi$$
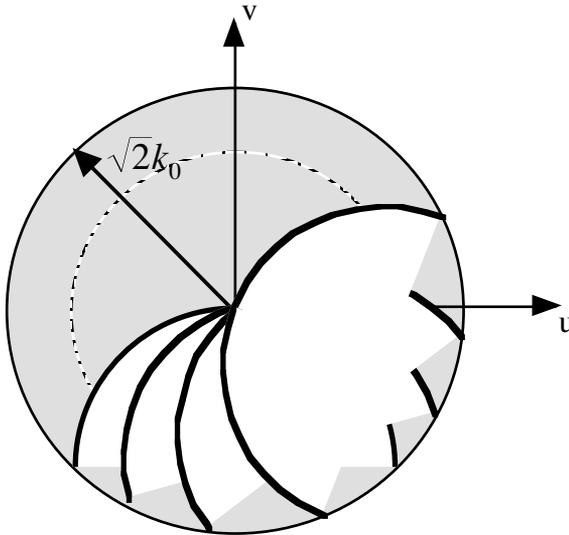
Figure 7.10: The entire 2D Fourier transform, $\Delta\tilde{c}(u,v)$, is sampled along a series of semicircular arcs by rotating the object keeping the transmitter and array fixed. Because u must lie within $-k_0$ to $+k_0$ the radius of the disc spanned by the semicircular sampling arcs is equal to $\sqrt{2}k_0$.

$$= \frac{1}{2\pi}\int_0^\pi \left\{\frac{1}{2\pi}\int_{-\infty}^\infty F(s,\varphi)|s|e^{jst}\,ds\right\}d\varphi$$

$$= \frac{1}{2\pi}\int_0^\pi \tilde{f}(x\cos\varphi + y\sin\varphi)d\varphi$$

where

$$\tilde{f}(x\cos\varphi + y\sin\varphi) = \left\{\frac{1}{2\pi}\int_{-\infty}^\infty F(s,\varphi)|s|e^{js(x\cos\varphi + y\sin\varphi)}\,ds\right\}$$

We shall use the above representation of 2D Fourier transform in polar coordinates. Note that in place of f(x,y) we have $\delta\tilde{c}$ and in place of $F(s,\varphi)$ we have $\Delta\tilde{c}$. Equation (7.13b) relates the Fourier transform of the scattered field to the object Fourier transform, that is,

$$\Delta\tilde{c}(u\cos\varphi + u'\sin\varphi, \ -u\sin\varphi + u'\cos\varphi)$$

$$= \frac{2}{jk_0^2} \sqrt{k_0^2 - u^2} \, e^{-j\sqrt{k_0^2 - u^2} \, l} P_\varphi(u)$$

where $|u| \le k_0$ *and* $u' = (\sqrt{k_0^2 - u^2} - k_0)$. Using the above relation we obtain the reconstruction

$$\delta\tilde{c}(x,y) = \frac{1}{2\pi} \int_0^\pi \delta\tilde{c}(x\cos\varphi + y\sin\varphi) d\varphi \qquad (7.14a)$$

where

$$\delta\tilde{c}(x\cos\varphi + y\sin\varphi) = \left\{ \begin{matrix} \dfrac{1}{2\pi} \displaystyle\int_{-k_0}^{k_0} \dfrac{2}{jk_0^2} \sqrt{k_0^2 - u^2} \, e^{-j\sqrt{k_0^2 - u^2} \, l} \\ \times P_\varphi(u)|u| e^{ju(x\cos\varphi + y\sin\varphi)} \, du \end{matrix} \right\} \qquad (7.14b)$$

According to (7.14b) the scattered field measured by the linear array is filtered with a filter whose transfer function is given by

$$H(u) = \sqrt{k_0^2 - u^2} \, e^{-j\sqrt{k_0^2 - u^2} \, l} |u| \qquad (7.14c)$$

The term of interest in the filter transfer function is $e^{-j\sqrt{k_0^2 - u^2} \, l}$ which represents backward propagation of the wavefield from the plane of observation to the target (see §1.2.4). For this reason the reconstruction algorithm described above is called the filtered backpropagation algorithm. Except for this difference the algorithm is quite similar to the filtered backprojection algorithm.

7.2.2 <u>Multisource Illumination</u>: There are many situations where it is not possible to turn an object around for multiple illuminations nor is it possible to turn around the source-array configuration, keeping the object fixed as the space around the object may not be accessible as is the case in geophysical imaging, nondestructive testing, remote monitoring, etc. In such a situation it is recommended to employ an array of sources, often arranged as a linear equispaced array. A typical example is borehole tomography which we have already considered in the previous section in the context of algebraic reconstruction. In this section we shall reconsider the same in the context of the backpropagation algorithm [8]. But, first, let us look at a simpler system, a circular array of transceivers of interest in medical imaging (see fig. 7.11).

<u>Circular Array</u>: The back scatter is lost in a linear array tomographic system; naturally, some potentially useful information is also lost, in the sense that only a part of the object spectrum lying within a disc of radius equal to $\sqrt{2}k_0$
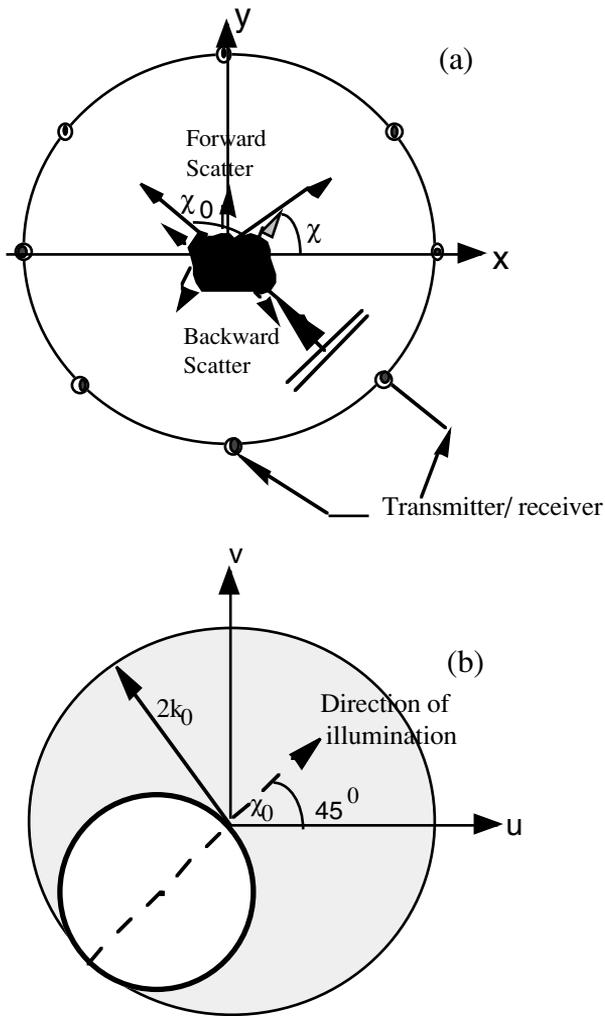
Figure 7.11: (a) An experimental setup of transceivers for tomographic imaging. (b) $\Delta C(u, v)$ is now sampled over a circumference of a circle of radius k_0 as shown. The angle of incidence of the plane wave illumination is 45°.

is utilized. An alternate tomographic system consisting of a circular array of sensors (transceivers) encircling the object is proposed [9]. In this configuration both forward scatter and backward scatter are captured. This results in the doubling of the area of the spectrum coverage, a disc of radius $2k_0$.

Arbitrarily shaped measurement boundaries were suggested in [10, 11] whose authors have shown that on a straight line boundary or on a circular boundary it is enough to measure either the diffracted field or its normal derivative. For a completely arbitrary boundary, we need both types of measurements [11]. A circular transducer array was used to illuminate an object with a pulse (broadband) from different directions [12, 13], and the back scatter alone, measured as a function of time, was employed for the purpose of object reconstruction. A circular array for ultrasound holographic imaging was used by Qin et al. [14] but they have approximated a circular array by a series of linear arrays and then applied the backpropagation algorithm. A circular array of transceivers was suggested for microwave diffraction tomography [15] where a near field diffraction phenomenon was used. The object Fourier transform was related to the scattered field through a two dimensional convolution relation. The scattered field measured with a large circular array surrounding the object (see fig. 7.11a) is proportional to the Fourier transform of the object profile taken on the circumference of a circle of radius equal to the wave number and centered at (-$k_0$cos χ_0, -$k_0$sin χ_0). This result is called here a Fourier diffraction theorem (FDT) for a circular array [9].

$$P_s(R, \chi, \chi_0) =$$

$$\frac{k_0^2}{4} e^{j(k_0 R + \frac{\pi}{4})} \sqrt{\frac{2}{\pi k_0 R}} \Delta C(k_0(\cos\chi - \cos\chi_0), k_0(\sin\chi - \sin\chi_0)) \quad (7.15)$$

The left hand side is simply the observed scattered field on a large circle. The right hand side is a Fourier transform of the object function which is evaluated on a circle of radius k_0 and centered at $k_x = -k_0 \cos\chi_0$ and $k_y = -k_0 \sin\chi_0$ (see fig. 7.11).

By changing the angle of incidence of the wavefront, χ_0, it is possible to cover the Fourier plane with a series of circles spanning a disc of radius equal to $2k_0$ (see fig. 7.12). The increased coverage has been possible because we captured the back scatter as well. Note the important difference is that the scattered field measured with a circular array, being in the far field region, directly yields the object Fourier transform. On the contrary, with a linear array we need to Fourier transform the observed field. This important result is first verified against the measured scattered field.

Verification of Fourier Diffraction Theorem: First we shall verify the FDT through an example where an exact scattered field as well as its object Fourier transform are known. Consider a liquid cylinder in water and assume that its refractive index is slightly above that of the water, $\delta n \leq \frac{\lambda}{4a}$ where δn is
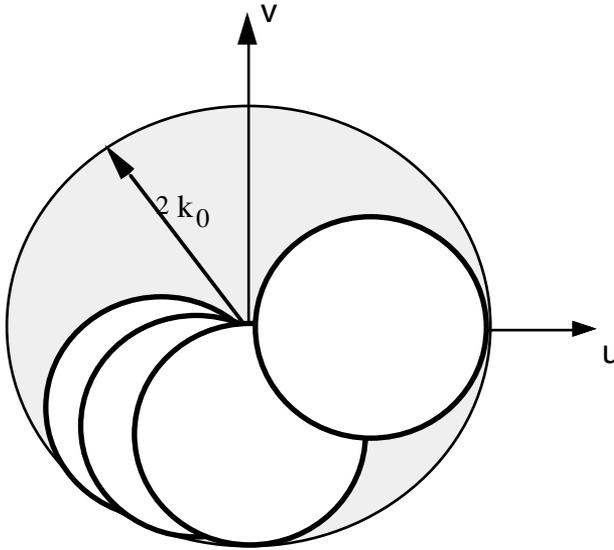
Figure 7.12: By changing the angle of incidence of the wavefront, χ_0, it is possible to cover the Fourier plane with a series of circles spanning a disc of radius equal to $2k_0$.

change in the refractive index, a is radius of the cylinder and λ is wavelength of illuminating wave (Born approximation). The scattered field due to a liquid cylinder, when it is illuminated by a plane wave, was experimentally studied by [16] and theoretically by [17]. We have computed the scattered field using the derivation given in [17]. Next, we evaluate the object profile Fourier transform

$$\tilde{O}(k_0(\sin\chi - \sin\chi_0), k_0(\cos\chi - \cos\chi_0)) = 2\pi\delta_n a \frac{J_1(ka)}{k}$$

where

$$k = \sqrt{(\sin\chi - \sin\chi_0)^2 + (\cos\chi - \cos\chi_0)^2} \ k_0$$

as a function of χ for a fixed direction of illumination; in the present case, $\chi_0 = 0^o$. The scattered field measured by a circular array is now compared with the Fourier transform of the object profile evaluated on a circle of radius k_0

centered at $(\dfrac{-k_0}{\sqrt{2}}, \dfrac{-k_0}{\sqrt{2}})$ (see).

7.2.3 <u>Bilinear Interpolation Algorithm</u>: As the Fourier transform of sound speed fluctuations is obtained from the scattered field either from a linear array or circular array, in principle it is possible, by inverse Fourier transformation, to estimate the sound speed fluctuations. In practice, however, this is not a trivial step. As the object Fourier transform is measured over a series of arcs it will be necessary to interpolate to the nearest square grid point and thus create a discrete Fourier transform matrix for inversion. Alternatively, the back propagation method originally developed for nondiffracting radiation may also be used. We shall use the interpolation approach.

Consider a circular array of transceivers. The angle of illumination χ_0 is varied over 360° by switching on a sensor to transmit mode one at a time and keeping the remaining sensors in the receive mode. The received field can be expressed as a function of two parameters, namely, the angle of illumination and the angular coordinate of each sensor, that is, (χ, χ_0). We must then map every point in (χ, χ_0) space onto (k_x, k_y) space. The reverse mapping, i.e., from (k_x, k_y) space onto (χ, χ_0) space, is, however, more convenient to use. The to-and-fro mapping functions are as follows:

$$\cos\chi - \cos\chi_0 = \frac{k_x}{k_0}, \ \ \sin\chi - \sin\chi_0 = \frac{k_y}{k_0} \qquad (7.16)$$

Solving the above equations for χ_0, we get the following inverse mapping functions:

$$\chi_0 = \tan^{-1}\left\{\frac{-k_y - k_x\sqrt{\dfrac{2}{p}-1}}{-k_x + k_y\sqrt{\dfrac{2}{p}-1}}\right\} \qquad \chi = \tan^{-1}\left\{\frac{k_y - k_x\sqrt{\dfrac{2}{p}-1}}{k_x + k_y\sqrt{\dfrac{2}{p}-1}}\right\} \qquad (7.17)$$

where $p = \dfrac{k_x^2 + k_y^2}{2k_0^2}$. Equations (7.16) and (7.17) together give a set of transformation equations that can be used to map from the k-plane into the χ-plane. Every point in the k-plane is mapped onto the χ-plane. The values of (χ, χ_0) thus obtained may not correspond to any of those points where the
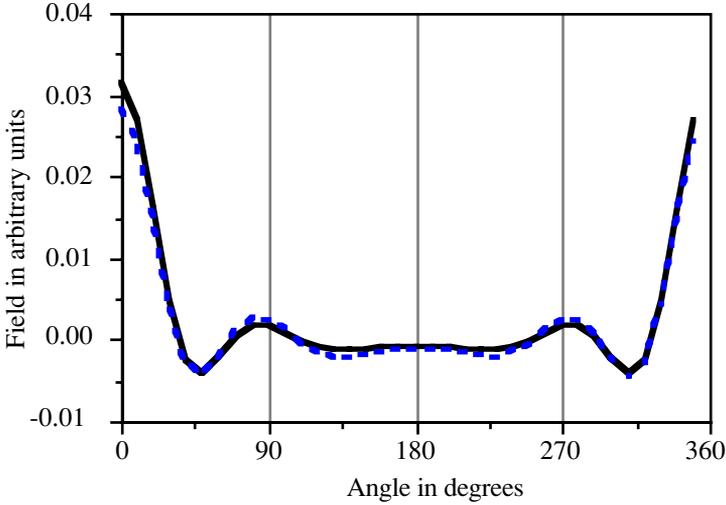
Figure 7.13: The scattered field measured by a circular array is now compared with the Fourier transform of a uniform cylindrical object evaluated on a circle centered at $(\frac{-k_0}{\sqrt{2}}, \frac{-k_0}{\sqrt{2}})$ as shown in fig. 7.11. The mean square error is 1.759×10^{-5}.

scattered field is observed; then we must take recourse to some form of interpolation. For example, bilinear interpolation is given by

$$O(\chi, \chi_0) = \sum_i \sum_j O(\chi_i, \chi_j) h_1(\chi - \chi_i) h_2(\chi_0 - \chi_j)$$

where $h_1(\chi) = 1 - \dfrac{|\chi|}{\Delta\chi}$, for $|\chi| \leq \Delta\chi$ otherwise $=0$ and $h_2(\chi_0) = 1 - \dfrac{|\chi_0|}{\Delta\chi_0}$, for $|\chi_0| \leq \Delta\chi_0$ otherewise $=0$. Here $\Delta\chi$ and $\Delta\chi_0$ are the sampling intervals.

Once the values of the Fourier transform are obtained over a rectangular grid in (k_x, k_y) space, the inverse two dimensional Fourier transform can be computed to obtain the object profile. The above algorithm is essentially an adaptation of the frequency domain interpolation algorithm, which is known to be very fast [1].

7.2.4 Imaging with a Circular Array: Since a circular array captures the entire diffracted energy, that is, both forward and backward scattered energy, a greater part of the object spectrum is utilized, indeed twice that of forward-scatter-only (linear array) setup. Consequently, we expect a better resolution of small inhomogeneities. To demonstrate this, we have c arried out the following
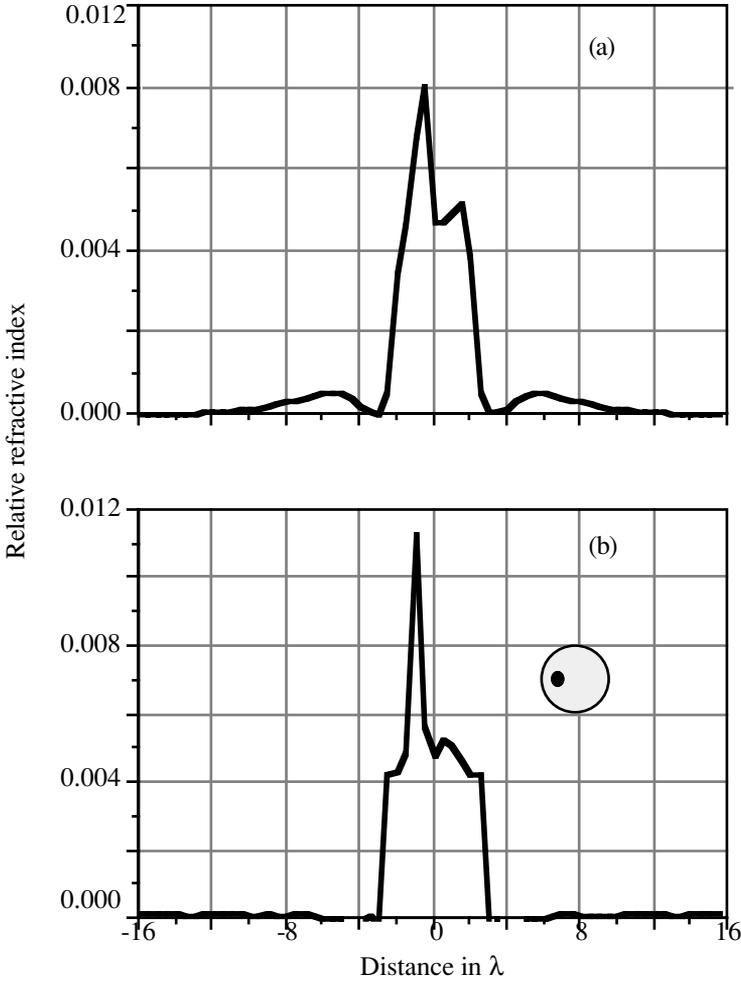
Figure 7.14: An example of improved performance of the circular array over linear array. A cylinder of radius 2λ with a small inhomogeneity of radius 0.25λ embedded in it (see inset in (b)) is used as a test target. (a) Reconstruction using a linear array and (b) using a circular array.(From [18] with permission.)

numerical experiment [18]. A small inhomogeneity of radius 0.25λ is embedded in a larger cylinder of radius 2λ having a refractive index contrasts with respect to the surrounding medium of 0.01 and 0.005, respectively (fig. 7.14b). The target is insonified with a narrowband plane wave radiation of wavelength 1.0λ. A circular array of 64 transceivers is assumed. For comparison we have also
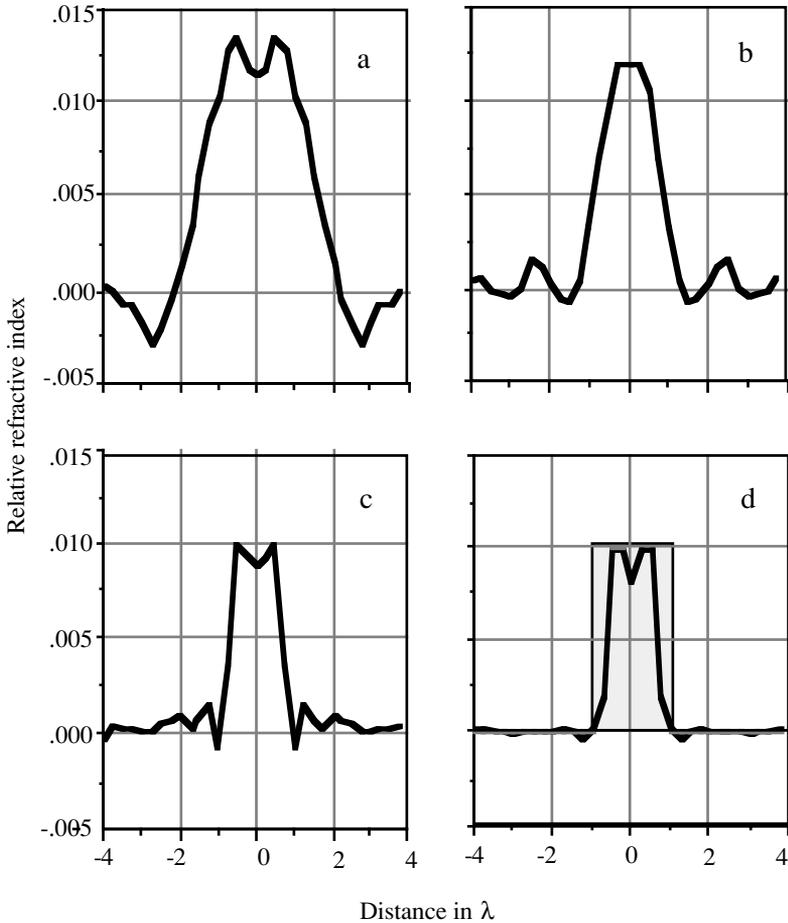
Figure 7.15: A comparison of performance of a linear array of finite size with that of a circular array. The number of illuminations in *all* cases was 64. (From [18] with permission.)

considered a linear array of the same length and one transmitter located on the broad side, and the scattered field was calculated using the object Fourier transform over semicircular arcs. For a circular array, however, the scattered field was computed using the exact solution given in [16]. The reconstruction (a central slice) of the target is shown in fig. 7.14. The reconstruction obtained using a linear array is shown in fig. (7.14a) and that obtained using a circular array is shown in fig. (7.14b). Clearly the circular array outperforms the equivalent linear array as the small inhomogeneity is more accurately located. Next, we would like to emphasize the role of the array size on object (fig. 7.15)
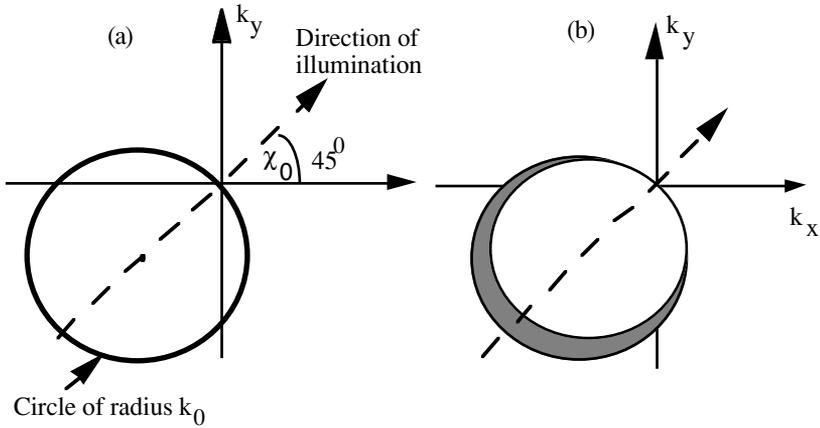
Figure 7.16: (a) For narrowband the scattered field is proportional to the object Fourier transform evaluated on a circle. (b) For finite band, the scattered field is proportional to the object Fourier transform evaluated inside a crescent shaped region. (From [18] with permission.)

reconstruction. When using a linear array, it is necessary that the array output be Fourier transformed before it is used for reconstruction. Consequently, the errors in the Fourier transformation due to finite size of the array will degrade the reconstruction. This effect is demonstrated in fig. 7.15. The first three figures (7.15a, b, c) were obtained using a linear array of three different sizes, namely, 64, 128, and 512 receivers spaced at $\frac{\lambda}{2}$ and 100λ away from the object, and a cylinder of radius 1λ with a refractive index contrast of 0.01. The scattered field was computed using the exact solution given in [16]. The reconstruction shown in fig. (7.15d) was obtained using a 64 element circular array (radius=100λ). The reconstruction obtained with the circular array is superior to that obtained with a linear array of a much larger size (512 receivers). Notice that the side lobes have practically disappeared .

§7.3 **Broadband Illumination**: The object of interest is illuminated with a broadband signal which results into a better coverage of the spatial spectrum of the object with fewer illuminations. We shall in this section examine the effectiveness of broadband illumination.

7.3.1 <u>Spectrum Coverage</u>: If the object is illuminated from one direction with a set of different frequencies covering a spectral band, the scattered field will correspond to a set of circular arcs covering a crescent shaped region in the object Fourier plane, as shown in fig. 7.16b. The radii of the inner and outer circles forming the crescent are related to the lower and upper cut-off frequencies. Let f_l and f_u be the lower and upper cut-off frequencies

respectively; the radii of the circles are $\dfrac{2\pi f_l}{c_0}$ and $\dfrac{2\pi f_u}{c_0}$, respectively. By suitably selecting the lower or upper cut-off frequencies it is possible to emphasize the low or high frequency spatial spectrum of the object. With a single illumination it is thus possible to cover a fraction r of the disc of maximum size of radius $2f_u$, where

$$r = \frac{f_u^2 - f_l^2}{4f_u^2} \tag{7.18}$$

Note that $r \le 0.25$, where the maximum is achieved when $f_l = 0$. To cover the entire disc we need four or more illuminations, but certainly far fewer than the number of illuminations for narrowband illumination. With more than two illuminations certain regions are likely to be covered more than once; thus the fraction of covered area will be less than that predicted by (7.18). For example, for four illuminations the area covered is shown in fig. 7.17 where we have assumed that $f_l = 1.5\ kHz$ and $f_u = 3.0\ kHz$. The fraction of the Fourier plane covered is 75%. It is further possible to increase this fraction, in particular, by decreasing the lower cut-off frequency. It is straightforward to analytically compute the area covered by one or more crescents, although it becomes difficult with increasing number of illuminations. In table 7.2 we list the fraction of area covered as a function of lower cut-off but keep the upper cut-off frequency fixed at 3 kHz. Note that one hundred percent coverage is not possible even when $f_l = 0$.

Further, to demonstrate the effect of the lower cut-off, we have carried out computer reconstruction using four illuminations and 64 sensors (circular array). The results are shown in fig. 7.18. There is an overall improvement in the reconstruction with decreasing lower cut-off frequency. Evidently, increasing the number of illuminations will also increase the coverage of the Fourier plane. For example, in table 7.3 we show how the coverage increases as we increase the number of illuminations. Here the lower and upper cut-off frequencies are held fixed at 250 and 3000 Hz respectively. Four to eight illuminations seem to be ideal as not much is gained by going beyond eight illuminations, which gives almost 95% coverage.

7.3.2 <u>Signal Processing Issues</u>: The major issue in tomographic imaging from the point of signal processing relates to the fact that the object Fourier transform is sampled on a nonstandard grid, such as polar rastor in nondiffracting tomography or circular arcs in diffraction tomography. All existing methods of reconstruction require interpolation to convert from nonstandard grid to standard square grid. Another issue relates to the fact that in any practical implementation of the tomographic imaging scheme only a
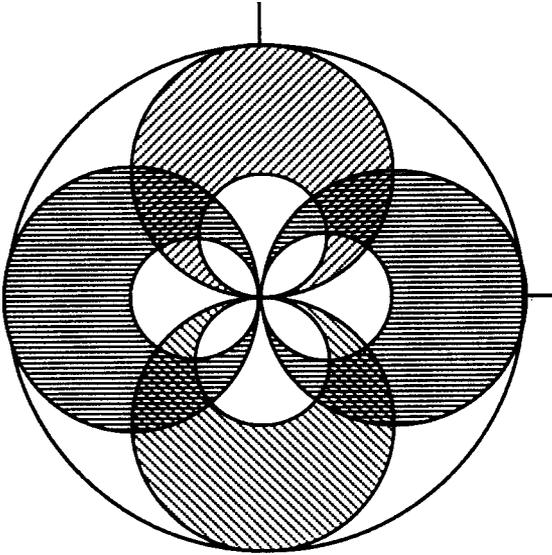
Figure 7.17: Broadband illumination provides a better coverage of the object Fourier transform. Just with four illuminations it is possible to get a 75% coverage. The lower and upper cut-off frequencies are 1.5 and 3.0 kHz, respectively.(From [18] with permission.)

Lower cutoff frequency(Hz)	Area covered %
2000	34.35
1500	75.00
750	80.72
500	81.37
250	81.71
0	82.00

Table 7.2: The fraction of the disc area covered with a broadband signal whose lower cutoff is varied and upper cutoff is held fixed (3.0kHz). We have assumed four illuminations.(Source [18] with permission.)

limited number of views, often covering a finite angular interval, are likely to be available, leaving large gaps in the Fourier plane. As shown earlier, broadband illumination can help to reduce the gaps. Signal processing tools have also been suggested for extrapolation of the observed Fourier transform into the missing gaps. A priori information about the object, such as a limit on the support of the object function either in space or frequency domain, does
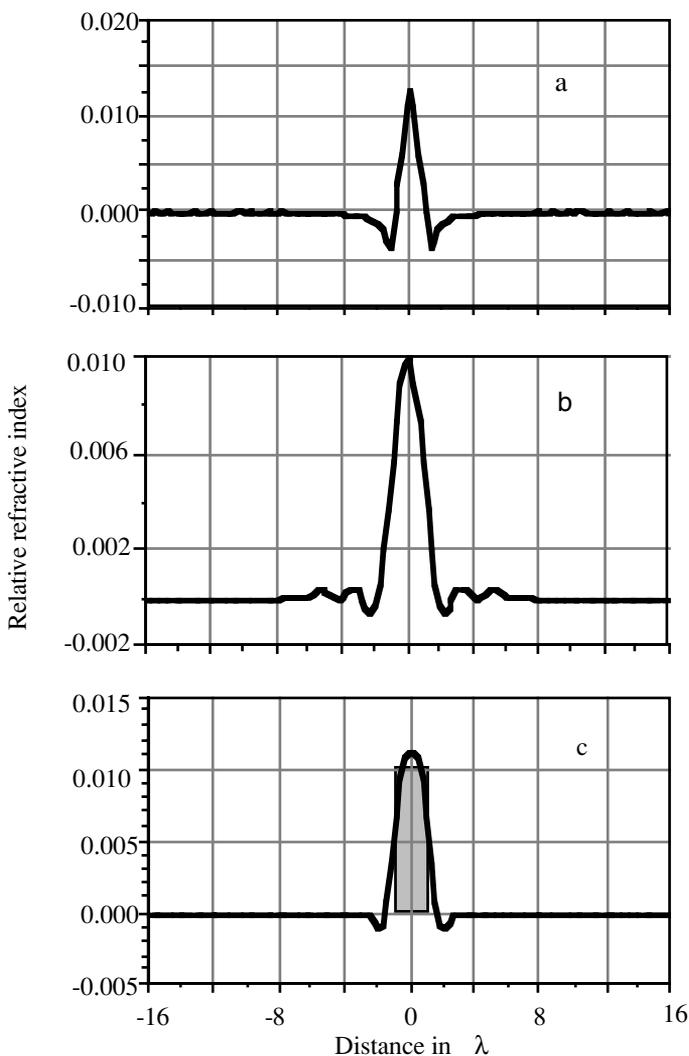
Figure 7.18: The effect of the lower cut-off frequency on the reconstruction. (a) 1500Hz, (b) 500Hz and (c) 250Hz. The upper cut-off frequency is 3000 Hz. (From [18] with permission.)

not help to uniquely reconstruct the object function from a limited number of samples on an algebraic contour in the Fourier plane [19]. Extrapolation outside the frequency domain, where the observations are available, has been attempted using the principle of maximum entropy [21, 22] which is familiar to the signal processing community as it is extensively used to extrapolate the covariance function for spectrum estimation [20].

Number of illuminations	Area covered %
1	24.83
2	49.66
4	81.71
8	94.9
12	97.2

Table 7.3: Fraction of the disc area covered with a broadband signal (250-3000Hz) for a different number of illuminations. (Source [18] with permission.)

7.3.3 <u>Cross Borehole Tomography</u>: In chapter 1 we have derived a relationship between the scattered wavefield from a target which has been illuminated by a source array positioned in one borehole and receiver array in another borehole (pp. 65-68). Consider a two dimensional target for which the relationship is given by (1.91) which we reproduce here for convenience,

$$F_1(u_1, u_2) = 2k_0^2 \Delta \tilde{c} (\frac{u_1}{d} + \frac{u_2}{d}, \sqrt{k_0^2 - (\frac{u_1}{d})^2} - \sqrt{k_0^2 - (\frac{u_2}{d})^2})$$

$$\times \frac{e^{-j\sqrt{k_0^2 - (\frac{u_1}{d})^2} L_a}}{\sqrt{k_0^2 - (\frac{u_1}{d})^2}} \frac{e^{-j\sqrt{k_0^2 - (\frac{u_2}{d})^2} L_b}}{\sqrt{k_0^2 - (\frac{u_2}{d})^2}} \qquad (1.91)$$

Let $d = \frac{\lambda_0}{2}$, $u_1' = \frac{2u_1}{\lambda_0}$, and $u_2' = \frac{2u_2}{\lambda_0}$. For $\pi \le u_1, u_2 \le -\pi$ it turns out that $k_0 \le u_1', u_2' \le -k_0$. A point in the Fourier plane, (u_1', u_2'), would correspond to a point (u,v) in the Fourier plane of $\Delta \tilde{c}$ where

$$u = u_1' + u_2'$$
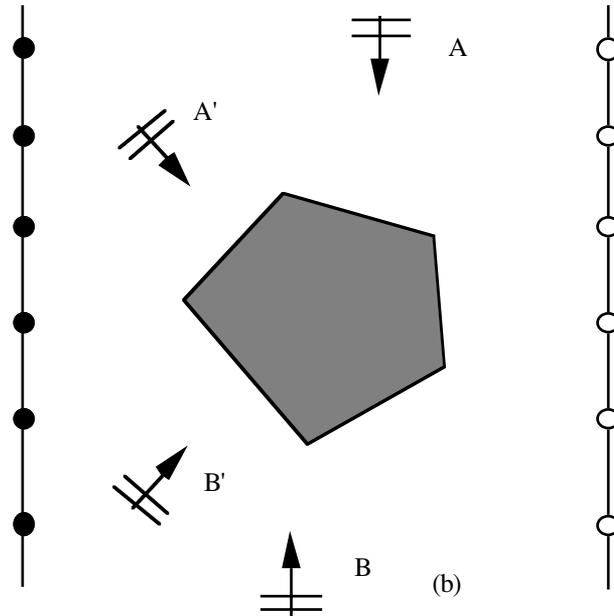$$v = \pm \left\{ \sqrt{k_0^2 - (u_1')^2} - \sqrt{k_0^2 - (u_2')^2} \right\} \qquad (7.19)$$

Eliminating u_2' from two equations in (7.19) we obtain

$$\left[v \pm \sqrt{k_0^2 - (u_1')^2} \right]^2 + (u - u_1')^2 = k_0^2 \qquad (7.20)$$

For a given value of u_1' equation (7.20) describes a circle with radius k_0 and centered at $(u_1', \pm\sqrt{k_0^2 - (u_1')^2})$, for example; for $u_1' = 0$ the two circles are
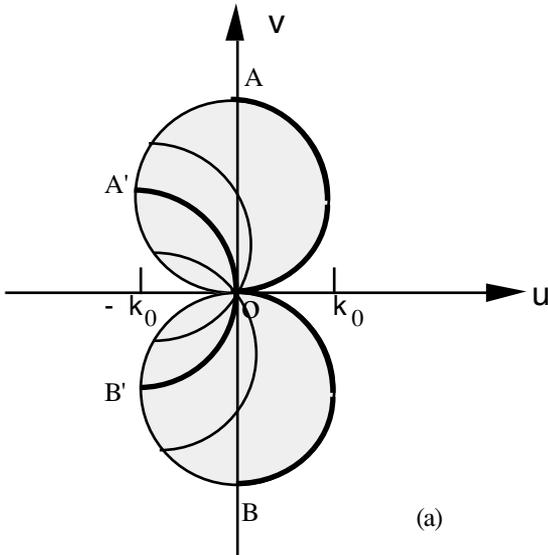
Figure 7.19: (a) The object Fourier transform lying inside two circular disks is scanned along a series of semicircular arcs. (b) As an example, consider four plane wavefronts A, A´, B, B´ emitted by distant sources in the source array. Corresponding circular arcs are shown in (a). The object Fourier transform is scanned on semicircle AO by wavefront A and on OB by wavefront B. Wavefronts A´ and B´scan the object Fourier transform on A´OB´.
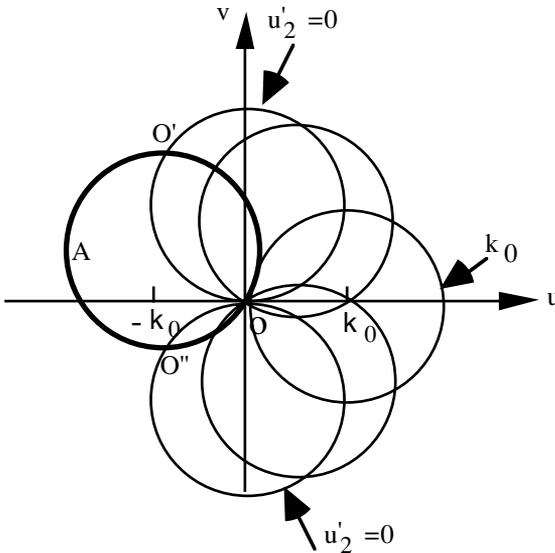
Figure 7.20: For a fixed u_1', the object Fourier transform is scanned over a circle A. For different values of u_2', that is, for different directions of the beam, a series of circles will intersect the circle A over an arc $O''OO'$. The object Fourier transform will be sensed over this arc only.

centered on the y-axis at $\pm k_0$. In fig. 7.19 we show different circles (arcs) for different values of u_1'; in particular, the thick arcs are for $u_1' = 0$ and $u_1' = -k_0$ and the thin arcs are for $u_1' = \pm \dfrac{k_0}{\sqrt{2}}$. Similarly, by eliminating u_1' from (7.19) we obtain

$$\left[v \pm \sqrt{k_0^2 - (u_2')^2}\right]^2 + (u - u_2')^2 = k_0^2 \tag{7.21}$$

which describes a circle with radius k_0 and centered at $(u_2', \pm\sqrt{k_0^2 - (u_2')^2})$. For a fixed u_2' the object Fourier transform is scanned over a circle, for example, A in fig. 7.20.

Let the receiving array (ULA) be beamed to receive the wavefield in some particular direction, that is, for some fixed u_2'. For $u_2' = 0$ the circle described by (7.21) will intersect circle A at two points, namely, O' and O (see fig. 7.20). The object Fourier transform is sensed only at these points of intersection. For different values of u_2', that is, for different directions of the beam, we can draw a series of circles which will intersect the circle A over an

arc O″OO′. Some of these circles are shown in fig. 7.20. There will be as many arcs as the number of sources. In the ideal situation of an infinitely long source array the arcs over which the object Fourier transform is sensed will fill

two circles shown in fig. 7.19. A few arcs corresponding to $u_2' = \pm k_0, \pm \dfrac{k_0}{\sqrt{2}}$

are also shown in this figure. In summary, in cross borehole tomography the object Fourier transform can be scanned over a pair of circles along a series of semicircular arcs as shown in fig. 7.19.

7.3.4 <u>Maximum Entropy Extrapolation</u>: The object Fourier transform scanned in cross borehole tomography (see fig. 7.19) skips a good portion of the Fourier plane, in particular along the u-axis. Consider a square, $4k_0 \times 4k_0$, superscribing the two circles. It is found that just about 39% of the square is included within the two circles. Consequently, the resolution particularly along the x-axis will be considerably deteriorated. This is true only for infinitely long source and sensor arrays. Additionally, since in any borehole tomographic experiment the array size is bound to be finite, there will be a further reduction of the scanned area [21]. To overcome the effect of undercoverage of the object Fourier transform it has been suggested to extrapolate the measured Fourier transform into an area where no measurements were possible, using the principle of maximum entropy [21, 22]. We shall briefly review this approach.

The aim of tomographic imaging is to determine the wave speed function $1 + \delta\tilde{c}(x, y)$. We have already shown how from the scattered field we can obtain the Fourier transform of the wave speed variations over a limited domain (fig. 7.19), hereafter referred to as Ω, which is completely determined by the geometry of experiment. Nothing is however known of the wave speed outside this domain. Of the many possible functions we choose one, which is consistent with the observed Fourier transform in the specified domain and is maximally noncommittal with regard to unavailable data. This is the principle of maximum entropy founded by Burg [23] and widely used in spectrum analysis [20]. The constrained optimization problem may be formulated as follows:

minimize:

$$H = \iint_\Gamma \left(1 + \delta\tilde{c}(x,y)\right)\ln\left(1 + \delta\tilde{c}(x,y)\right)dxdy \qquad (7.22)$$

subject to

$$\Delta\tilde{c}(u,v) = \Delta\tilde{c}(u,v)\big|_{measured\ in\ \Omega} \qquad (u,v) \in \Omega \qquad (7.23)$$

$$\partial H = \iint_\Gamma \partial\delta\tilde{c}(x,y)\left[1 + \ln\left(1 + \delta\tilde{c}(x,y)\right)\right]dxdy$$

$$= \frac{1}{4\pi^2} \int_{-k_0}^{k_0} \int_{-k_0}^{k_0} \partial\Delta\tilde{c}(u,v) \int\int_{\Gamma} \left[1 + \ln\left(1 + \delta\tilde{c}(x,y)\right)\right] e^{j(ux+vy)} dxdydudv$$

H will be minimized when $\partial H = 0$. Since $\Delta c(u,v)$ is already specified in $(u,v) \in \Omega$, $\partial\Delta\tilde{c}(u,v) = 0$ in Ω. Therefore, maximization of entropy requires

$$\int\int_{\Gamma} \left[1 + \ln\left(1 + \delta\tilde{c}(x,y)\right)\right] e^{j(ux+vy)} dxdy = 0 \qquad \text{in } (u,v) \notin \Omega \quad (7.24)$$

The solution is obtained by requiring to alternatively satisfy (7.23) and (7.24). The algorithm for the constrained optimization has the following steps:

(1) Compute $q(u,v) = \int\int_{\Gamma} \left[1 + \ln\left(1 + \delta\tilde{c}(x,y)\right)\right] e^{j(ux+vy)} dxdy$

(2) Set q(u,v)=0   for $(u,v) \notin \Omega$

(3) Compute FT $\left\{ \delta\tilde{c}(x,y) = e^{IFT\{q(u,v)\}-1} - 1 \right\}$

(4) Set $\Delta\tilde{c}(u,v) = \Delta\tilde{c}(u,v)\big|_{measured}$ *for* $(u,v) \in \Omega$

(5) Compute $\delta\tilde{c}(x,y) = IFT\left\{\Delta\tilde{c}(u,v)\right\}$ and go to step 1

The procedure is terminated when the reconstructed function meets some criterion of convergence. The algorithm is demonstrated through an example. A square object of size (5x5) with wave speed c (=1+0.1) is embedded in a background with wave speed, c_0=1.0. The Fourier transform of the object was computed over a grid of 64x64. Let us assume that cross borehole geometry permits measuring the object Fourier transform over a figure of 8 (see fig. 7.19a) where the radius of the pass disc is four. The filtered object Fourier transform was then used in the maximum entropy reconstruction algorithm described above. The reconstructed object (a horizontal cross section) after four iterations is shown in fig. 7.21, but it did not change much even after 40 iterations. Much of the spectrum lost during filtering remains unrecoverable except what lies between the upper and lower discs.

§7.4 **Reflection Tomography:**
The wave field returned by an object may be considered either as a reflection at the surface of discontinuity in physical parameters (wave speed, density, etc.) or backscattering from deep inside the object due to variable physical parameters. In this section we shall deal with the latter situation and reserve the former to
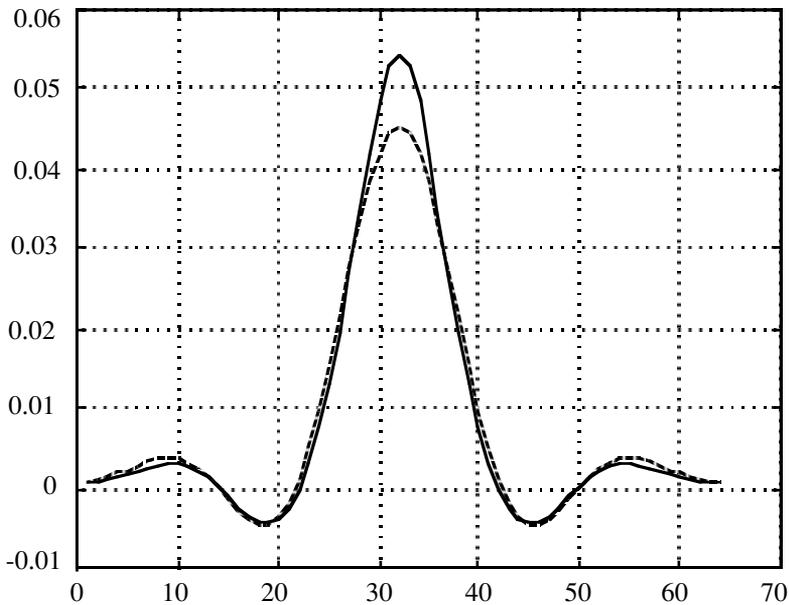
Figure 7.21: Maximum entropy reconstruction, dashed line (--) filtered object and solid line (__) maximum entropy reconstruction. There is only a marginal improvement in the maximum entropy reconstructed object.

be dealt with in chapter 8. The backscatter from an inhomogeneous object may be related to the inhomogeneity inside the object under the usual assumption of weak scattering (Born approximation). Indeed, the backscattering coefficient in frequency domain may be easily derived from (7.15) by setting $\chi = -(180 - \chi_0)$. We obtain

$$P_s(R, \chi_0) = \frac{k_0^2}{4} e^{j(k_0 R + \frac{\pi}{4})} \sqrt{\frac{2}{\pi k_0 R}} \Delta C \, (-2k_0 \cos\chi_0, -2k_0 \sin\chi_0) \quad (7.25)$$

where R now stands for distance to the center of the object whose size is assumed to be much smaller than R, so that the far field approximation hold good. The experimental setup is shown in fig. 7.22a. The backscatter coefficient at a fixed frequency, $P_s(R, \chi_0)$, is proportional to the object Fourier transform at spatial frequency $(-2k_0 \cos\chi_0, -2k_0 \sin\chi_0)$ where χ_0 is the angle of illumination (see fig. 7.22b). If we now use a broadband signal for illuminating the object (keeping χ_0 fixed) and Fourier decompose the
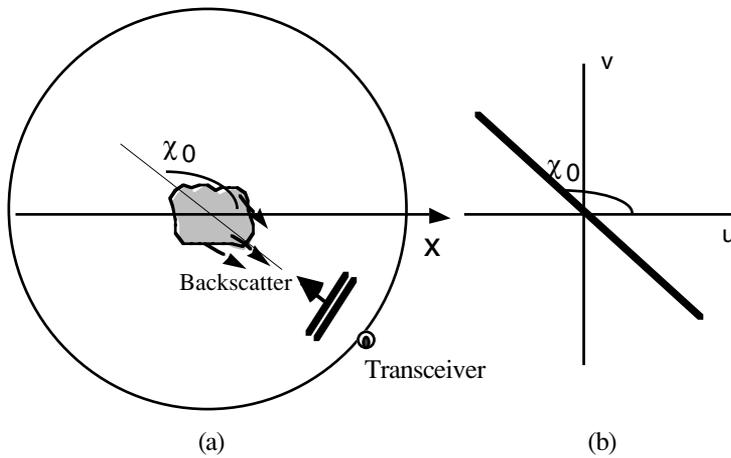
Figure 7.22: (a) A single transceiver is used to illuminate an object and receive the backscatter. A plane wavefront is assumed to be incident at an angle χ_0. (b) Backscatter coefficients are proportional to the object Fourier transform over a radial line as shown.

received signal we shall obtain the object Fourier transform over a radial line at an angle χ_0 (see fig. 7.22b). By illuminating the object repeatedly over 360°, either by physically taking the transceiver around the object or rotating the object around its axis, we can cover the entire Fourier plane. This commonly used experimental setup was suggested by [13]. The above result is akin to the Fourier slice theorem of nondiffracting tomography (§7.1). Naturally, many of the reconstruction methods developed for transmission (nondiffracting) tomography, in particular, backprojection, backpropagation and interpolation methods, can be used in the present case of reflection tomography. Additionally, there is further similarity between the transmission tomography and reflection tomography. As noted in §7.1 the projection of an object is equal to the line integral of some physical quantity (e.g., absorption in X-ray tomography) over the ray path. A similar physical insight can be given to reflection tomography.

7.4.1 <u>Line Integral</u>: Let a broadband plane wavefront be incident on a scattering object. A series of echoes will be emitted as the wavefront penetrates the object. At any time instant the receiver will receive echoes from all scattering elements which lie on a surface (see fig. 7.23). Let f(x,y) be the reflectivity function

$$p_s(\rho, \chi_0) = \int_s f(x, y) ds$$

$$= \int_0^\infty \int_0^{2\pi} f(r,\theta)\delta(\sqrt{r^2 + R^2 + 2rR\cos(\theta - \chi_0)} - \rho)rdrd\theta \quad (7.26)$$

Different variables appearing in (7.26) are illustrated in The limits on integrals suggest that the object is of infinite size. But in practice the object is finite and the transceiver is placed outside the object. To overcome this difficulty we shall assume that the reflectivity is zero outside the domain of the object.

For fixed χ_0 we have a waveform which is a function of time or ρ (=ct). We then compute a Fourier transform of the waveform $p_s(\frac{\rho}{c}, \chi_0)$. After simplification we obtain

$$P_s(k, \chi_0) =$$
$$\int_0^\infty \int_0^{2\pi} f(r,\theta)\exp(-jk(\sqrt{r^2 + R^2 + 2rR\cos(\theta - \chi_0)}))rdrd\theta \quad (7.27)$$

Under the assumption R>>r the exponent in (7.27) may be expanded in bionomial series. Retaining the first two terms in the binomial expansion

$$\sqrt{r^2 + R^2 + 2rR\cos(\theta - \chi_0)}$$
$$\approx R + r\cos((\theta - \chi_0) + \frac{r^2}{2R}(1 - \cos^2(\theta - \chi_0)))$$

we obtain

$$P_s(k, \chi_0) = e^{-jkR} \int_0^\infty \int_0^{2\pi} f(r,\theta) e^{-jk(\frac{r^2}{2R}(1-\cos^2(\theta-\chi_0)))} e^{-jkr\cos(\theta-\chi_0)} rdrd\theta$$

$$(7.28)$$

It is easy to see that

$$r\cos(\theta - \chi_0) = x\cos(\chi_0) + y\sin(\chi_0)$$

and

$$r^2(1 - \cos^2(\theta - \chi_0)) = (x\sin(\chi_0) - y\cos(\chi_0))^2$$
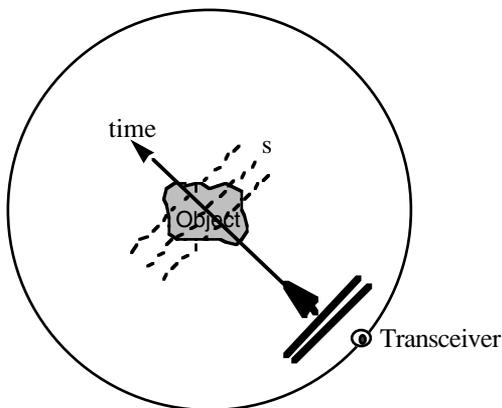
Substituting in (7.28) we obtain

Figure 7.23: At any time instant the receiver will receive echoes from all scattering elements which lie on a surface, s.
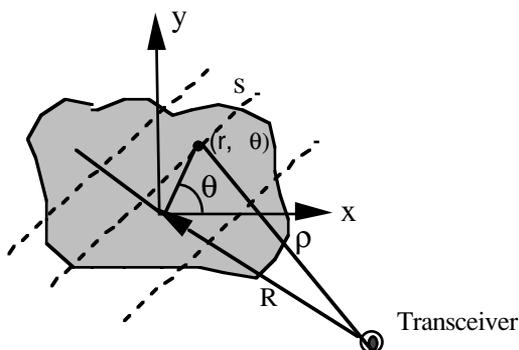

Figure 7.24: Symbols used in (7.26) are explained in the figure above.

$$P_s(k,\chi_0) = e^{-j\omega R} \times$$
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-jk\frac{(x\sin(\chi_0)-y\cos(\chi_0))^2}{2R}}e^{-jk(x\cos(\chi_0)+y\sin(\chi_0))}dxdy \qquad (7.29a)$$

If we assume that the object size is much smaller than R, the middle term in (7.29a) may be set to 1, leading to a simple Fourier transform relation,

$$P_s(k,\chi_0) = F(k\cos(\chi_0),k\sin(\chi_0)) \qquad (7.29b)$$

where F(.) is the Fourier transform of f(x,y). When this assumption is not acceptable the presence of a middle term can be accounted for through an iterative procedure described by [24].

7.4.2 Spherical Wavefronts: Plane wave illumination is not always practical. Often point sources at a finite distance are used giving rise to spherical wavefronts penetrating a scattering object. While the basic results derived in relation to a simplified plane wave model hold good, the spherical wavefront model, though mathematically more complex, yields more accurate results. We shall derive the exact result in one case where the transceivers are placed on a horizontal plane above the scattering object. Consider a uniform two dimensional array of transceivers and a three dimensional object with arbitrary speed variation (see fig. 7.25). The transceivers are sequentially fired (in any order) and all returned signals are arranged with common zero time, that is, as if all transceivers were fired at the same time instant. Interestingly in this setup it is practical to think of using a single transceiver and move it from place to place. Consider a weakly inhomogeneous (speed fluctuations only) object illuminated by a point source with the scattered field being received by a sensor kept close to the source. From every scattering volume element a backscatter radiation reaches the detector as modeled. The first order back scatter may be derived from equation (1.77), which we reproduce here for convenience,

$$f_1(\mathbf{r}, t) = \frac{1}{4\pi} \int\int\int_{-\infty}^{+\infty} \frac{2k_0^2 \delta\tilde{c} e^{j(k_0|\mathbf{r}-\mathbf{r}'|)}}{|\mathbf{r}-\mathbf{r}'|} f_0(\mathbf{r}', t) dx' dy' dz'$$

where $f_0(\mathbf{r}', t)$ is an illuminating wavefield, which, for a point source emitting a sinusoid, is given by

$$f_0(\mathbf{r}', t) = \frac{e^{j(k_0|\mathbf{r}_s - \mathbf{r}'|)}}{|\mathbf{r}_s - \mathbf{r}'|} e^{-j\omega_0 t}$$

where $\mathbf{r}_s$ is source position vector. Since the source and detector are at the same location, $\mathbf{r} = \mathbf{r}_s$. Using the point source illumination expression given above in (1.77) we obtain

$$f_1(\mathbf{r}, t) = \frac{1}{4\pi} e^{-j\omega_0 t} \int\int\int_{-\infty}^{+\infty} \frac{2k_0^2 \delta\tilde{c} e^{j2(k_0|\mathbf{r}-\mathbf{r}'|)}}{|\mathbf{r}-\mathbf{r}'|^2} dx' dy' dz' \quad (7.30a)$$

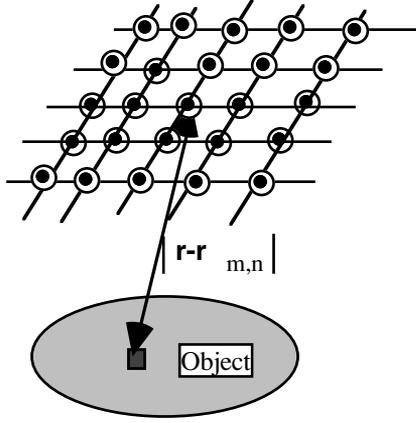The scattered field on a z=0 surface in rectangular coordinates may be written as

Figure 7.25: Uniform planar array of transceivers above a three dimensional object.

$$f_1(x,y,z=0,\omega_0)=$$

$$\frac{1}{4\pi}\int\int\int_{-\infty}^{+\infty}\frac{2k_0^2\delta\tilde{c}(x',y',z')e^{j2\left(k_0\left[(x-x')^2+(y-y')^2+(z-z')^2\right]^{\frac{1}{2}}\right)}}{\left[(x-x')^2+(y-y')^2+(z')^2\right]}dx'dy'dz' \quad (7.30b)$$

which we shall express in a form that enables us to use a result (1.78), derived in chapter 1 on page 61,

$$\frac{\partial}{\partial\omega_0}\left(\frac{f_1(x,y,z=0,\omega_0)}{\omega_0^2}\right)=$$

$$\frac{j}{\pi c^3}\int\int\int_{-\infty}^{+\infty}\frac{\delta\tilde{c}(x',y',z')e^{j2\left(k_0\left[(x-x')^2+(y-y')^2+(z-z')^2\right]^{\frac{1}{2}}\right)}}{\left[(x-x')^2+(y-y')^2+(z')^2\right]^{\frac{1}{2}}}dx'dy'dz' \quad (7.31)$$

Now using the result (1.78) in (7.31) we go into the frequency domain,

$$\frac{\partial}{\partial\omega_0}\left(\frac{f_1(x,y,z=0,\omega_0)}{\omega_0^2}\right)=$$

$$\frac{1}{2\pi^2c^3}\int\int_{-\infty}^{+\infty}\frac{\Delta\tilde{c}(u,v,\sqrt{4k_0^2-u^2-v^2})}{\sqrt{4k_0^2-u^2-v^2}}e^{j(ux+vy)}dudv \quad (7.32)$$

From (7.32) we can get the Fourier transform of the speed fluctuations in terms of the Fourier transform of the wavefield observed on the surface,

$$\Delta \tilde{c}(u, v, \sqrt{4k_0^2 - u^2 - v^2}) =$$
$$\frac{c^3}{2} \sqrt{4k_0^2 - u^2 - v^2} \; \frac{\partial}{\partial \omega_0} \left(\frac{F_1(u, v, \omega_0)}{\omega_0^2} \right) \tag{7.33}$$

Thus, the Fourier transform of the speed variations is derived from the Fourier transform of the backscatter measured on a plane surface. It is interesting to observe that the Fourier transform thus computed actually corresponds to the Fourier transform of the object on a sphere centered at the origin and with radius equal to $2k_0$ (see fig. 7.26). A broadband signal will be necessary to cover the entire Fourier transform of the speed variation function.

§7.5 Object Shape Estimation:

If the boundary of an object is piecewise linear, the corner points are sufficient for pattern recognition, image compression and coding, shape analysis, etc., [25]. The corner detection algorithms work on spatial image data in the form of a photograph. The sensor arrays are used for corner detection from the scattered wavefield (acoustic or electromagnetic). When an object, whose refractive index is slightly different with respect to that of the surrounding medium, is illuminated with a plane wave the scattered field measured around the object is proportional to the Fourier transform of the object. Thus, the shape information is buried in the scattered field. It is of some interest in medical diagnosis, in subsurface imaging, and in nondestructive testing to be able to recognize the shape of the buried object from the scattered acoustic or electromagnetic field, particularly when only a few limited views are permitted. We shall show that when the object is binary, convex and having a nondegenerate polygonal cross section, the scattered field is a sum of sinusoids, a function of wave number and corners of the polygon. The object is illuminated with a broadband plane wave and the scattered field is measured as a function of wavenumber. The frequencies of the sinusoids are estimated from the scattered field using an algorithm described in [20].

It is shown in §7.2 that the scattered field measured by a circular array is proportional to the 2D Fourier transform of the object profile taken on the circumference of a circle of radius equal to the wave number and centered at $(-k_0 \cos\chi_0, -k_0 \sin\chi_0)$ where χ_0 is angle illumination (see fig. 7.11). By changing the direction of illumination (0 to 360°) the object Fourier transform is scanned over a disk of radius k_0. When the object is binary (i.e., refractive index is constant throughout the object) the interest is in the shape of the object. The shape information may be directly obtained from the scattered field
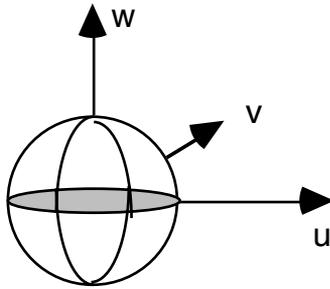
Figure 7.26: The Fourier transform of a reflected signal (echo) corresponds to the Fourier transform of the object on a sphere centered at the origin and radius equal to $2k_0$.

or 2D Fourier transform of the object. This approach was taken by Milanfar and co-workers [26] in the context of ray tomography where the input data are projections of object.

7.5.1 <u>Fourier Transform of Binary Convex Polygonal Object</u>: Consider the evaluation of the 2D Fourier transform over a p-sided binary and convex polygonal domain (see fig. 7.27). Take any point inside the polygon and join it to all corners forming p triangles which lie entirely inside the polygon and make this as the origin of the coordinate system

$$
\begin{aligned}
F(u,v) &= \iint\limits_{over\ polygon} e^{j(ux+vy)}dxdy \\
&= \sum_{n=1}^{p} \iint\limits_{over\ n^{th}\ triangle} e^{j(ux+vy)}dxdy
\end{aligned}
\tag{7.34}
$$

To evaluate the integral over n^{th} triangle refer to fig. 7.27b where we show the integration along a narrow strip under the rotated coordinate system such that the new the x-axis is perpendicular to n^{th} side. Note that this is valid only for convex objects (for nonconvex objects, it's not possible to drop a perpendicular from the origin to at least one edge, such that the perpendicular lies entirely within the object) that are nondegenerate. The triangle is then covered by a series of strips. Equation (7.34) reduces to

$$
F(u,v) = \sum_{n=1}^{p} \int_{0}^{\rho_n} \int_{-x'\tan\phi_{2n}}^{x'\tan\phi_{1n}} e^{\left\{\begin{array}{l} -j((u\cos\theta_n + v\sin\theta_n)x' \\ +(v\cos\theta_n - u\sin\theta_n)y') \end{array}\right\}} dx'dy'
\tag{7.35}
$$

where $x'=(x\cos\theta_k+y\sin\theta_k)$ and $y'=(y\cos\theta_k-x\sin\theta_k)$. Evaluate the integral in (7.35) first with respect to y' followed by integration with respect to x'. We obtain

$$F(u,v) = \sum_{n=1}^{p} \frac{[e^{-j(u'-v'\tan\phi_{1n})\rho_n} - 1]}{v'(u' - v'\tan\phi_{1n})} - \frac{[e^{-j(u'+v'\tan\phi_{2n})\rho_n} - 1]}{v'(u' + v'\tan\phi_{2n})} \qquad (7.36)$$

where $u' = (u\cos\theta_n + v\sin\theta_n)$ and $v' = (v\cos\theta_n - u\sin\theta_n)$. We shall now rewrite (7.36) by replacing θ_n, ϕ_{1n}, and ϕ_{2n} in terms of the coordinates of the two corners corresponding to the nth side, namely, (a_n,b_n) and (a_{n+1},b_{n+1}). The following relations are used for this purpose:

$$\rho_n = a_n\cos\theta_n + b_n\sin\theta_n = a_{n+1}\cos\theta_n + b_{n+1}\sin\theta_n$$

$$a_n = \rho_n(\cos\theta_n + \sin\theta_n \tan\phi_{1n})$$

$$b_n = \rho_n(\sin\theta_n - \cos\theta_n \tan\phi_{1n})$$

$$a_{n+1} = \rho_n(\cos\theta_n - \sin\theta_n \tan\phi_{2n})$$

$$b_{n+1} = \rho_n(\sin\theta_n + \cos\theta_n \tan\phi_{2n})$$

We obtain

$$F(u,v) = \sum_{n=1}^{p} \rho_n \left[\frac{[e^{-j(ua_n + vb_n)} - 1]}{v'(ua_n + vb_n)} - \frac{[e^{-j(ua_{n+1} + vb_{n+1})} - 1]}{v'(ua_{n+1} + vb_{n+1})} \right] \qquad (7.37)$$

Our goal is to determine (a_n,b_n) and (a_{n+1},b_{n+1}) from (7.37). This may be achieved by expressing (7.37) on the $k_y = 0$ and $k_x = 0$ axes. We get the following equations
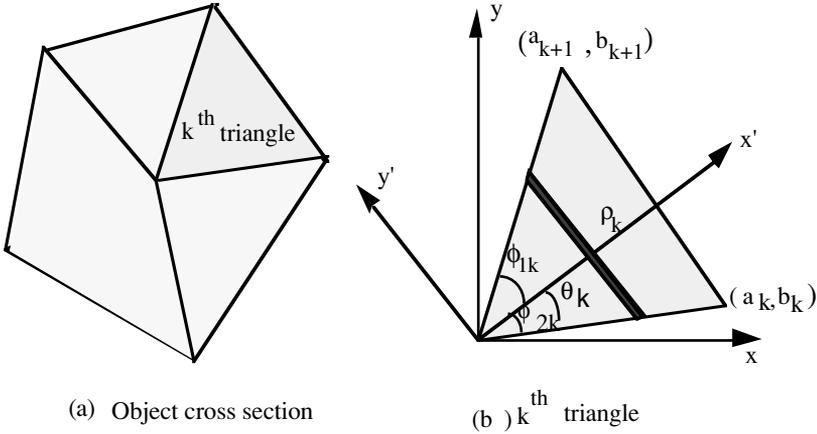
Figure 7.27: To evaluate the Fourier transform of a polygonal object we consider each triangle (From [27] with permission from IEEE (© 1998 IEEE).)

$$u^2 F(u, v=0) = -\sum_{n=1}^{p} \rho_n \{\frac{e^{-jua_n}-1}{a_n \sin\theta_n} - \frac{e^{-jua_{n+1}}-1}{a_{n+1}\sin\theta_n}\}$$

$$v^2 F(u=0, v) = \sum_{n=1}^{p} \rho_n \{\frac{e^{-jvb_n}-1}{b_n \sin\theta_n} - \frac{e^{-jvb_{n+1}}-1}{b_{n+1}\sin\theta_n}\}$$

(7.38)

The above equations may be solved by modeling them as a sum of sinusoids and using the well known Prony's algorithm or its more modern versions [20]. From the coefficients in the exponents of the complex sinusoids we obtain (a_n, a_{n+1}) and (b_n, b_{n+1}) but we are yet to pair them, that is, select the right x and y coordinate pair which will form a valid corner. We note that $F(u, v=0)$ and $F(u=0, v)$ represent backscatter due to a broadband illumination along the x- and y-axes, respectively (see (7.22b)).

7.5.2 <u>Pairing Algorithm</u>: In the previous section we saw how to obtain the x- and y-coordinates of the corners of the polygon. This alone will not suffice to define a unique convex polygon. We need some additional information on how to pair a given x-coordinate with the right y-coordinate from the list of estimated y-coordinates. This problem is resolved by using an additional illumination at an angle θ

$$k^2 F(k\cos\theta, k\sin\theta) =$$

$$\sum_{n=1}^{p} \Gamma_n \left\{ \begin{array}{c} \dfrac{[e^{-jk(a_n \cos\theta + b_n \sin\theta)} - 1]}{(a_n \cos\theta + b_n \sin\theta)} \\ -\dfrac{[e^{-jk(a_{n+1} \cos\theta + b_{n+1} \sin\theta)} - 1]}{(a_{n+1} \cos\theta + b_{n+1} \sin\theta)} \end{array} \right\} \qquad (7.39)$$

where $\Gamma_n = \dfrac{\rho_n}{\sin(\theta - \theta_n)}$. From the back scatter due to an illumination at angle

$\theta \ (\neq 0 \ or \ \dfrac{\pi}{2})$, we can estimate as described in [20, 27] the coefficients in the

exponents of the complex sinusoids. Thus, we get the additional information in the form of linear combination of the x- and y-coordinates of the corners, $(a_n \cos\theta + b_n \sin\theta)$, $n = 1, 2, \ldots p$. The steps in the pairing algorithm are as below:

1) Generate a list of x-coordinates, y-coordinates and the linear combination of the x- and y-coordinates. It is presumed that the list is not in the same order as the indexed corners.

2) Take the first element from the x-coordinate list and any one element from the y-coordinate list and form a linear combination, $(a_1 \cos\theta + b_n \sin\theta)$

$n = 1, 2, \ldots p$

3) Compare the result of the linear combination with those estimated with

$\theta \ (\neq 0 \ or \ \dfrac{\pi}{2})$ illumination. The best match (within the limits of estimation

error) will indicate the correct choice of b_n.

4) Take the next element from the x-coordinate list and go to step (2).

For the purpose of illustration we consider a square object of size (6m,6m), rotated by 30 deg. and shifted away from the origin by (5m,5m). It is illuminated from three directions, 0, π and $\pi/6$. The x- and y-coordinates got from the noiseless scattered field in the first two directions and their linear combination (θ=30 deg) are shown in table 7.4 and those estimated from the scattered field got in the third direction are shown in table 7.5. The application of the pairing algorithm is illustrated in table 7.6. The best match with the estimated coefficients is shown in column three in bold figures and the corresponding y-coordinate is shown in the last column.

We may encounter the problem of repeated x- or y-coordinates or their projections. The projections of two corners may overlap or come very close to each other depending upon the orientation of the object. As shown in fig. 7.28, for a square object depending upon the orientation, the adjacent projections (e.g., x_1 and x_2) may come close to each other or overlap. The problem of

x	y	$x\cos\theta + y\sin\theta$
9.0981	6.0981	10.9282
3.9019	9.0981	7.9282
0.9019	3.9019	2.7321
6.0981	0.9019	5.7321

Table 7.4: The x- and y-coordinates and their linear combination (θ=30 deg) are shown in the above table.(Source: [27] with permission from IEEE (© 1998 IEEE).)

x	9.0981, 6.0981, 3.9019, 0.9019
y	9.0981, 6.0981, 3.9019, 0.9019
$x\cos\theta + y\sin\theta$	10.9282, 5.7321, 7.9282, 2.7321

Table 7.5: The estimated projections from the scattered field (noise free) are shown (Source: [27] with permission from IEEE (© 1998 IEEE).)

x	y	$x\cos\theta + y\sin\theta$	Best match for x
9.0981	9.0981	12.4282	6.0981
	6.0981	**10.9282**	
	3.9019	9.8302	
	0.9019	8.3301	
6.0981	9.0981	9.8302	0.9019
	6.0981	8.3301	
	3.9019	7.2321	
	0.9019	**5.7321**	
3.9019	9.0981	**7.9282**	9.0981
	6.0981	6.4290	
	3.9019	5.3301	
	0.9019	3.8302	
0.9019	9.0981	5.3301	3.9019
	6.0981	3.8302	
	3.9019	**2.7321**	
	0.9019	0.5000	

Table 7.6: A numerical illustration of the pairing algorithm. Q=30 deg.(Source: [27] with permission from IEEE (© 1998 IEEE).)

repeated projection can be resolved by selecting another direction of illumination whenever the number of sinusoids estimated in x, y differs. The
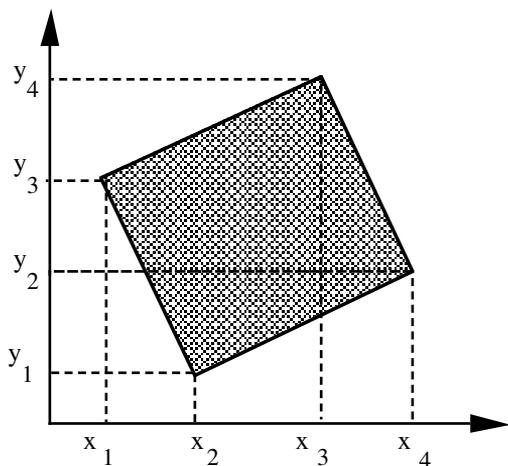
Figure 7.28: A square object and the projections of its corners on the x- and y-axes. Note that, depending upon the orientation, the adjacent projections (e.g., x_1 and x_2) may come close to each other(From [27] with permission from IEEE (© 1998 IEEE).).

number of sinusoids observed in all three illuminations must be equal. In practice it may be necessary to illuminate an unknown object along several directions and estimate the sinusoids along each direction. From this set choose three directions, preferably two orthogonal directions, having an equal number of sinusoids. The number of sinusoids which may be determined from a finite data in the presence of noise is indeed a complex problem and hence it is outside the scope of this book. The reader may like to refer to a book, for example [20], or current literature on this topic.

When there are two or more objects, as the center of coordinates would lie outside all but one object, it is necessary to modify (7.31), which was derived under the assumption that the center lies inside the object. Also, there may be some ambiguity in the process of constructing the object shape even when all x- and y-coordinates are correctly paired. The amplitude of the sinusoid corresponding to a corner can then be used to resolve such an ambiguity in addition to the fact that the objects are convex and the number of objects is known (see [27] for more details).

7.5.3 <u>Performance Analysis</u>: The performance of the object shape estimation procedure has been investigated through numerical experiments [27]. For this we have considered a square object (refractive index contrast equal to 0.01) of size $8x8m^2$ and rotated by 9 degrees with respect to the x-axis (see fig. 7.28). The object was illuminated with a broadband plane wavefront whose wave
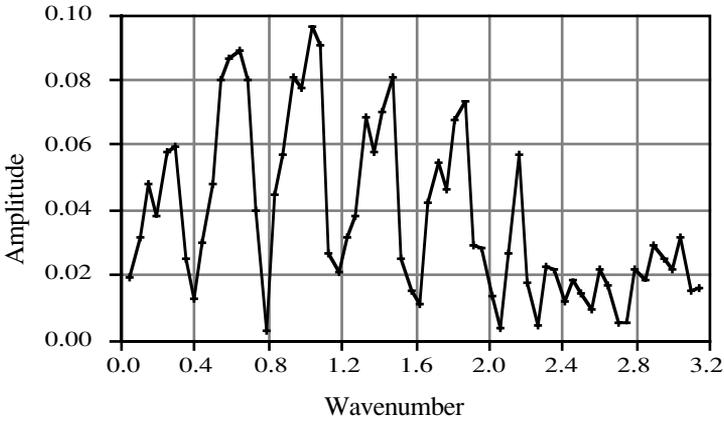
Figure 7.29: Computed backscatter from a weakly scattering square object (shown in fig 7.28) illuminated by a broadband signal.(From [27] with permission from IEEE (© 1998 IEEE).)

number is varied from $\frac{\pi}{64}$ to π in steps of $\frac{\pi}{64}$ along three directions, namely, x-axis, y-axis and a radial direction at an angle θ=30 degrees. The backscatter at each wavenumber was computed using the Fourier transform approach described in [1]. A typical example of backscatter caused by x-axis illumination is shown in fig. 7.29. To this scattered field sufficient white Gaussian noise was added so that the snr became equal to a specified figure. Here the snr is defined as ten times the logarithm (base 10) of the ratio of the average scattered energy to noise variance. The corners and also the T_n' were estimated using the procedure given in [27]. The mean square error (MSE) in the estimation of coordinates of the corners was studied as a function of snr. The results, obtained by averaging over fifty independent experiments, are shown in fig. 7.30a. Notice that MSE rises very rapidly for snr below 8 dB. This is largely on account of the fact that the projections of two adjacent corners (e.g., x_1 and x_4, and x_2 and x_3 in fig. 7.29) are close to each other; in this example they are 1.2515 meters apart. For a different orientation, say, at 6° when the separation becomes 0.8362 meters, the MSE rapidly rises for snr below 15 dB. The estimation error (MSE) also depends upon the number of corners in a polygonal object. The numerical results are shown in fig 7.30b.

§7.6 **Exercises**:
1. What is the essential difference between an array of sensors used for DOA estimation (chapters 2 and 5) or for signal waveform estimation (chapters 3 and 6) and the array used in nondiffracting radiation tomography?
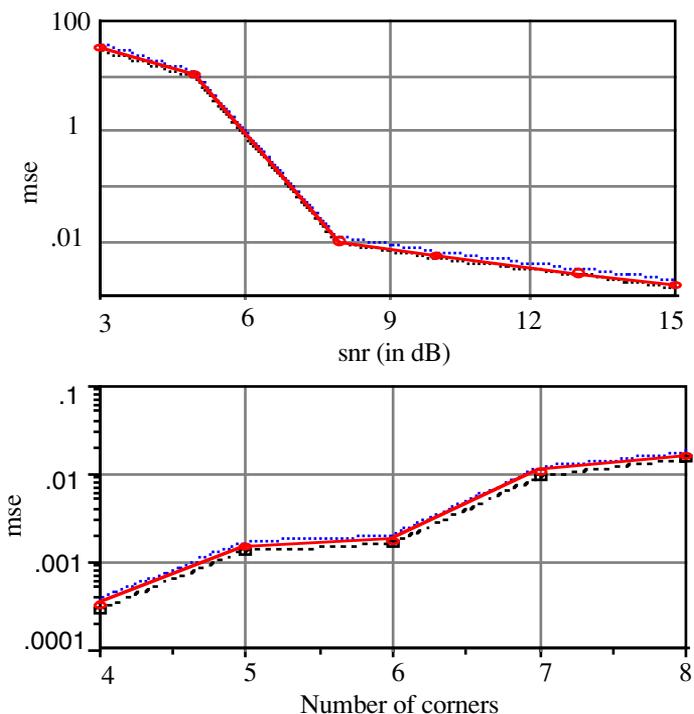
Figure 7.30: (a) Mean square error (m^2) in the estimation x- and y-coordinates as a function snr. (b) Mean square error (m^2) as a function of the number of corners (snr=20dB). The dotted lines show the error bounds for 95% confidence.

2. Show that the filter function used in the backpropagation algorithm, equation (7.14c), reduces to that used in the backprojection algorithm (page 368) for the radiation of wave length much smaller than the scale of speed/density fluctuations.

3. A rectangular object (2D) is illuminated by a plane wavefront as shown in fig. 7.31 below. An array of sensors measures the relative delays. Assume that the speed of propagation inside the object is constant and that there are no diffraction effects. (a) Sketch the delay profile (delay as a function of sensor position) for different object orientations, $\theta=0^o$ (as seen in the figure above), $\theta=45^o$, and $\theta=90^o$. (b) How would you estimate the size of the object?

4. Consider a simple arrangement of sources and sensors in two borehole as shown in fig. 7.32. Compute the sampling points in the Fourier plane. See [28] on how to compute the sampling points in a more general case.
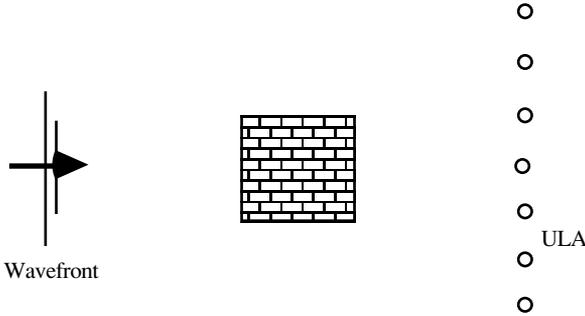
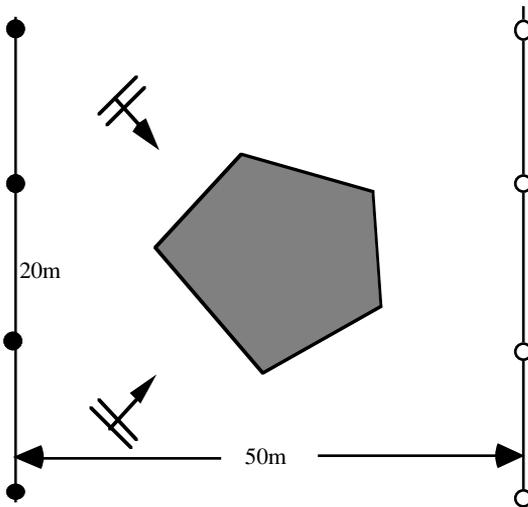Figure 7.31: A rectangular object (2D) is illuminated by a plane wavefront

Figure 7.32: A cross borehole tomography experimental setup.

References

1. A. C. Kak, Tomographic Imaging with diffracting and non diffracting sources, *in Array Signal Processing*, Ed. S. Haykin, Prentice-Hall, Englewood Cliffs, NJ, pp. 351-428, 1985.
2. S. Ivansson, Seismic borehole tomography- theory and computational methods, Proc. of IEEE, vol. 74, pp. 328-338, 1986.
3. M. Gustavsson, S. Ivansson, P. Moren, and J. Pihl, Seismic borehole tomography- measurement system and field studies, Proc. IEEE, vol. 74, pp. 339-346, 1986.

4. K. Iwata and R. Nagata, Calculation of refractive index distribution from interferograms using the Born and Rytov's approximation, Japan, J. Appl. Phy., vol. 14, Suppl. 14-1, pp. 379-383, 1974.

5. R. K. Mueller, M. Kaveh, and G. Wade, Reconstructive tomography and applications to ultrasonics, Proc. IEEE, vol. 67, pp. 567-587, 1979.

6. E. Wolf, Three dimensional structure determination of semi-transparent objects from holographic data, Optics Cummunication, vol. 1, pp. 153-156, 1969.

7. A. J. Devaney, Diffraction Tomography, *in* Inverse Methods in Electromagnetic Imaging, Part 2, W. M. Boerner et. al. (Eds), pp. 1107-1135 D. Reidel Pub Co., 1985.

8. A. J. Devaney, Geophysical diffraction tomography, IEEE Trans., vol. GE-22, pp. 3-13, 1984.

9. Prabhakar S. Naidu, A. Vasuki, P. Satyamurthy, and L. Anand, Diffraction tomographic imaging with circular arra, IEEE Trans., UFFC-42, pp. 787-789, 1995.

10. R. Porter, Determination of structure of weak scatterers from holographic images, Opt. Commun. vol. 39, pp. 362-364, 1981.

11. A. J. Devaney and G. Beylkin, Diffraction tomography using arbitrary transmitter and receiver surfaces, Ultrasonic Imaging, vol. 6, pp. 181-193 1984.

12. S. J. Norton, Tomographic reconstruction of acoustic reflectivity using circular array: Exact solution, J. Acoust. Soc. Amer. vol 67, pp. 1266-1273, 1980.

13. S. J. Norton and M. Linzer, Ultrasonic reflectivity tomography with circular transducer arrays, Ultrasonic Imaging, vol. 1, pp. 154-184, 1979.

14. Z. Qin, et al., Circular array ultrasound holographic imaging using the linear array approach, IEEE Trans., vol. UFFC-36, pp. 485, 1989.

15. J. M. Ruis, M. Ferrando, L. Jofre, E. De Los Reyes, A. Elias, and A. Broquet, Microwave tomography: An algorithm for cylindrical geometries, Electronics Letters, vol. 23, pp. 564-565, 1987.

16. P. Tamarkin, Scattering of an underwater ultrasonic beam from liquid cylindrical obstacles, J. Acoust. Soc. Am., vol. 21, pp. 612-616, 1949.

17. S. J. Bezuszka, Scattering of underwater plane ultrasonic waves by liquid cylindrical obstacles, J. Acoust. Soc. Am., vol. 25, pp. 1090-1095, 1953.

18. A. Vasuki and P. S. Naidu, Broadband tomographic imaging with circular array, Acoustic Letters (UK), vol. 21, pp. 144-150, 1998.

19. J. L. Sanz, On the reconstruction of band-limited multidimensional signals from algebraic contours, IBM Research Report, RJ 4351, (47429), 1984.

20. P. S. Naidu, Modern Spectrum Analysis of Time Series, CRC Press, Boca Raton, Fl, 1996.

21. Tien-when Lo, G. L. Duckworth, and M. N. Toksoz, Minimum cross entropy seismic diffraction tomography, J. Acoust. Soc. of Am., vol. 87, pp. 748-756, 1990.

22. A. Mohammad-Djafari and G. Demoment, Maximum entropy Fourier synthesis with application to diffraction tomopraphy, Applied Optics, vol. 26, pp. 1745-1754, 1987.

23. J. P. Burg, Maximum entropy spectrum analysis, Presented at the International Meeting of Exploration Geophysicists, Orlando, Fl, 1967.

24. C. Q. Lang and W. Xiong, An iterative method of ultrasonic reflection mode tomography, IEEE Trans on Medical Imaging, vol. 13, pp. 419-425, 1994.

25. G. Medioni and Y. Yasumoto, Corner detection and curve representation using cubic B-spline,'Computer Vision Graphics and Image Processing, Vol 39, pp. 267-278, 1987.

26 P. Milanfar, W. C. Karl and A. S. Willsky, Reconstructing binary polygonal objects from projections: A statistical view. Graphical Models and Image Processing, vol. 56, pp. 371-391, 1994.

27. A. Buvaneswari and P. S. Naidu, Estimation of shape of binary polygonal object from scattered field, IEEE Trans. on Med Imaging, vol. 7, pp. 253-257, 1998.

28. D. Nahamoo, S. X. Pan, and A. C. Kak, Synthetic aperture diffraction tomography and its interpolation free computer implementation, IEEE Trans., vol. SU-31, pp. 218-226, 1984.

Chapter Eight

Imaging by Wavefield Extrapolation

We have seen in chapter 2 that the wavefield on a horizontal plane can be derived from the wavefield observed on another horizontal plane given that the space lying between the two planes is homogeneous and free from any sources. This operation is called extrapolation of wavefield; forward extrapolation is when we go away from the sources and backward extrapolation is when we go toward the sources (see fig. 8.1). Wavefield extrapolation enables us to map the source distribution provided it is known that all sources are confined to a layer. This problem is known as inverse source problem [1]. A slightly different situation arises in the scattering problem. An external source induces a field on the surface of a scatterer which in turn will radiate wavefield, known as scattered field, back into the space. This scattered field contains information about the scatterers. The inverse scattering problem pertains to extraction of information about the scatterers from the scattered field. The tomographic imaging covered in chapter 7 falls in the realm of the inverse scattering problem. In the present chapter we seek a means of reconstructing a layered (but not necessarily horizontally layered) medium using the reflected wavefield. This problem is of great significance in seismic exploration where it is commonly known as migration. An image of subsurface reflectors can also be achieved through focused beamformation, which gives an estimate of the reflected energy received from a subsurface point. The focused beamformation is based on a ray theoretic description of the wavefield, as in optics, but the migration is based on diffraction properties of the wavefield. Both approaches lead to similar results. For imaging, an essential input is the wave speed which, fortunately, has to be estimated from the observed wavefield only.

§8.1 **Migration**:
The interface between two homogeneous layers may be considered as a thin layer of point sources (scatterers). This forms the basis of the exploding reflector model [2]. The wavefield observed on the surface of earth can be extrapolated downward into the earth. The interfaces separating homogeneous layers reflect or scatter wave energy. Such an interface may be modeled as a surface with point scatterers whose density is proportional to the impedance discontinuity. To image an interface, that is, to map the impedance discontinuity, it is necessary to compute the distribution of the wave energy on an interface. This problem has been treated as a boundary value problem [3, 4] or an initial value problem [5]. As a boundary value problem we solve the wave equation in homogeneous half space with a boundary condition that the wavefield is given on the surface of the earth (z=0).
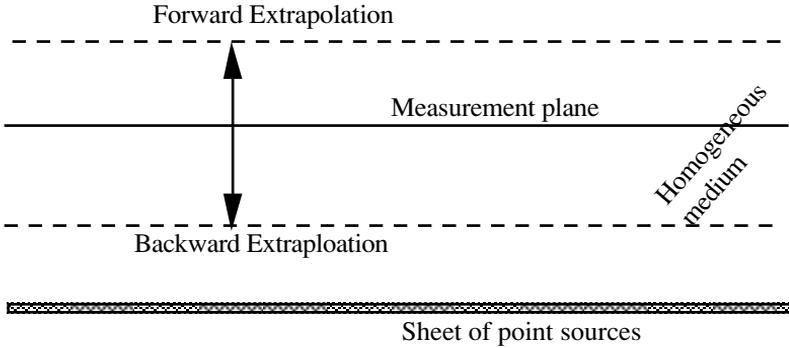
Figure 8.1: The wavefield measured on a horizontal plane can be extrapolated upward (forward extrapolation) or downward toward the source (backward extrapolation).

The boundary value problem has been solved in the time domain by Claerbout [3] by solving a finite difference equation with a time varying boundary condition and in the frequency domain by Stolt [4] by expressing the extrapolation as a filtering problem. Extrapolation in the z direction may also be expressed as propagation in backward time. Note that in the wave equation the double derivative with respect to z differs from the double derivative with respect to time only in a scale factor given by $-c^2$. Thus, starting from some time instant, the wavefield observed at the surface (z=0) is propagated backward in time, a process that is equivalent to extrapolation of the wavefield to a lower level (z<0). The boundary value problem may be reformulated as a source problem with zero boundary condition but driven by an external source which is given as a time reversed output of each receiver [6]. In another approach extrapolation is posed as an initial value problem but marching backward [5]. In this approach the time axis is scaled by the wave speed which converts a recorded seismic section into a wavefield throughout the subspace as it might have appeared at the latest recording time. Thus the converted wavefield is next propagated backward in time.

8.1.1 <u>Imaging Conditions</u>: Imaging requires two steps, namely, (i) extrapolation in space or reverse propagation in time and (ii) an imaging condition, that is, how to decide when an image has been formed. In optical imaging convergence of all rays emerging from a point to another point (image point) is the imaging condition. In seismic or acoustic imaging, the imaging condition commonly used is when the depropagated field reaches the starting time which is the time when the scatterer was illuminated or excited. This information can be found given the wave speed and the distance from the source [see fig. 8.2]. It is also possible to set the excitation time to zero provided the scattered wavefront and illuminating wavefront travel along the same path but
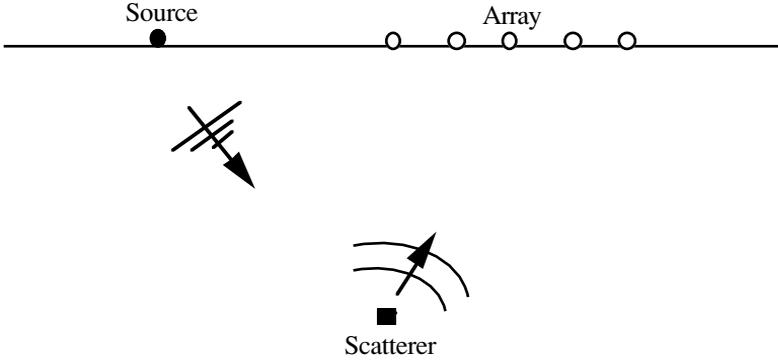
Figure 8.2: Depropagate a wavefront to a time instant when the illuminating wavefront hits the scatterer. When this happens, the scattered wave field is found to be concentrated around the scattering point.

in opposite directions. This forms the basis of the popular migration principle called exploding reflector which we shall describe in detail in the next section. Qualitatively speaking, imaging is focusing of wave energy. An imaging condition based on how well the wave energy is focused at a point is also a likely candidate as an imaging condition. Indeed, in seismic imaging, it has been suggested that when p-waves and s-waves are focused at the same point, an image of the point is obtained [7].

8.1.2 <u>Downward Continuation of Sources and Sensors</u>: The source and sensor arrays are normally placed on the same surface. It is possible to analytically compute the field when both source and sensor arrays are relocated onto another plane given the field on the observation plane. For simplicity we shall assume that both the observation plane and the plane onto which the source and sensor arrays are relocated are horizontal. We are already familiar with continuation of wavefield from one plane to another (see chapter 1 and also later in this chapter). Continuation of a source array requires an additional concept of reciprocity which states that when the positions of an omnidirectional source and an omnidirectional sensor are interchanged the observed field remains unchanged [8]. To extrapolate the source array keeping the sensor array fixed we need to interchange the source and the sensor arrays and then apply the wavefield extrapolation algorithm. By virtue of the principle of reciprocity the result of the above approach will be same as that of actual relocation of the source array. In actual application the source and the sensor arrays are relocated alternatively in small steps. As the source and the sensor arrays are continued downwards towards a reflector at some stage the two arrays will completely coincide when they reach the reflector after a lapse of time equal to the one way travel time. Occurrence of such a coincidence of source and sensor may be used as a condition for imaging. This phenomenon is illustrated in fig. 8.3.
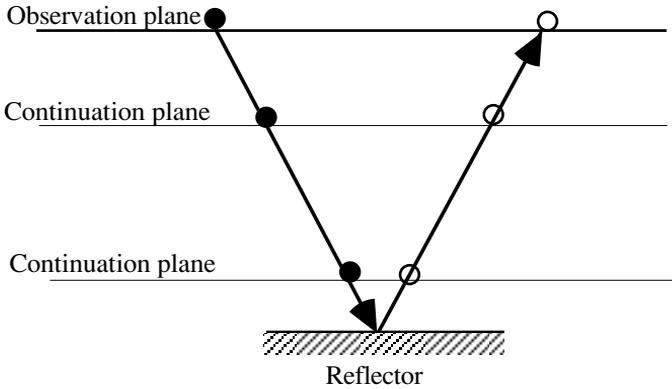
Figure 8.3: The wavefield measured in one plane is continued to another plane as if the source and the sensor arrays are located on that plane. It may be recalled that as the wavefield is continued the wave energy actually propagate along a ray.

§8.2 Exploding Reflector Model:

The wavefield, measured on a horizontal surface, is propagated backward in time in order to arrive at the reflector location giving an estimate of the wavefield as existing at the reflector. The seismic traces are first stacked with proper delays so that the output corresponds to a hypothetical sensor kept close to the source. A set of such stacked seismic traces may be modeled as a wavefield observed in an imaginary experiment in which small charges are placed on a reflector and all of them are exploded at the *same* time instant. The wavefield is assumed to propagate upwards and reach the surface at time t_0. Conversely, the wavefield when propagated backwards will reach the reflector point after t_0 time units. It is assumed that there are no multiple reflections, surface waves, refractions, etc. Indeed, during the process of stacking, since the array is focused downward, much of the interference would be attenuated. The exploding reflector model, also known as the Loewenthal model [2], consists of tiny charges placed on a reflector and fired at the same time instant (fig. 8.4). The quantity of charge placed at a given point on the interface is proportional to the reflection coefficient at that point.

Let the interface be described by a function, $z = g(x, y)$, where the acoustic impedances above and below the interface are constant but different (see fig. 8.4). The reflection coefficient for vertical incidence is given by $r_0 = \dfrac{(\rho_2 c_2 - \rho_1 c_1)}{(\rho_2 c_2 + \rho_1 c_1)}$. The reflectivity function may be written as
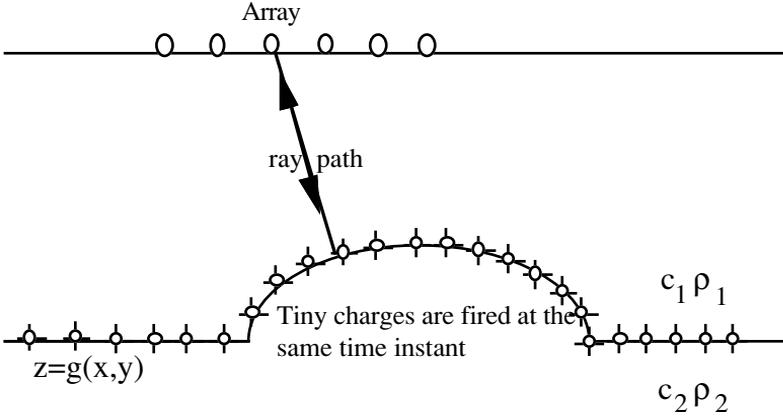
Figure 8.4: Exploding reflector model, also known as the Loewenthal model. Tiny charges are placed on a reflector and fired at the same time instant. The ray paths are perpendicular to the reflector. The wavefront, at the time of firing, that is, t=0, coincides with the reflector.

$$r(x,y,z) = r_0 \delta(z - g(x,y)) \qquad (8.1)$$

The wavefield in a medium bounded by the earth's surface above and the interface below will satisfy the following boundary condition and initial conditions:

Boundary condition: $f(x,y,z,t)\big|_{z=0} = f_0(x,y,t)$, that is, pressure field observed on the surface.

Initial condition: $f(x,y,z,t)\big|_{t=0} = r(x,y,z)\delta(t)$, that is, pressure field generated by the exploding charges on the interfaces.

The wavefront at time instant t=0 is the interface itself. As the time progresses the wavefront travels upward toward the surface. The wavefield observed at the surface acts as a boundary condition, and the shape of the wavefront at t=0 is the initial condition which, in practice, is not known. The boundary condition (when given over an infinite plane) is enough to solve the wave equation. The wavefield thus obtained at all points within the space bounded from above by the observation plane and from below by the initial wavefront and for all t in the range $0 \le t < \infty$. Observe that the wavefield at t=0 is actually the pressure field generated by setting off charges on the interface and everywhere else it is zero. Conversely, if the wavefield observed on the surface is propagated

backward in time till we reach the time instant t=0 we shall indeed reach the surface where the charges were set off. This is the rationale for imaging a reflector by propagation backward in time.

8.2.1 Initial Value Problem: Alternatively, an equivalent description of imaging is through a solution of initial value problem which may also be expressed as an inverse source problem [9]. The wave equation with the right hand side equal to a source creating a wavefield is

$$\nabla^2 f = \frac{1}{c_0^2}\frac{d^2 f}{dt^2} + r(x,y,z)\delta'(t)$$

where r(x,y,z) as before stands for the reflectivity function and $\delta'(t)$ is the derivative of $\delta(t)$. At each point an explosive charge proportional to r(x,y,z) is set off at t=0. The waves propagate unhindered by other reflecting interfaces (no multiple reflections). The solution of the inhomogeneous wave equation on the z=0 surface is given by

$$f(x,y,z=0,\omega)$$

$$= \frac{j\omega}{4\pi}\int\limits_{-\infty}^{+\infty}\!\!\!\int\!\!\!\int r(x',y',z')\frac{e^{jk\sqrt{(x-x')^2+(y-y')^2+(z')^2}}}{\sqrt{(x-x')^2+(y-y')^2+(z')^2}}dx'dy'dz' \quad (8.2)$$

which may be further simplified following the procedure used in obtaining (1.81)

$$f(x,y,z=0,\omega) = \frac{\omega}{4\pi^2}\int\limits_{-\infty}^{+\infty}\!\!\!\int \frac{R(u,v,w)}{w}e^{j(ux+vy)}dudv \quad (8.3a)$$

$$F_0(u,v,\omega) = \omega\frac{R(u,v,w)}{w} \quad (8.3b)$$

where $w = -sgn(\omega)\sqrt{k^2 - u^2 - v^2}$ for upgoing waves. From (8.3b) we can obtain the unknown reflectivity from the surface pressure field

$$R(u,v,w) = \frac{w}{\omega}F_0(u,v,\omega) \quad (8.4)$$

Next we shall show how the same result (that is, (8.4)) can be obtained as a boundary value problem (see (8.7)).

§8.3 **Extrapolation in** $\omega-k$ **Plane:**

Recall the integral representation of a wavefield in a homogeneous medium (see chapter 1, (1.19)) which we reproduce here for quick reference,

$$f(x,y,z,t) = \frac{1}{8\pi^3} \int\limits_{-\infty}^{-\infty} \int \int F_0(u,v,\omega) e^{+j\sqrt{k^2-u^2-v^2}\,z} e^{-j(ux+vz-\omega t)} dudvd\omega$$

We have chosen the positive sign in $e^{\pm j\sqrt{k^2-u^2-v^2}\,z}$ as the wavefield is propagated from surface to the interface where charges are placed, that is, propagation is towards the source; hence, as per our convention (chapter 1), +ve is chosen.

8.3.1 <u>Downward Continuation</u>: The wavefield measured on the surface may be continued downwards to any depth and for all times (see (1.26)). Using the initial condition in the exploding charge model the wavefield at time t=0 is equal to the reflectivity function,

$$r(x,y,z) = f(x,y,z,t)\big|_{t=0}$$

$$= \frac{1}{8\pi^3} \int\limits_{-\infty}^{-\infty} \int \int F_0(u,v,\omega) e^{+j\sqrt{k^2-u^2-v^2}\,z} e^{-j(ux+vy)} dudvd\omega \qquad (8.5)$$

Note that in (8.2) $k = \dfrac{\omega}{c}$ where c is the wave speed in the medium above the interface and it is assumed to be known. Further, we relate the temporal frequency to the vertical spatial frequency, w. Since $k = \dfrac{\omega}{c} = -w\sqrt{1+\dfrac{s^2}{w^2}}$,

where $s = \sqrt{u^2 + v^2}$, we can express $w = -\dfrac{\omega}{c}\sqrt{1-(\dfrac{sc}{\omega})^2}$ for an upgoing

wave and $d\omega = -\dfrac{c}{\sqrt{1+\dfrac{s^2}{w^2}}} dw$. Using these results in (8.5) we obtain [10]

$$r(x,y,z) =$$

$$\frac{c}{8\pi^3} \int\limits_{-\infty}^{\infty} \int \int \frac{1}{\sqrt{1+\dfrac{s^2}{w^2}}} F_0(u,v,-cw\sqrt{1+\dfrac{s^2}{w^2}}) e^{-j(ux+vy+wz)} dudvdw \qquad (8.6a)$$

Equation (8.6a) is applicable only when the sources actually replace the reflecting interfaces, but in practice the wave excitation is done on the surface and the wave propagates into the medium and it is then reflected at an interface back to the surface. In this process, since the wavefield travels first down and then up, the travel time is doubled; equivalently, the wave speed may be halved. Hence, we have

$$r(x,y,z) =$$

$$\frac{c}{16\pi^3} \int\int\int_{-\infty}^{\infty} \frac{1}{\sqrt{1+\dfrac{s^2}{w^2}}} F_0(u,v,-\frac{c}{2}w\sqrt{1+\frac{s^2}{w^2}})e^{-j(ux+vy+wz)}\,dudvdw \quad (8.6b)$$

where $w = -\dfrac{2\omega}{c}\sqrt{1-(\dfrac{sc}{2\omega})^2}$. This result agrees with that given in [9].

Computing the inverse Fourier transform on both sides of (8.6) we obtain

$$R(u,v,w) = \frac{c}{\sqrt{1+\dfrac{s^2}{w^2}}} F_0(u,v,-cw\sqrt{1+\frac{s^2}{w^2}}) \quad (8.7a)$$

or

$$R(u,v,w) = \frac{1}{\sqrt{1+\dfrac{s^2}{w^2}}}\frac{c}{2} F_0(u,v,-\frac{c}{2}w\sqrt{1+\frac{s^2}{w^2}}) \quad (8.7b)$$

The wavefield in a homogeneous (also in horizontally layered) medium has a radial symmetry in the (x,y) plane. For this case the appropriate extrapolation equation in polar coordinates is (1.27b), reproduced here for convenience,

$$f(r,z,t) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty} F_0(s,\omega)e^{j\omega t}\,d\omega \int_0^{\infty} sJ_0(sr)e^{\pm j(\sqrt{k^2-s^2}\,z)}\,ds$$

where $F_0(s,\omega)$ is the Fourier transform of the surface wavefield having a radial symmetry. Let us rewrite (1.27) in terms of plane wave decomposition

(PWD). We map s domain to γ domain where $s = k \sin(\gamma)$ and rewrite (1.27b) as

$$f(r,z,t) =$$

$$\frac{1}{8\pi^2} \int\limits_{-\infty}^{\infty} k^2 e^{j\omega t} d\omega \int\limits_{0}^{\infty} F_0(\sin(\gamma),\omega) e^{\pm j(kz\cos(\gamma))} \cos(2\gamma) J_0(k\sin(\gamma)r) d\gamma$$

(8.8)

$F_0(\sin(\gamma),\omega)$ in (8.8) or more specifically its inverse Fourier transform, $F_0(\sin(\gamma),t)$, may be obtained from (1.32), that is, by slant stacking or Radon transform. Recall that $F_0(\sin(\gamma),t)$ is a plane wave incident at angle γ, a result of plane wave decomposition of a point source. $F_0(\sin(\gamma),\omega)$ is also known as the angular spectrum of the wavefield on surface. The angular spectrum at depth z is given by

$$F(\sin(\gamma),z,\omega) = F_0(\sin(\gamma),\omega) e^{\pm j(kz\cos(\gamma))}$$

(8.9)

8.3.2 <u>Layered Medium</u>: We can extend the wavefield extrapolation problem from a single layer to a two layer medium. Each layer is separated by a plane interface, either horizontal as in fig. 8.5a or inclined as in fig. 8.5b. In each layer the density and wave speed are constant. We shall assume that zero offset processing has removed all multiply reflected waves.

Let $f_1(x,y,z.t)$ and $f_2(x,y,z.t)$ be the wavefields produced by the exploding charges placed on the interface I and interface II respectively. The total field is given by $f(x,y,z.t) = f_1(x,y,z.t) + f_2(x,y,z.t)$. The wavefield observed on the surface z=0 is given by $f_0(x,y,0,t) = [f_1(x,y,z.t) + f_2(x,y,z.t)]\big|_{z=0}$. The wavefield generated by exploding charges on two interfaces is given by

$$f(x,y,z,t)\big|_{t=0} = r_1 \delta(z - g_1(x,y)) + r_2 \delta(z - g_2(x,y))$$

(8.10)

where r_1 and r_2 are reflection coefficients and $g_1(x,y)$ and $g_2(x,y)$ surfaces separating the two layers. It may be noted that by removing the multiple reflections we have decoupled the two interfaces; in effect we have linearized the propagation effects. Extension to the N-layer medium is straightforward when all layers are decoupled.

For extrapolation it is necessary that the correct propagation speed of the material in each layer is used. First, extrapolation to the first interface is carried out using (1.26). The Fourier transform of the wavefield observed on the surface is multiplied with the propagation filter function

$\exp(-j\sqrt{k_1^2 - u^2 - v^2}\,z)$, $0 \le z \le z_1$ where $k_1 = \dfrac{\omega}{c_1}$. Extrapolation to the second interface is obtained by multiplying with a filter function, $\exp(-j\sqrt{k_2^2 - u^2 - v^2}\,z)$, $z_1 \le z \le z_2$ where $k_2 = \dfrac{\omega}{c_2}$.

$$f_2(x,y,z,t)\big|_{t=0}$$

$$= \frac{1}{8\pi^3} \int\limits_{-\infty}^{-\infty}\int\int F_0(u,v,\omega)e^{-j\sqrt{k_1^2-u^2-v^2}\,z_1}e^{-j\sqrt{k_2^2-u^2-v^2}\,(z-z_1)}e^{-j(ux+vy)}\,dudvd\omega$$

$$= \frac{1}{8\pi^3} \int\limits_{-\infty}^{-\infty}\int\int \frac{F_0(u,v,\pm c_2 sgn(\omega)\sqrt{w^2+u^2+v^2}\,)}{e^{-j\left\{\sqrt{k_1^2-u^2-v^2}-\sqrt{k_2^2-u^2-v^2}\right\}z_1}} \frac{c_2 w}{\sqrt{w^2+u^2+v^2}} e^{-j(ux+vy-wz)}\,dudvdw$$

$$= r_2(x,y,z)$$

$$(8.11)$$

Note that $r_1(x,y,z)$ is zero in the second layer.

8.3.3 <u>Sloping Interface</u>: We shall now look at an example of a single sloping reflector as shown in fig. 8.6 [10]. We shall assume a ULA of transceivers oriented along the slope. For simplicity let us assume that the transceiver radiates a spike and receives, after a delay, the same spike. The seismic data are assumed to have been stacked so that we have a zero-offset data corresponding to a field from hypothetical charges placed on the interface (exploding reflector model described in §8.2). The zero-offset data may be expressed as

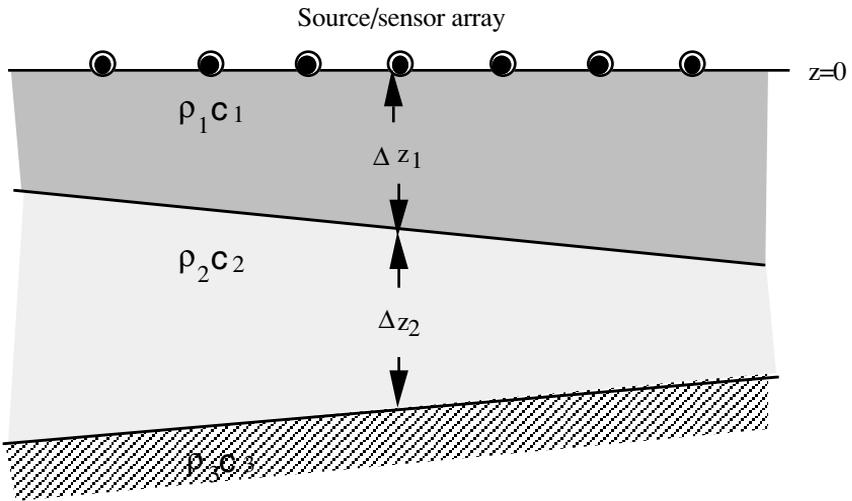$$f(md,t) = \delta(t - m\tau_0 - t_0), \quad m = 0,1,\dots M-1$$

where $\tau_0 = \dfrac{d}{c}\sin\theta_0$, and $t_0 = \dfrac{z_0}{c}\cos\theta_0$. Note that z_0 is a depth to the interface below sensor m=0. The Fourier transform of the zero-offset data may be obtained assuming the array size is very large,

$$F_0(u,\omega) = e^{-j\omega t_0}\delta(ud + \omega\tau_0) \tag{8.12}$$

Using (8.12) in (8.6a), that is, its 2D equivalent, we obtain after simplification
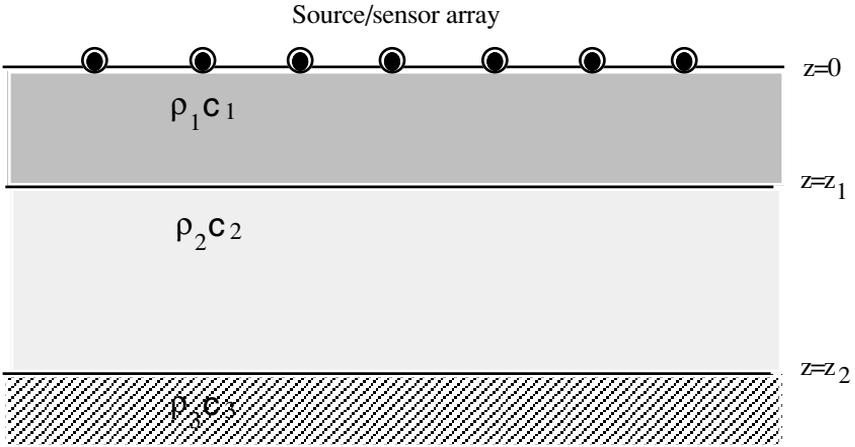
Figure 8.5: Wavefield extrapolation in two layer medium. (a) Horizontal layers and (b) inclined layers. ⦿ represents transceiver.

$r(x,z)$

$$= \frac{c}{4\pi^2} \int\int_{-\infty}^{-\infty} \frac{1}{\sqrt{1 + \frac{s^2}{w^2}}} e^{jc\sqrt{w^2+u^2}\,t_0} \delta(ud + c\tau_0\sqrt{w^2+u^2})$$
$$e^{-j(ux+wz)} du\, dw \tag{8.13}$$

$$= \frac{c}{2\pi} \int_{-\infty}^{-\infty} \cos\theta_0 e^{-j\frac{ct_0}{\cos\theta_1}w} e^{-j(-\tan\theta_0 x+z)w} dw$$

$$= c\cos\theta_0 \delta(-\tan\theta_0 x - z_0 + z)$$

From (8.13) the reflector is given by

$$z = \tan\theta_0 x + z_0 \tag{8.14}$$

It is interesting to note that the slope of the interface is not equal to the slope of the line joining the spike arrivals in the zero-offset seismic data (see exercise 3 at the end of this chapter).

8.3.4 <u>Depropagation of Wavefield</u>: Extrapolating a wavefield backward in time is known as depropagation; when continued it is possible to home onto the starting point. In the exploding reflector model all scatter points are simultaneously fired. We shall now show that the field observed on the surface is equal to the field generated at scatter points and then propagated to the surface. By depropagating, we hope to obtain the field generated at the exploding reflectors. The zero offset data provides only a blurred image of the reflectors, but by depropagating the blurring can be reduced [5]. Consider a 2D model with uniform wave speed with reflecting facets. The wavefield in a uniform medium is given, in the frequency domain, by

$$f(x,z,t) = \frac{1}{4\pi^2} \int\int_{-\infty}^{+\infty} F(u,\omega) e^{j(-ux+\sqrt{k^2-u^2}\,z)} e^{j\omega t} du\, d\omega$$

$$= \frac{c}{4\pi^2} \int\int_{-\infty}^{+\infty} \frac{1}{\sqrt{1+\frac{u^2}{v^2}}} F(u,\omega) e^{-j(ux+vz)} e^{-j\sqrt{u^2+v^2}\,ct} du\, dv \tag{8.15}$$

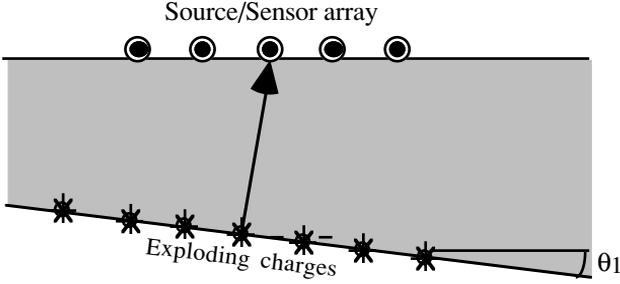where $v = -sgn(\omega)\sqrt{k^2 - u^2}$. The field on the surface, z=0, is given by

Figure 8.6: A sloping reflector and source/sensor (transceiver) array on the surface. By a process of stacking, the seismic field is reduced to a field generated by hypothetical exploding charges on the sloping interface.

$$f(x, z = 0, t) = \frac{c}{4\pi^2} \int\int_{-\infty}^{+\infty} \frac{1}{\sqrt{1 + \dfrac{u^2}{v^2}}} F(u, \omega) e^{-j(ux + \sqrt{u^2 + v^2} ct)} \, du \, dv \quad (8.16)$$

The mathematical framework for imaging by wavefield extrapolation is covered in great detail in [11, Berkhout].

The wavefield at the exploding reflector at a depth z and backpropagated to the surface at time $t = \dfrac{z}{c}$ is given by

$$f(x, z, t = \frac{z}{c}) = \frac{c}{4\pi^2} \int\int_{-\infty}^{+\infty} \frac{1}{\sqrt{1 + \dfrac{u^2}{v^2}}} F(u, \omega) e^{-j(ux + vz)} e^{-j\sqrt{u^2 + v^2} z} \, du \, dv \quad (8.17)$$

Since the wavefield is by and large vertically propagating, $u \ll v$ (8.17) may be approximated as

$$f(x, z, t = \frac{z}{c}) \approx \frac{1}{4\pi^2} \int\int_{-\infty}^{+\infty} c F(u, \omega) e^{-j(ux + 2vz)} \, du \, dv \qquad (8.18)$$

The wavefield observed on the surface (z=0) is mapped into the (x,z) plane by substituting $\dfrac{z}{c/2}$ for t. Note that the time axis of the recorded seismic data

refers to two way time and hence wave speed is halved as is common in seismic data processing. We obtain

$$f(x, z = 0, t = \frac{z}{c/2}) \approx \frac{1}{4\pi^2} \int\int_{-\infty}^{+\infty} cF(u, \omega)e^{j(-ux+2vz)} du dv \qquad (8.19)$$

Thus, from (8.18) and (8.19) we have an important approximate result which provides an initial wavefield for depropagation,

$$f(x, 0, t = \frac{z}{c/2}) \approx f(x, z, t = \frac{z}{c}) \qquad (8.20)$$

In words, the wavefield observed on the surface with its time axis mapped into a z-axis ($t = \frac{z}{c/2}$) is approximately equal to the wavefield at the reflector which is then propagated to the surface.

Now given $f(x, 0, t = \frac{z}{c/2})$ or $f(x, z, t = \frac{z}{c})$ we need to depropagate till we reach the point of reflection. To depropagate by Δt time units we shall use (8.15) with approximation $u << v$

$$f_d(x, z, t - \Delta t)$$

$$\approx \frac{1}{4\pi^2} \int\int_{-\infty}^{+\infty} cF(u, \omega)e^{-j(ux-vz)}e^{-jvc\Delta t} du dv \qquad (8.21)$$

$$= \frac{c}{4\pi^2} \int\int_{-\infty}^{+\infty} FT\left\{ f(x, 0, t = \frac{z}{c/2}) \right\} e^{-j(ux-vz)}e^{-jvc\Delta t} du dv$$

where the subscript d stands for depropagated field (see fig. 8.7)

8.3.5 <u>Relation to Diffraction Tomography</u>: We shall now show that a close relation exists between imaging by extrapolation and diffraction tomography which we have dealt with in the previous chapter. Consider a weakly inhomogeneous object (speed fluctuations only) illuminated by a point source and the scattered field being received by a sensor kept close to the source. We have shown in chapter 7 (pp. 400-402) that the Fourier transform of the speed variations evaluated on a spherical surface of radius 2k is related to the Fourier transform of the field received by an array of transceivers (see fig. 7.25). The relationship (7.33) is reproduced here for convenience:
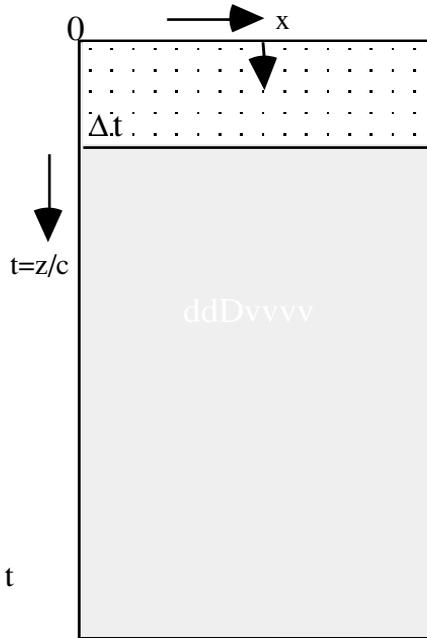
Figure 8.7: The time axis of the wavefield measured on the surface is mapped onto the z-axis (the wave speed is assumed to be known). The wavefield thus obtained in the (x, z) plane is now depropagated in steps of Δt. The depropagated segment corresponds to the desired reflectivity function over a segment of depth, $\Delta t\, c$.

$$\Delta\tilde{c}(u,v,\sqrt{4k_0^2 - u^2 - v^2}) = \frac{c^3}{2}\sqrt{4k_0^2 - u^2 - v^2}\,\frac{\partial}{\partial\omega_0}\left(\frac{F_1(u,v,\omega_0)}{\omega_0^2}\right)$$

It may be observed that the zero offset field measured in seismic exploration is essentially an output of an array of transceivers which may be considered as a mathematical model of the zero offset field measurement setup.

To relate $\Delta\tilde{c}(u,v,\sqrt{4k_0^2 - u^2 - v^2})$ to the reflectivity, we shall use (8.7a) where we have shown the relationship between the Fourier transform of the reflectivity and the wavefield on the surface (zero offset). It can be shown that the partial derivative appearing in (7.33) may be given by

$$\frac{\partial}{\partial \omega}\left(\frac{F_1(u,v,\omega)}{\omega^2}\right) = -\frac{1}{2}\frac{\partial}{\partial w}\left(\frac{R(u,v,w)}{\sqrt{w^2+s^2}\,w}\right)\frac{\sqrt{w^2+s^2}}{w} \qquad (8.22)$$

where $R(u,v,w)$ is the spatial Fourier transform of r(x,y,z). Substituting (8.22) in (7.33) we obtain

$$\Delta \tilde{c}(u,v,\sqrt{4k_0^2-u^2-v^2})$$
$$= -\frac{c^3}{4}\sqrt{w^2+s^2}\,\frac{\partial}{\partial w}\left(\frac{R(u,v,w)}{\sqrt{w^2+s^2}\,w}\right) \qquad (8.23)$$

After carrying out the required differentiation we obtain the following result:

$$\Delta \tilde{c}(u,v,w)$$
$$= \frac{c^3}{4}\left(\frac{R(u,v,w)}{(w^2+s^2)}+\frac{R(u,v,w)}{w^2}-\frac{R_w(u,v,w)}{w}\right) \qquad (8.24)$$

where $R_w(u,v,w)$ stands for the derivative of $R(u,v,w)$ with respect to w.

8.3.6 <u>Continuation of Sources and Sensors</u>: Earlier in §8.1 we had mentioned that imaging can be realized by continuing both the source and sensor arrays down to the reflector. This is particularly useful for imaging when zero offset data are not available [12]. We shall now derive filters for the downward continuation of the source and sensor arrays. The transfer function of a filter to continue the wavefield to a plane Δz below the observation plane may be obtained from (1.18)

$$H(u_r,v_r,\Delta z) = \exp(+j\sqrt{(k^2-u_r^2-v_r^2)}\Delta z) \qquad (8.25a)$$

where the subscript r refers to the receiver coordinate. In the first step the sources are held fixed on the z=0 plane (observation plane) and sensor array is displaced downward. Next we interchange the source array with the displaced receiver array. By virtue of the reciprocity theorem the wavefield at the new position of the sensor array due to the source array also at its new position will be equal to the previously downward continued field. In the second step we downward continue the previously continued field but the continuation is done in source coordinate space. The transfer function for downward continuation is given by

$$H(u_s,v_s,\Delta z) = \exp(+j\sqrt{(k^2-u_s^2-v_s^2)}\Delta z) \qquad (8.25b)$$
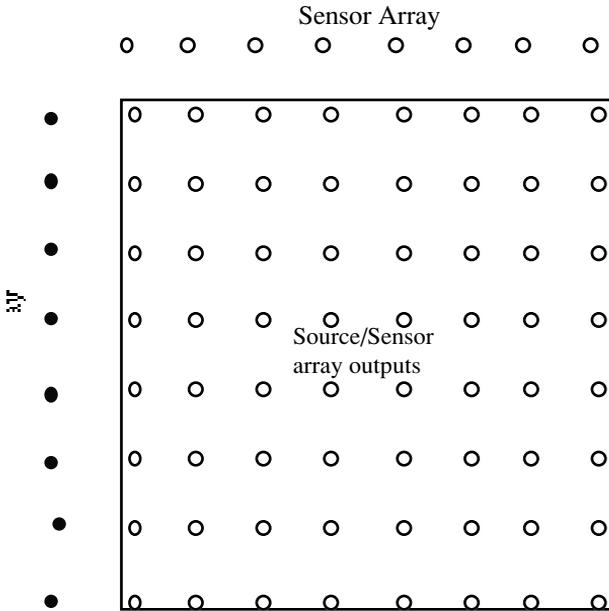
Figure 8.8: M^2 outputs source/receiver ULAs are arranged over a square grid. All outputs in a row are known as source gathers and all outputs in a column are known as sensor gathers.

where the subscript s refers to source coordinate. The downward continuation is alternatively done in the source and sensor spaces. In practice the downward continuation is carried out as follows. Assume that we have two source and sensor ULAs each having M devices. There will be M^2 outputs which we like to arrange on a square grid as shown in fig. 8.8. All outputs in a row (known as source gathers) are continued downwards in the sensor space and all outputs in a column (known as sensor gathers) are continued in the source space.

§8.4 **Focused Beam**:
When a source is in the far field, the directions of arrival (DOA), azimuth and elevation are of interest. We have already seen how a beam is formed in a given direction (chapters 4 and 5). On the other hand, when a source is in the near field region, a beam may be formed to receive energy not only from a given direction but also from a given point. This is akin to focusing in an optical system. In seismic exploration the array size is of the same order as the depth of reflectors; therefore, it is often inappropriate to assume a far field or plane wavefront condition which is required for the purpose of imaging, that is,
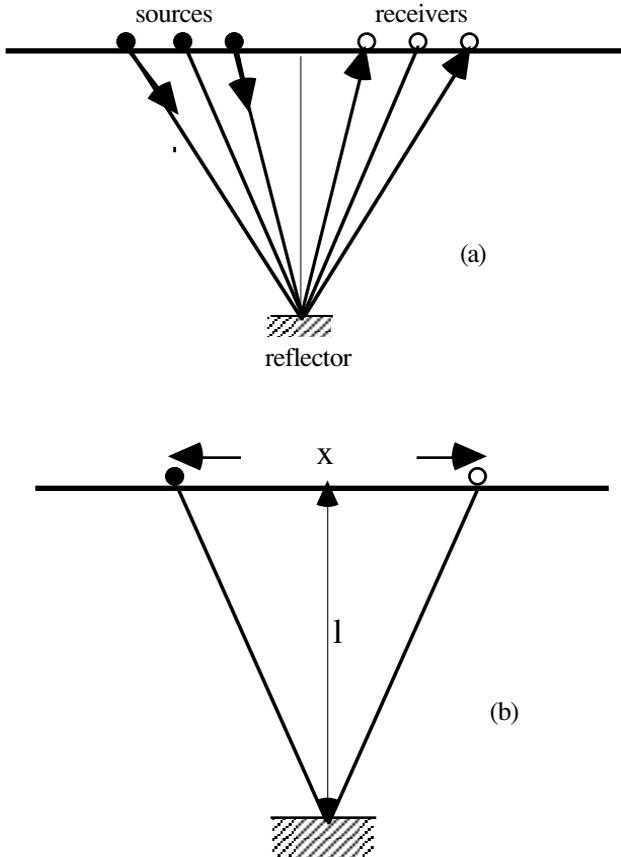
Figure 8.9: Common depth point (CDP) setup. (a) The source and receiver arrays are arranged in such a manner that a signal is always reflected from the same reflecting element. (b) Source-receiver pair.

to form a focused beam to receive reflected energy from a given reflector.

8.4.1 <u>Zero-offset Wavefield</u>: In reflection seismic exploration it is often desired to position both source and detector at the same location, although in practice this cannot be achieved and some amount of offset is always present. The main advantage of this arrangement is that the returned signal is a normally reflected signal; consequently, there is no conversion of wave energy into s-waves (see chapter 1, p. 28). The zero offset wavefield can however be obtained through the array processing approach. For this we must focus a beam at the foot of a perpendicular drawn from the array mid-point to the reflector. A typical example

of a source and receiver setup is shown in fig. 8.9. All signals reaching the sensor array emanate from the same reflecting element. In seismic parlance these signals are called common depth point (CDP) gathers.

Common Depth Point: A linear array of sources and detectors are so arranged that the received signal always comes from the same reflecting element. This arrangement is illustrated in fig. 8.9a, and is commonly known as a common depth point (CDP) setup. The total time of travel from the source to the receiver is a function of separation or offset between the source and the receiver. When the reflecting element is horizontal the travel time is given by

$$T_x^2 = T_0^2 + \frac{x^2}{c^2} \tag{8.26}$$

where x stands for separation between the source and receiver , $T_0 = \dfrac{2l}{c}$, and l is depth to the reflecting element (see fig 8.9b). A plot of T_x vs x is very useful for it enables us to estimate T_0 and the wave speed.

Let us now consider a sloping reflecting element but the source and receiver arrays are, as before, on a horizontal surface (see fig. 8.10). Let the slope of the reflecting element be α. To compute the travel time to a downslope or an upslope sensor we consider the image of the source and compute the distance from the image to the sensor. For a downslope sensor,

$$\begin{aligned} T_{x_+} &= \frac{(ir_+)}{c} \\ &= \frac{\sqrt{(x_+)^2 + 4l^2 - 4lx_+ \cos(\angle isr_+)}}{c} \end{aligned} \tag{8.27}$$

and squaring on both sides of (8.27) we obtain

$$T_{x_+}^2 = \left(\frac{x_+}{c}\right)^2 + T_0^2 + 2T_0 \frac{x_+}{c} \sin(\alpha) \tag{8.28}$$

Similarly, we have for an upslope receiver

$$T_{x_-}^2 = \left(\frac{x_-}{c}\right)^2 + T_0^2 - 2T_0 \frac{x_-}{c} \sin(\alpha) \tag{8.29}$$

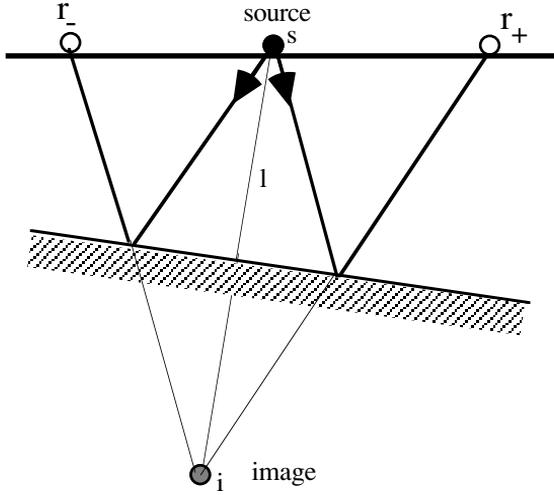The upslope and downslope travel times are shown in fig. 8.11.

Figure 8.10: Inclined reflector. There are two sensors, one sensor, r_+, is downslope and another sensor, r_-, is upslope. To compute the travel time we consider an image of the source. The time of travel to the downslope sensor is equal to $(i\,r_+)/c$ and similarly for the upslope sensor.
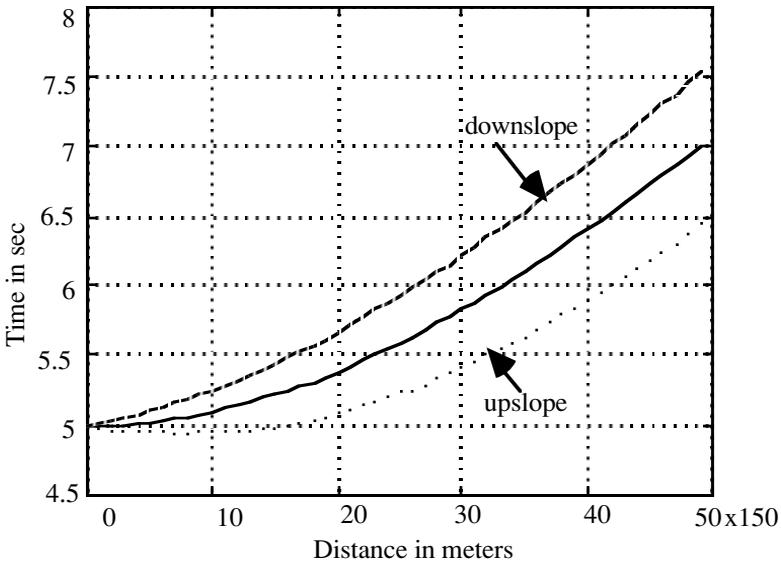

Figure 8.11: Downslope and upslope travel times (eqs. 8.21 & 8.22) in seconds as a function of distance in meters. Wave speed $=1500$m/s, $T_0=5$sec. Solid curve is for horizontal reflector (8.19).

where $T_0 = \dfrac{2l}{c}$ and l is the perpendicular depth to the reflector below the source position (see fig. 8.9). Note that (8.28) and (8.29) reduce to (8.26) for $\alpha = 0$ and for $x \to 0$, $T_x \to T_0$. Rearranging terms in (8.28) and (8.29) we can express them in a compact form

$$T_{x_\pm}^2 = \left(T_0 \pm \frac{x_\pm}{c}\sin(\alpha)\right)^2 \pm \left(\frac{x_\pm}{c}\right)^2 \cos(\alpha) \qquad (8.30)$$

Let the downslope and the upslope sensor be equidistant from the source, that is, $x_+ = x_- = \bar{x}$. The average of the squares of downslope and upslope travel times turns out to be independent of the slope of the reflector,

$$\frac{T_{x_+}^2 + T_{x_-}^2}{2} = \left(\frac{\bar{x}}{c}\right)^2 + T_0^2 \qquad (8.31)$$

Similarly, the difference of the squares of downslope and upslope travel times is given by

$$\frac{T_{x_+}^2 - T_{x_-}^2}{4} = T_0 \frac{\bar{x}}{c}\sin(\alpha) \qquad (8.32)$$

8.4.2 <u>Layered Medium</u>: We now consider a layered medium overlying a reflector. The overlying medium is modeled as a stack of uniform layers (see fig. 8.12). The round trip travel time T_x and the receiver position x are the two parameters of interest. They are given by

$$T_x = \sum_{k=1}^{N} \frac{2\Delta z_k}{c_k \sqrt{(1 - p^2 c_k^2)}} \qquad (8.33a)$$

$$x = p \sum_{k=1}^{N} \frac{2\Delta z_k c_k}{\sqrt{(1 - p^2 c_k^2)}} \qquad (8.33b)$$
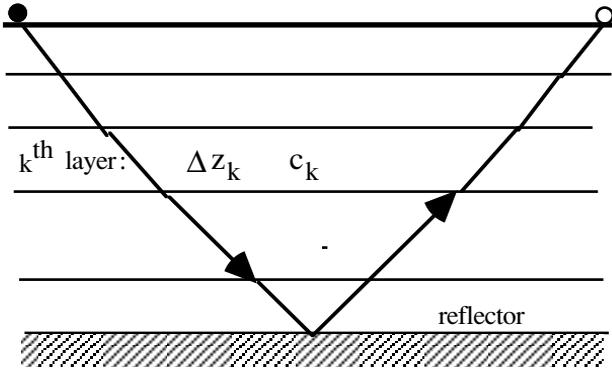
Figure 8.12: Layered medium overlying a reflector. Each layer is homogeneous with constant thickness and speed. The ray path consists of a series of linear segments. At an interface between two layers the Snell's law (see p.) must be satisfied.

where p is the ray parameter ($p = \dfrac{\sin(\theta_k)}{c_k}$) where θ_k is the angle of incidence in the kth layer (see page for more information on the ray parameter) and N stands for the number of layers. Let us expand T_x as a function of x in Taylor's series,

$$T_x = T_0 + \frac{d^2 T_x}{dx^2}\bigg|_{x=0} \frac{x^2}{2!} + \frac{d^4 T_x}{dx^4}\bigg|_{x=0} \frac{x^4}{4!} + \qquad (8.34)$$

Since, for a horizontally layered medium, T_x is a symmetric function of x, only even terms are retained in the Taylor's series expansion (8.27). Further, the second derivative can be shown to be equal to $\dfrac{d^2 T_x}{dx^2}\bigg|_{x=0} = \dfrac{1}{\sum\limits_{k=1}^{N} \Delta z_k c_k}$.

Equation (8.34) simplifies to

$$T_x \approx T_0 + \frac{x^2}{2\sum\limits_{k=1}^{N} \Delta z_k c_k} \qquad (8.35)$$

Upon squaring on both sides of (8.35) and retaining only the second order terms we obtain

$$T_x^2 \approx T_0^2 + \frac{x^2 T_0}{\sum\limits_{k=1}^{N} \Delta z_k c_k}$$

which after rearranging reduces to a form identical to (8.26),

$$T_x^2 \approx T_0^2 + \frac{x^2}{\dfrac{1}{T_0} \sum\limits_{k=1}^{N} \dfrac{\Delta z_k}{c_k} c_k^2} = T_0^2 + \frac{x^2}{c_{rms}^2} \qquad (8.36)$$

where $c_{rms}^2 = \dfrac{1}{T_0} \sum\limits_{k=1}^{N} \dfrac{\Delta z_k}{c_k} c_k^2$ is known as a root mean square speed of the layered medium. Indeed we may replace a layered medium by a uniform medium having a speed equal to c_{rms}. For the purpose of focused beamformation we may use the root mean square speed.

8.4.3 <u>Focusing</u>: To form a focused beam we must have many CDP gathers which are obtained by means of a specially designed source-receiver array. For example, consider a linear array where every location is occupied either by a source or a sensor. The array is fired as many times as the number of locations. From every reflecting element we obtain a number of gathers equal to the number of receivers. This is illustrated in fig. 8.13. A ULA of four sensors is headed by a source. In position (A) the source is fired and the reflected signal is received by sensor #1. The entire receiver-source array is moved laterally by half sensor spacing as in position (B) and the source is fired once again. The reflected signal is received by sensor #2. This procedure is continued as in (C) and (D). Thus, we get four CDP gathers.

Let $T_1, T_2, \ldots T_N$ be the round-trip travel time from the source to sensor #1, #2,...#N, respectively. Let $f_0(t)$ be the signal transmitted by the source and $f_i(t), i = 1,2,\ldots N$ be the reflected signals received by four receivers. Since these are the delayed versions of $f_0(t)$ we can express them as

$$f_i(t) = f_0(t - T_i), i = 1,2,\ldots N \qquad (8.37)$$

Focusing involves firstly correcting for delayed reception and secondly summing coherently after correction. Let the delay correction be given by
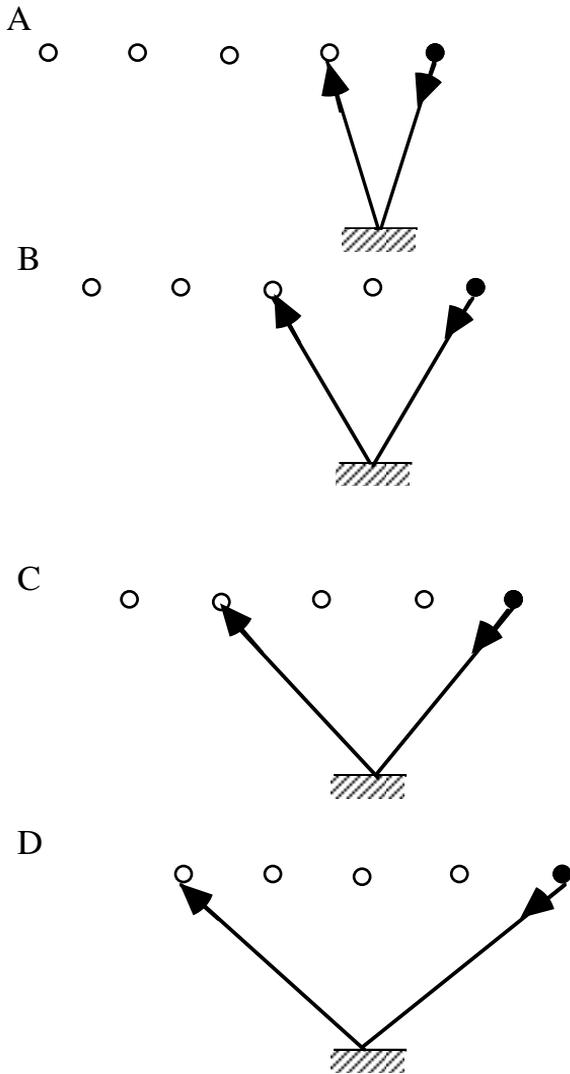
Figure 8.13: A reflecting element is illuminated by a source at different angles. The reflected signal is received by one of the sensors as shown in the figure. The entire receiver-source array is moved laterally by half sensor spacing.

$\hat{T}_1, \hat{T}_2, ..., \hat{T}_N$ which are computed using (8.26) for an assumed depth to the reflector,

$$\hat{T}_n = \hat{T}_0 \sqrt{(1 + (\frac{\Delta tn}{\hat{T}_0})^2)} \qquad (8.38)$$

where $\hat{T}_0 = \dfrac{2\hat{l}}{c}$, $\hat{l}$ is assumed depth to the reflector and $\Delta t = \dfrac{d}{c}$. Recall that d stands for sensor spacing. We assume that the wave speed c is known. The CDP gathers are coherently summed after correction for the delays computed from (8.38). As before we shall assume that the source emits a broaband signal. The coherently summed output may be expressed as follows:

$$g(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \frac{1}{N} \sum_{n=0}^{N-1} e^{-j\omega\left[T_0\sqrt{(1+(\frac{\Delta tn}{T_0})^2)} - \hat{T}_0\sqrt{(1+(\frac{\Delta tn}{\hat{T}_0})^2)}\right]} e^{j\omega t} d\omega$$

$$(8.39a)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) H_N(\omega) e^{j\omega t} d\omega$$

where

$$H_N(\omega) = \frac{1}{N} \sum_{n=0}^{N-1} e^{-j\omega\left[T_0\sqrt{(1+(\frac{\Delta tn}{T_0})^2)} - \hat{T}_0\sqrt{(1+(\frac{\Delta tn}{\hat{T}_0})^2)}\right]} \qquad (8.39b)$$

is the filter transfer function. A numerical example of the transfer function is shown in fig. 8.14. A horizontal reflector is assumed at a depth corresponding to round-trip time equal to 5 seconds. The sensor spacing, measured in units of propagation time, $\Delta t = \dfrac{d}{c} = 0.1$ seconds.

8.4.4 <u>Depth of Focus</u>: The response function has a finite width. A sharp reflector now appears as a diffused zone whose width is known as the depth of focus, analogous to that in optics. Ideally one would like the depth of focus to be as narrow as possible. For the purpose of quantitative measure we shall define the depth of focus as a distance between two 3 dB points on the response function. The depth of focus, measured in the units of round-trip propagation time (=$\dfrac{2\Delta l}{c}$, where Δl is depth of focus), and the array aperture, also measured in terms of the propagation time (=$\dfrac{x}{c}$), are shown in fig. 8.15. Notice that the minimum occurs when the aperture size is about four times the round-trip propagation time, in this case five seconds. With further increase in
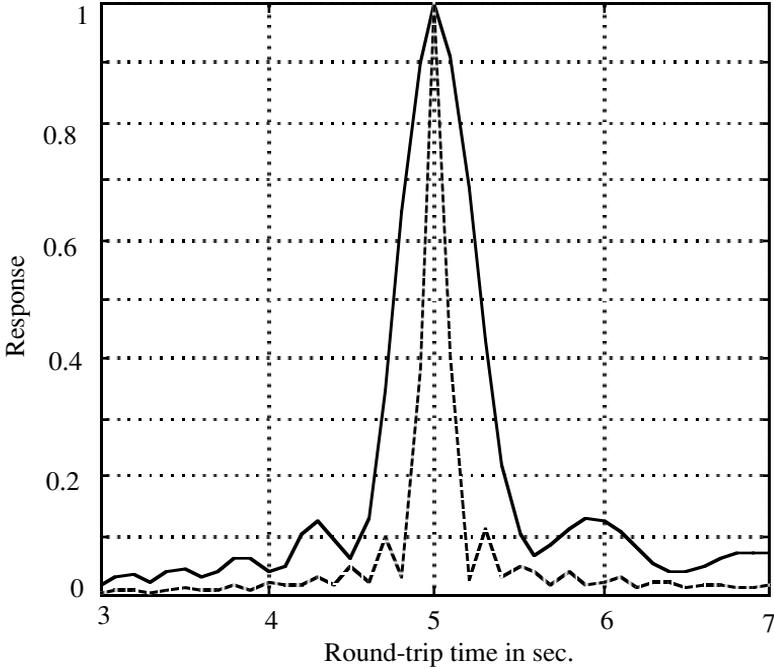
Figure 8.14: Response function of focused beamformation. The reflector is at a depth corresponding to 5 sec round-trip time. The array aperture is measured in propagation time ($\frac{x}{c}$). Solid line is for array aperture of 2.5 seconds and dashed line for 5.0 seconds. Further the angular frequency is assumed to be 100 radians.

the aperture size the depth of focus rapidly deteriorates. The minimum depth of focus appears to be independent of depth to the reflector; however, the required array aperture increases rapidly as the reflector depth increases. The dependence of the depth of focus on frequency is significant, as shown in fig. 8.16. Notice that the depth of focus becomes very narrow beyond about 50Hz.

8.4.5 <u>Inclined Reflector</u>: We consider a sloping reflector. A small segment of the reflector is illuminated at different angles by means of a source-receiver array similar to the one used for a horizontal reflector. As the array is laterally shifted the point of reflection changes, but only slightly, depending upon the slope (see fig. 8.17). When the slope is zero all CDP gathers emanate from the same point on the reflecting element. For gentle slope the spread of the reflecting points is assumed to be small, small enough to allow the assumption of a constant slope.
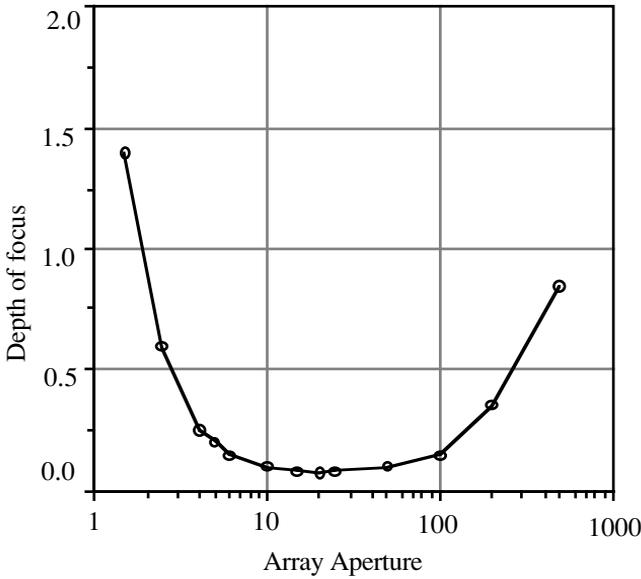
Figure 8.15: Depth of focus in units of round-trip propagation time (sec) as a function of array aperture, also measured in units of propagation time (sec).
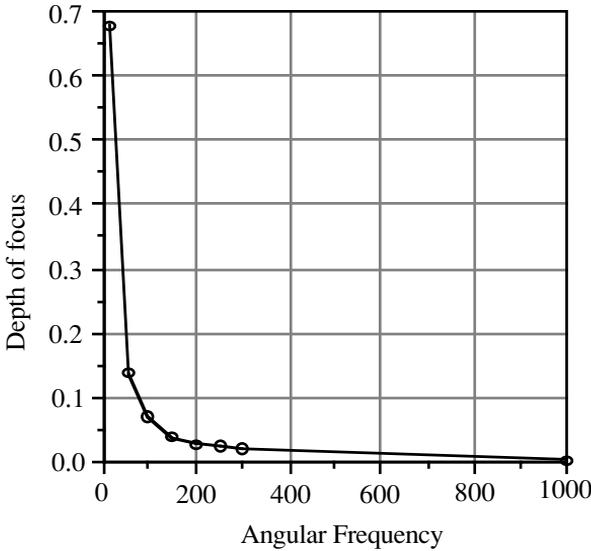

Figure 8.16: Depth of focus vs angular frequency. The reflector is at a depth of 5 seconds (round-trip travel time). The array aperture is held fixed at 20 seconds (propagation time).
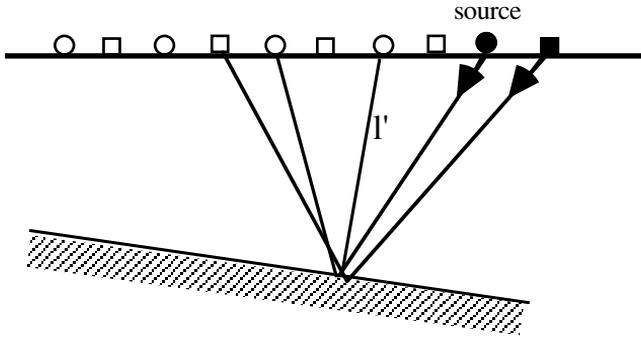
Figure 8.17: Common depth point (CDP) gathers from a sloping reflector. The reflecting element is at a depth of l' (round trip travel time=$\dfrac{2l'}{c}$). Two positions of source-sensor array are shown; position #1: circles and position #2: squares. Notice the displacement of the reflecting point.

The response of the focused beamformer as a function of the slope and the return travel time is shown in fig. 8.18. We have assumed that $\dfrac{2l'}{c} = 5$ and array aperture, in units of propagation time, is equal to $\dfrac{x}{c} = 20$ seconds. There are 200 CDP gathers spaced at an interval equal to 0.1 sec. The result shown in fig. 8.18 is perhaps the best one can expect as we have assumed the optimum array size.

8.4.6 Relation Between Focusing and Downward Extrapolation: Focusing appears to achieve what the downward extrapolation does in migration. Indeed both are related in the sense that focusing is a simplified version of extrapolation. In order to see this relationship let us examine the impulse response function of the downward extrapolation filter whose transfer function is given by

$$H(\omega, u, v) = \exp(j\Delta z \sqrt{((\frac{\omega}{c})^2 - u^2 - v^2)}) \qquad (8.40a)$$

To get the impulse response function we compute the 2D inverse Fourier transform of (8.40a). The result is given in [13]
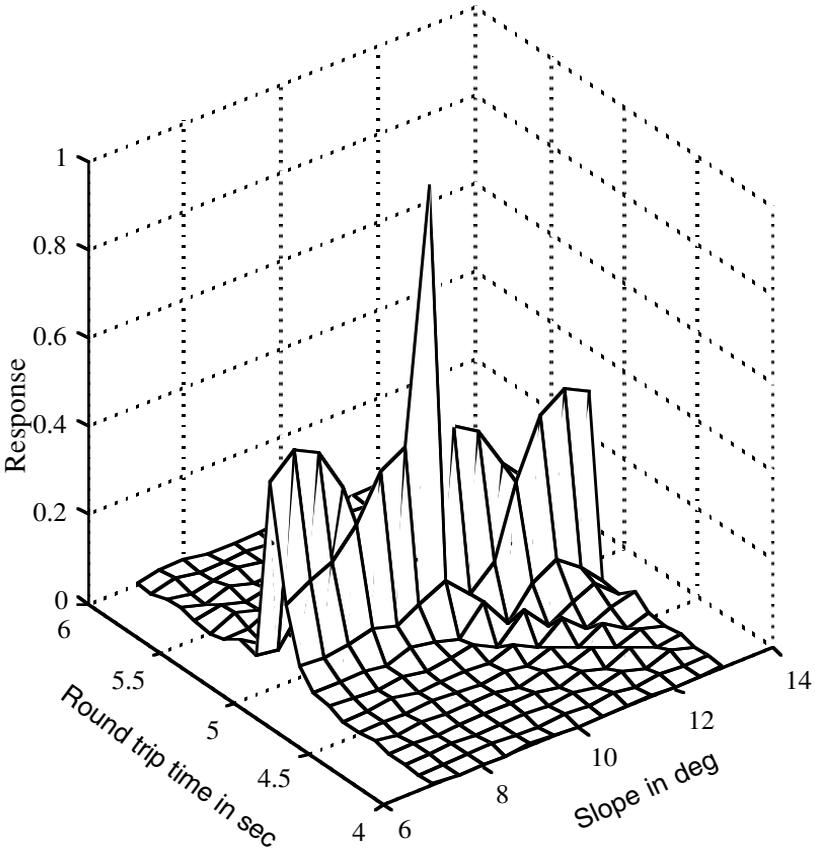
Figure 8.18: A mesh plot of the response function of the focused beamformer as a function of slope of the reflector and round-trip time. Array aperture is 20 seconds (propagation time). Angular frequency assumed is equal to 100 radians.

$$h(r,\omega) = \frac{1}{2\pi} \frac{\partial}{\partial z} \left[\frac{\exp(j\frac{\omega}{c}r)}{r} \right] = \frac{1}{2\pi} \frac{j\frac{\omega}{c}r + 1}{r^3} \exp(j\frac{\omega}{c}r) \qquad (8.40b)$$

where $r = \sqrt{x^2 + y^2 + \Delta z^2}$. Let us rewrite the exponential term in (8.40b) as

$$\exp(j\frac{\omega}{c}r) = \exp(j\omega\sqrt{\frac{\rho^2}{c^2} + \frac{\Delta z^2}{c^2}}) = \exp(j\omega\sqrt{T_0^2 + \frac{\rho^2}{c^2}}) \qquad (8.40c)$$

where $\rho^2 = x^2 + y^2$. The phase delays introduced by the downward extrapolation filter are identical to those used in the focused beamformer. The difference, however, lies in the amplitude term. In place of a variable amplitude, we use a constant amplitude in the focused beamformer.

8.4.7 Focused Beamformation for Imaging: A focused beam enables us to estimate the amount of scattered or reflected power from a scattering volume element or a reflecting surface element located at a point in space. Consider a volume element illuminated by a point source with the scattered waves being received by an ULA (see fig. 8.19). To form a focused beam we need to compute the travel time from the source to the volume element and from the volume element to a sensor. For simplicity let us assume that the background wave speed is constant so that ray paths are straight lines. In any realistic problem, however, the ray paths are more likely to be curved. It will be necessary to trace the rays before we are able to compute the travel times. Such a situation was considered in [14]. Let $f_n(t), n = 1, 2, \ldots N$ be the sensor outputs. The focused beam output is given by

$$g(t) = \frac{1}{N} \sum_{n=1}^{N} f_n(t - t(so) - t(or_n))$$

where $t(so)$ stands for the travel time from the source to the volume element and $t(or_n)$ stands for the travel time from the volume element to the nth sensor. The scattered power from the volume element is given by $\int_{T_s} |g(t)|^2 dt$ where T_s stands for signal duration. The process is repeated at every point in the space (x-z plane). The resulting map provides an image of the scattering strength.

The question of resolution needs to be answered. The size of the volume element is controlled, along the x-axis, by the array aperture (inversely proportional to array aperture), along the z-axis, by depth of focus and finally along the y-axis, by the bandwidth of the signal (inversely proportional to bandwidth) [14]. An interesting application of the focusing by means of back propagation of wavefield to detection of an acoustic source in a room without considering the reflections from the walls is reported in [15]. An array of microphones is deployed in a 3D space surrounding the source(s) of sound energy. The output from each microphone is back propagated to a point in
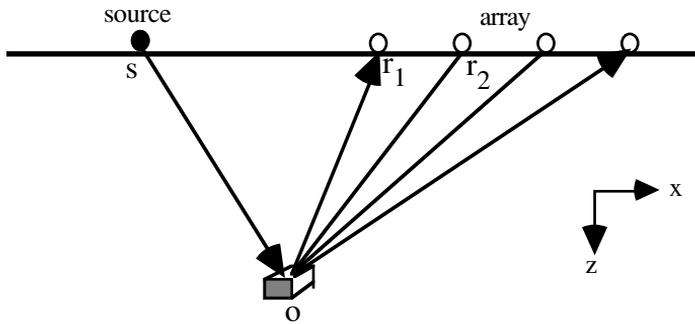
Figure 8.19: The array output is focused to a point o where a scattering volume element is presumed to be located. The power output is proportional to the scattering strength of the volume element.

3D space. Since we assume that no reflections from the walls or scattering exist, the back propagation is simply equal to delaying the microphone output by an amount equal to the travel time along a straight line path from the sensor to the assumed position of the source. The delayed output is cross-correlated with similarly delayed outputs from other sensors. An average of the cross-correlation is a measure of the acoustic energy present at the selected point. Ideally, the averaged cross-correlation function will have a sharp peak at the true sound source position. In practice, it is shown in [15] that a broad peak stands at the true sound source position.

§8.5 **Estimation of Wave Speed**:

We have assumed that the average background wave speed is known and this information was used for the purpose of imaging. We shall now look into how the wave speed can be estimated from the recorded wavefield itself. Firstly, the medium is assumed to be homogenous for which one needs to estimate just one unknown, namely, the wave speed of the homogenous medium. Secondly, the medium is a stack of layers for which we need to estimate the root mean square speed defined earlier (8.36). It turns out that the approach remains the same for both.

8.5.1 <u>Wave Speed from CDP Gathers</u>: In the context of CDP gathers the round-trip travel time is given (8.26) for a homogeneous medium and by (8.36) for a layered medium. The only difference is that in place of wave speed of a homogenous medium we have root mean square speed, c_{rms}. To estimate the wave speed or rms wave speed consider the travel time at two distinct sensor locations and compute the difference of the squares

$$T_{x_1}^2 - T_{x_2}^2 = \frac{x_1^2 - x_2^2}{c^2} \tag{8.41}$$

from which the wave speed can be obtained as

$$c = \sqrt{\frac{x_1^2 - x_2^2}{T_{x_1}^2 - T_{x_2}^2}} \tag{8.42}$$

For estimation of c_{rms} we use an identical approach. When the reflector is sloping, we use the relation between the round-trip travel time and the source sensor separation given by (8.28 & 8.29). Consider two sensors, equidistant from the source, one located downslope and the other upslope. From (8.31) it is clear that the slope of the reflector does not affect the average round-trip travel time. We can now use (8.31) in place of (8.26) for the estimation of wave speed.

Interval Speed: There is an interesting possibility of estimating the wave speed of a particular layer. Consider c_{rms} at two different depths having N and N+1 layers. From (8.36) we have

$$c_{rms}^2\big|_N = \frac{1}{T_0\big|_N} \sum_{k=1}^{N} \frac{\Delta z_k}{c_k} c_k^2$$

$$c_{rms}^2\big|_{N+1} = \frac{1}{T_0\big|_{N+1}} \sum_{k=1}^{N+1} \frac{\Delta z_k}{c_k} c_k^2$$

By subtraction we obtain an estimate of the wave speed in the N+1st layer

$$c_{N+1} = \frac{c_{rms}^2\big|_{N+1} T_0\big|_{N+1} - c_{rms}^2\big|_N T_0\big|_N}{\Delta z_{N+1}} \tag{8.43}$$

Now consider an interval containing p layers. Again from (8.36) we obtain after subtraction

$$\sum_{k=1}^{p} c_{N+k} \Delta z_{N+k} = c_{rms}^2\big|_{N+p} T_0\big|_{N+p} - c_{rms}^2\big|_N T_0\big|_N \tag{8.44a}$$

which may be rewritten in a form,

$$\frac{1}{T_0\big|_{N+p} - T_0\big|_N} \sum_{k=1}^{p} \frac{\Delta z_{N+k}}{c_{N+k}} c_{N+k}^2 = \frac{c_{rms}^2\big|_{N+p} T_0\big|_{N+p} - c_{rms}^2\big|_N T_0\big|_N}{T_0\big|_{N+p} - T_0\big|_N}$$

By definition, the right hand side is the interval rms wave speed in the interval containing p layers

$$c_{rms}^2\big|_p = \frac{1}{T_0\big|_{N+p} - T_0\big|_N} \sum_{k=1}^{p} \frac{\Delta z_{N+k}}{c_{N+k}} c_{N+k}^2$$

$$= \frac{c_{rms}^2\big|_{N+p} T_0\big|_{N+p} - c_{rms}^2\big|_N T_0\big|_N}{T_0\big|_{N+p} - T_0\big|_N}$$

$$(8.44b)$$

8.5.2: <u>Estimation in Presence of Errors</u>: Accurate knowledge of the wave speed is essential not only for imaging but also for characterization of the medium. The geophysical literature is proliferated with references on wave speed estimation. Several different approaches have been tried to estimate this important parameter. Briefly, correlation between the sensor outputs was used to compute the so-called velocity spectrum in [16, 17]; minimum entropy approach was suggested in [18], wavefield extrapolation in [19] and maximum likelihood estimate of rms wave speed was suggested in [20]. Our interest shall be limited to just two approaches, namely, maximum likelihood and focusing with optimum wave speed.

<u>Maximum Likelihood</u>: We model the measurement error in T_x, as an additive random variable, that is,

$$\hat{T}_{x_i} = \sqrt{T_0^2 + \frac{x_i^2}{c_{rms}^2}} + \tau_i \qquad (8.45)$$

where τ_i is a measurement error in $\hat{T}_{x_i}$, assumed to be a zero mean Gaussian random variable. We introduce the following vectors:

$$\hat{\mathbf{T}}_x = col\{\hat{T}_{x_1}, \hat{T}_{x_2}, ..., \hat{T}_{x_N}\}$$

$$\mathbf{T}_m = col\left\{\sqrt{T_0^2 + \frac{x_1^2}{c_{rms}^2}}, \sqrt{T_0^2 + \frac{x_2^2}{c_{rms}^2}}, ..., \sqrt{T_0^2 + \frac{x_N^2}{c_{rms}^2}}\right\}$$

$$\tau = col\{\tau_i, i = 1, 2, ..., N\}$$

Further, let c_{rms}^2 be a random variable with probability density function pdf(c_{rms}^2). We would like to maximize conditional probability density function $pdf\left(c_{rms}^2/\hat{\mathbf{T}}_x\right)$. Since

$$pdf\left(c_{rms}^2/\hat{\mathbf{T}}_x\right) = \frac{pdf(\hat{\mathbf{T}}_x/c_{rms}^2)pdf(c_{rms}^2)}{pdf(\hat{\mathbf{T}}_x)} \qquad (8.46)$$

it is sufficient to maximize the numerator in (8.46). For this note that

$$pdf(\hat{\mathbf{T}}_x/c_{rms}^2) = pdf(\tau)$$

$$(8.47a)$$

$$= \frac{1}{(2\pi)^{\frac{N}{2}}\sqrt{\mathbf{C}}}\exp\left[-\frac{1}{2}(\hat{\mathbf{T}}_x - \mathbf{T}_m)^H\mathbf{C}^{-1}(\hat{\mathbf{T}}_x - \mathbf{T}_m)\right]$$

where $\mathbf{C}$ is covariance matrix of measurement errors. Further, c_{rms}^2 is also assumed to be a Gaussian random variable with mean equal to $\bar{c}_{rms}^2$ and standard deviation σ

$$pdf(c_{rms}^2) = \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{1}{2}\left[\frac{(c_{rms}^2 - \bar{c}_{rms}^2)}{\sigma}\right]^2) \qquad (8.47b)$$

Substitute (8.47a) and (8.47b) in the numerator of (8.46), which is then maximized with respect to c_{rms}^2. We shall maximize the logarithm (natural) of the numerator

$$\max\left\{(\hat{\mathbf{T}}_x - \mathbf{T}_m)^H\mathbf{C}^{-1}(\hat{\mathbf{T}}_x - \mathbf{T}_m) + \left[\frac{(c_{rms}^2 - \bar{c}_{rms}^2)}{\sigma}\right]^2\right\}_{w.r.t.c_{rms}^2} \qquad (8.48)$$

Differentiate (8.48) with respect to c_{rms}^2 and set the derivative to zero to obtain

$$3\hat{\mathbf{T}}_x^H\mathbf{C}^{-1}\frac{\partial\mathbf{T}_m}{\partial(c_{rms}^2)} - 3\mathbf{T}_m^H\mathbf{C}^{-1}\frac{\partial\mathbf{T}_m}{\partial(c_{rms}^2)} - 2\frac{(c_{rms}^2 - \bar{c}_{rms}^2)}{\sigma^2} = 0 \qquad (8.49)$$

where

$$\frac{\partial \mathbf{T}_m}{\partial(c_{rms}^2)} = -\frac{1}{2c_{rms}^4} \times$$

$$col\left\{\frac{x_1^2}{\sqrt{T_0^2 + \dfrac{x_1^2}{c_{rms}^2}}}, \frac{x_2^2}{\sqrt{T_0^2 + \dfrac{x_2^2}{c_{rms}^2}}}, \dots, \frac{x_N^2}{\sqrt{T_0^2 + \dfrac{x_N^2}{c_{rms}^2}}}\right\} = \frac{-1}{2c_{rms}^4}\mathbf{x} \qquad (8.50)$$

Substituting for the derivative in (8.50) we obtain

$$\frac{3}{4c_{rms}^4}\left[\mathbf{T}_m^H \mathbf{C}^{-1}\mathbf{x} - \hat{\mathbf{T}}_x^H \mathbf{C}^{-1}\mathbf{x}\right] = \frac{(c_{rms}^2 - \bar{c}_{rms}^2)}{\sigma^2} \qquad (8.51)$$

We shall now introduce a few approximations. Let the measurement errors be uncorrelated. Then, the covariance function becomes $\mathbf{C} = diag\left\{\sigma_{\tau_1}^2, \sigma_{\tau_2}^2, \dots, \sigma_{\tau_N}^2\right\}$ where $\sigma_{\tau_i}^2$ is the variance of the error in i^{th} measurement. With $\sigma^2 \to \infty$ the right hand side of (8.51) becomes zero. Equation (8.51) may be simplified to

$$\sum_{i=1}^{N}\frac{x_i^2}{\sigma_{\tau_i}^2}\left[1 - \frac{\hat{T}_{x_i}}{\sqrt{T_0^2 + \dfrac{x_i^2}{c_{rms}^2}}}\right] = 0 \qquad (8.52)$$

which is equal to the minimum mean square error estimate [20]. Numerical simulations presented in [20] indicate that the minimum mean square error estimate is very close to the maximum likelihood estimate which is computationally far more involved. Since the unknown quantity c_{rms}^2 in (8.52) occurs inside the square root term, it is not possible to explicitly solve for it. To overcome this, we shall introduce an approximation $T_0^2 << \dfrac{x^2}{c_{rms}^2}$, which

enables us to replace the square root term by $T_0(1 + \dfrac{1}{2}\dfrac{x_i^2}{T_0^2 c_{rms}^2})$ and then solve

for the unknown,

Measurement errors (std)	Mean wave speed	std of wave speed estimate
0.01 sec	1532.2 m/s	16.3 m/s
0.05 sec	1535.1 m/s	102.8 m/s

Table 8.1: Mean and standard deviation of wave speed estimates given by (8.46). Following parameters were assumed: T_0=5.0 sec, c=1500 m/s, Maximum source-receiver separation = 1500 meters, Number of source-receiver pairs=100.

$$\hat{c}_{rms}^2 \approx \frac{1}{2} \frac{\displaystyle\sum_{i=1}^{N} \frac{\hat{T}_{x_i} x_i^2}{T_0^3} \frac{x_i^2}{\sigma_{\tau_i}^2}}{\displaystyle\sum_{i=1}^{N} \frac{x_i^2}{\sigma_{\tau_i}^2} \left[\frac{\hat{T}_{x_i}}{T_0} - 1 \right]} \tag{8.53}$$

The wave speed estimate given by (8.53), on account of approximation $T_0^2 << \frac{x^2}{c_{rms}^2}$, suffers from a large error even for small measurement error. This is demonstrated in table 8.1. The denominator in (8.53) is strongly influenced by the measurement errors for limited array aperture. As noted in the context of focused beamformation, for best depth of focus the array aperture was required to be four times the depth to the reflector. It is understood that a similar requirement exists for wave speed estimation.

8.5.3 Focusing with Optimum Wave Speed: In the previous section (§ 8.3) we saw how CDP gathers are focused at the point of reflection. We were required to know the correct wave speed. We shall now assume that the wave speed is unknown. In fact, we wish to find an optimum speed for which the focus is the best, in the sense that the magnitude of the focused signal is the highest possible. The response function of a focused beamformer, when the wave the speed is unknown, may be obtained from (8.32b)

$$H_N(\omega) = \frac{1}{N} \sum_{n=0}^{N-1} e^{-j\omega \left[T_0 \sqrt{(1+(\frac{\Delta tn}{T_0})^2)} - \hat{T}_0 \sqrt{(1+(\frac{\Delta \hat{t}n}{\hat{T}_0})^2)} \right]} \tag{8.54}$$

where $\Delta \hat{t} = \frac{\Delta x}{\hat{c}}$, $\hat{c}$ is the assumed speed and Δx is the basic unit of source and receiver separation. Note that $\Delta t = \frac{\Delta x}{c}$ is the actual propagation time.
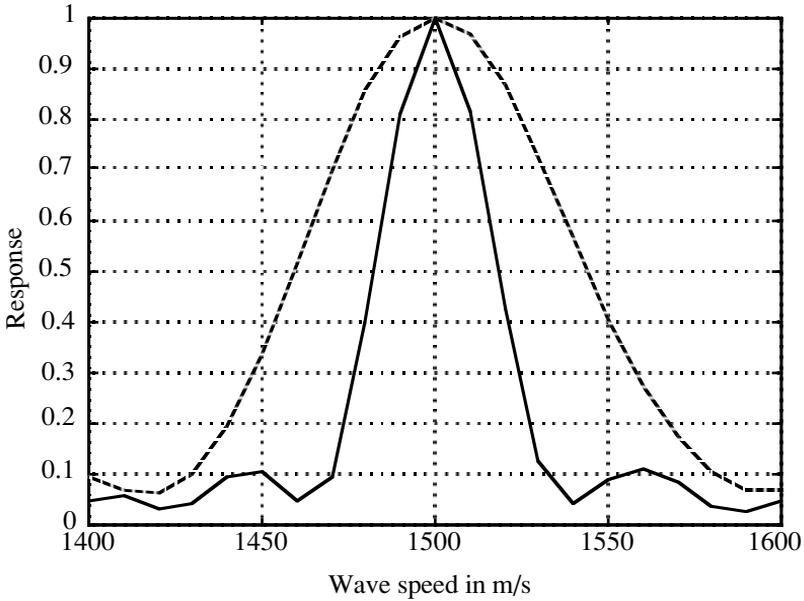
Figure 8.20: Response (magnitude square) of focused beamformer as a function of wave speed. Maximum source sensor separation: 8.5 km (solid line) and 4.5 km (dashed line). Angular frequency=62.84.
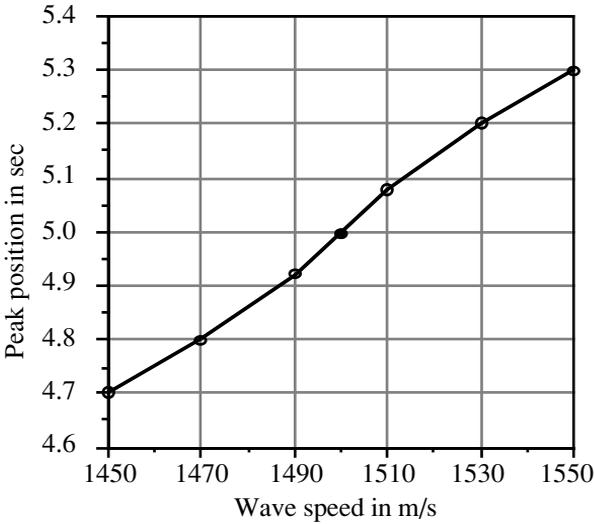

Figure 8.21: Focused beam position as a function of wave speed in m/s. All other parameters are as in fig. 8.20. The maximum sensor separation is 4.5 km.

The maximum value of the response function (4.54) is one and it is achieved only when the exponent is equal to zero, which requires $\hat{T}_0 = T_0$ and $\Delta\hat{t} = \Delta t$ or $\hat{c} = c$. For the present we assume that $\hat{T}_0 = T_0$ and evaluate $\left|H_N(\omega)\right|^2$ for a series of assumed wave speeds. The position of the peak changes with the assumed speed. The correct position is obtained only when the assumed wave speed is equal to the actual speed. The variation of the peak position is illustrated in fig. 8.21.

§8.6 Exercises

1. A point scatterer is located at a depth l unit below the observation surface. A transceiver, placed on the surface, records the round trip-travel time. Show that this travel time is exactly the same as in the CDP experiment with a reflector replacing the scattering point.

2. Common depth point (CDP) seismic gathers containing an echo from a horizontal reflector are cross-correlated. Show that the maximum of the cross-correlation function lies on a hyperbolic time-distance curve. This property is the basis for a technique of seismic speed estimation known as velocity spectrum [21].

3. Show that the slope of the line joining all reflected signals in a zero-offset seismic data over an incline reflector is related to the slope of the reflector through the following relation, $\tan\alpha = \sin\theta_0$ where α is the slope of the line joining reflected signals and θ_0 is slope of the reflector.

References

1. A. J. Devaney, Inverse source and scattering problem in ultrasonics, IEEE Trans., SU-30, pp. 355-64, 1983.

2. D. Loewenthal, L. Lu, R. Robertson, and J. Sherwood, The wave propagation applied to migration, Geophy. Prospecting, vol. 24, pp. 380-399, 1976.

3. J. F. Claerbout, Toward unified theory of reflector mapping, Geophysics, vol. 31, pp. 467-481, 1971.

4. R. H. Stolt, Migration by Fourier transforms, Geophysics, vol. 43, pp. 23-48, 1978.

5. D. Lowenthal and I. D. Mufti, Reverse time migration in spatial frequency domain, Geophysics, vol. 48, pp. 627-635, 1983.

6. G. A. McMechan, Migration by extrapolation of time dependent boundary values, Geophysical Prospecting, vol.31, pp. 413-420, 1983.

7. R. Sun and G. A. McMechan, Pre-stack reverse time migration, Proc. IEEE, vol. 74, pp. 457-467, 1989.

8. J. Claerbout, *Fundamentals of Geophysical Data Processing*, McGraw Hill, New York, 1972.

9. G. Cheng and S. Coen, The relationship between Born inversion and migration for common midpoint stack data, Geophysics, vol. 49, pp. 2117-2131, 1984.

10. E. A. Robinson, Frequency-domain analysis of multodimensional time series data, in Handbook of Statistics, Eds: D. R. Brillinger and P. R. Krishnaiah, vol. 3, pp. 321-342, Elsevier Science Pub. B. V., Amsterdam, 1983.

11. A. J. Berkhout, *Seismic Migration- Imaging of Acoustic Energy by Wave Field Extrapolation*, Elsevier Amsterdam, 1980.

12. P. S. Schultz and J. W.C. Sherwood, Depth migration before stack, Geophysics, vol. 45, pp. 376-393, 1980.

13. W. A. Schneider, Integral formulation for migration in two and three dimensions, Geophysics, vol. 43, pp. 49-76, 1978.

14. R. A. Phinney and D. M. Jurdy, Seismic Imaging of deep crust, Geophysics, vol. 44, pp.1637-1660, 1979.

15. K. Kido, H. Noto, A. Shima, and M. Abe, Detection of acoustic signal by use of cross spectra between every pair of sensors distributed in space. IEEE, ICASSP-86, pp. 2507-2510, 1986.

16. M. T. Taner and F. Koehler, Velocity spectra Digital computer derivation and applications of velocity functions, Geophy, vol. 34, pp. 859-881, 1969.

17. W. A. Schneider and M Backus, Dynamic correlation analysis, Geophy, vol. 33, pp. 105-126, 1968.

18. D. De Vries and A. J. Berkhout, Velocity analysis based on minimum entropy, Geophysics, vol. 49, pp. 2132-2142, 1984.

19. O. Yilmaz and R. Chambers, Migration velocity analysis by wave field extrapolation, vol. 49, pp. 1664-1674, 1984.

20. R. L. Kirlin, L.A.Dewey, and J. N. Bradely, Optimum seismic velocity estimation, Geophysics, vol. 49, pp. 1861-1661, 1984.

21. L. C. Wood and S. Treitel, Seismic signal processing, Proc. IEEE, vol. 63, pp. 649-661, 1975.