THE UNIVERSITY OF CALGARY

ANGLE OF ARRIVAL ESTIMATION IN THE OUTDOOR RADIO ENVIRONMENT

by

RICHARD WALTER KLUKAS

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF ELECTRICAL AND

COMPUTER ENGINEERING

CALGARY, ALBERTA

APRIL, 1993

National Library of Canada

Acquisitions and Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada

Direction des acquisitions et des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

ISBN 0-315-83164-2

THE UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled, **"ANGLE OF ARRIVAL ESTIMATION IN THE OUTDOOR RADIO ENVIRONMENT"**, submitted by Richard Walter Klukas in partial fulfillment of the requirements for the degree of Master of Science.

Supervisor, Dr. M. Fattouche
Dept. of Elec. and Comp. Engg.

Dr. A. Sesay
Dept. of Elec. and Comp. Engg.

Dr. S. T. Nichols
Dept. of Elec. and Comp. Engg.

Dr. G. Lachapelle
Dept. of Geomatics Engg.

Date: April 6, 93

ABSTRACT

AOA estimation using MUSIC, an eigendecomposition based superresolution algorithm, is investigated for the multipath radio environment. Bearing estimation is necessary for finding the location of a mobile cellular telephone by triangulation. The outdoor, UHF, multipath radio channel is assumed to consist of numerous clusters and a LOS component. Three different techniques to identify the LOS cluster and estimate its AOA are presented. Simulated data is used to evaluate the performance of these three techniques as well as a technique to estimate the AOA component within a cluster. The results indicate that good accuracy can be achieved in a multipath environment. In order to eliminate all ambiguities, a two dimensional array must be used. Whereas a virtual array created by travel along a straight path requires at least two antennas, it is demonstrated that a single antenna with a non-linear trajectory can create a sufficiently two dimensional virtual array.

ACKNOWLEDGMENTS

I would like to thank Dr. M. Fattouche for his enthusiasm, encouragement, guidance and support throughout this work. Thanks also to Dr. A. Sesay for discussions regarding the various aspects of the academic life and to Dr. G. Lachapelle for information relating to navigation and location finding.

The financial support of the Natural Sciences and Engineering Research Council of Canada, the University of Calgary, and TRLabs is gratefully acknowledged. Further thanks to TRLabs for the facilities provided.

I also thank my wife Deborah, my daughter Belinda, and my parents Helmut and Tina Klukas, for providing me with a home and family life conducive to my academic pursuits.

DEDICATION

I dedicate this thesis to my wife Deborah,

a most precious gift from God

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS & ABBREVIATIONS

α	complex scalar
$\boldsymbol{\alpha}$	vector of complex scalars
$\mathbf{a}(\theta)$	array steering vector for angle θ
a_i	response of array element to i^{th} wavefront or amplitude of i^{th} sinusoid
$a_n(\theta)$	complex gain of the n^{th} array element for angle θ
$\mathbf{A}$	matrix of steering vectors
AOA	Angle Of Arrival
AR	Autoregressive modelling
ARMA	Autoregressive - Moving Average model
β_n	n^{th} eigenvector of data correlation matrix
$\mathbf{b}$	noise subspace vector with minimum Euclidean norm
BW	standard beamwidth
$\mathbf{B}^l$	l^{th} power of time of arrival diagonal matrix
c	radio wave propagation speed
cm	centimetres
C	high level programming language
C/A	Course Acquisition code
CDMA	Code Division Multiple Access
CW	Continuous Wave
δ_k	complex delay of the k^{th} signal over the entire array span
$\boldsymbol{\delta}$	vector of complex delays
d	physical spacing between array elements or antennas

$d_{j,k}$	element of data matrix in the j^{th} row and k^{th} column
d_n	n^{th} data point in a data series
dB	decibel
dBW	power relative to one Watt
$\mathbf{D}$	data matrix
DGPS	Differential Global Positioning System
E	ensemble average
f_s	sampling frequency
f_{jk}	differential Doppler shift for the j^{th} and k^{th} arrivals
$f(\mathbf{x}\vert\theta_{AOA})$	likelihood function
$\mathbf{f}(\omega)$	frequency scanning vector
FBLP	Forward-Backward Linear Prediction method
FDMA	Frequency Division Multiple Access
GPS	Global Positioning System
$\mathbf{I}$	identity matrix
IS-54	digital cellular mobile radio standard
kHz	kilohertz
km	kilometre
km/hr	kilometres per hour
K	number of signals
λ	signal wavelength
λ_c	carrier wavelength
λ_n	n^{th} eigenvalue of data correlation matrix
L	number of forward subarrays
LORAN C	LOng RAnge Navigation system
LOS	Line Of Sight

μs	microsecond
m	number of coherent signals
ms	millisecond
M	number of subarray elements or FIR filter order
MA	Moving Average model
ME	Maximum Entropy
MHz	megahertz
MLE	Maximum Likelihood Estimation
MUSIC	MUltiple Signal Identification and Classification
MVDR	Minimum Variance Distortionless Response
$\mathbf{n}(t)$	array noise vector
$n_j(t)$	noise at j^{th} array element
$\mathbf{n}_l(t)$	noise vector for l^{th} forward subarray
$\tilde{\mathbf{n}}_l(t)$	noise vector for l^{th} backward subarray
N	number of array elements or data points in a time series elements
ω	angular frequency
ω_i	Doppler frequency shift of i^{th} sinusoid
ω_o	centre angular frequency
ω_{MUSIC}	normalized Doppler shift of a peak in the MUSIC spectrum
ϕ_{ab}	signal phase difference between antennas a and b
ϕ_e	electrical phase angle
ϕ_i	phase of i^{th} sinusoid
P	Precision code
P	number of physical elements
PEF	Prediction Error Filter

PETRA	PEak TRacking Algorithm
QUIKTRAK	a hyperbolic, time difference, automatic vehicle location system
$r_j(t)$	signal received at the j^{th} array element
$\mathbf{r}(t)$	array output vector
$\mathbf{r}_l^b(t)$	output vector of the l^{th} backward subarray
$\mathbf{r}_l^f(t)$	output vector of l^{th} forward subarray
R_{jk}	correlation between the output signals of the j^{th} and k^{th} array elements
$\mathbf{R}$	ensemble averaged data correlation matrix
$\hat{\mathbf{R}}$	estimate of data correlation matrix
$\tilde{\mathbf{R}}$	forward/backward smoothed data correlation matrix
$\mathbf{R}^b$	backward spatially smoothed data correlation matrix
$\mathbf{R}^f$	forward spatially smoothed data correlation matrix
RF	radio frequency
σ^2	noise variance
$s_i(t)$	i^{th} wavefront signal
$\mathbf{s}(t)$	matrix of signal vectors
S	number of samples per array element
$\mathbf{S}$	signal correlation matrix
$\tilde{\mathbf{S}}$	forward/backward smoothed signal correlation matrix
$\mathbf{S}^b$	backward spatially smoothed signal correlation matrix
$\mathbf{S}^f$	forward spatially smoothed signal correlation matrix
S.A.	Selective Availability
Sa	second dimension of subarrays
SNR	Signal to Noise Ratio

SURP	Simulation of the Urban Radio Propagation Channel software package
τ_i	interelement time of arrival difference for arrival angle i
θ	spatial angle
θ_{AOA}	angle of arrival
t_a	code transmission time
t_o	code arrival time
Tx	transmitter
TDMA	Time Division Multiple Access
v	array velocity
V	number of virtual elements or mobile's velocity
$\mathbf{V}_N$	matrix of all noise eigenvectors
$\mathbf{x}$	observed array signal vector

CHAPTER 1
INTRODUCTION

1.1 Location Finding

The art of location finding has been a popular field of study for centuries. Man has always wanted to know where he was in order to navigate to some other point. Equally important to man has been the location of natural and man-made objects. As a result, a myriad of different methods have been developed over time and the choice of the method has depended on the application. In the age of precise location finding, the methods have progressed from the sextant, which relied on natural satellites, to modern day systems which use artificial or man-made satellites.

The performance of a location finding method is not necessarily a fixed parameter. Continual research is producing and improving technology which in turn may be used to improve location finding methods which were developed long ago. This is precisely the aim of this thesis; not to present a new or novel method, but to apply recently developed technology to an age old technique. In this way, an older method which might have been overshadowed by more recent techniques, may be rejuvenated and given new life for modern applications.

1.2 Objective and Motivation

The main objective of this thesis is to estimate the angle of arrival (AOA) of radio waves in a multipath environment. If the bearings of multiple transmitters can be determined in this way, the location of the receiving antenna may be determined by triangulation. The intended application is to locate a mobile cellular telephone subscriber within the outdoor cellular telephone geographical grid. The base stations act as the transmitters and the AOA of the radio waves arriving at the mobile are used to determine its location.

The location of a mobile in the cellular system is required for two reasons. To initiate a telephone call, the mobile must know with which cell site to begin communication and vice versa. Secondly, while a call is in progress, the mobile may very likely travel from cell to cell. The call must therefore be handed off from base station to base station. As the number of cellular subscribers increases, the geographical size of the cells must decrease in order to increase capacity. As a result, the current methods of tracking the location of mobiles by paging and look up tables, will become far too slow to be practical. Faster means of determining which cell a subscriber is currently in, must be developed.

Decreasing cell size will also place limits on the required accuracy of a location finding system. It is anticipated that in the future, microcells will have radii in the order of 1 km. Therefore, the positional accuracy required would be 100 metres or less.

Although triangulation has been unpopular in the past for high accuracy applications, new superresolution algorithms, which are able to estimate AOA very accurately, may change this. The algorithm investigated in this thesis is MUSIC (MUltiple Signal Identification and Classification). A triangulation system using MUSIC is advantageous in that it is narrowband in nature. The transmitted signals need only be tones. In addition, very little additional RF hardware, if any, would be required. The result would be a low cost system which could be implemented immediately.

The outdoor radio propagation channel is characterized by multipath. Reflections and diffraction due to man made objects and topography, cause the transmitted signal to follow various indirect paths in addition to the line of sight (LOS) path. The result is a number of highly correlated signals with various arrival angles. In some cases the LOS path may not even exist.

Consequently, the multipath nature of the radio propagation channel is a serious problem when attempting to locate by triangulation. In order to triangulate in a multipath environment one must assume that a LOS path does exist. This assumption made, two questions must then be answered: 1) which of the many arriving signals followed the LOS path, and 2) what is the AOA of that LOS arrival. In this thesis three techniques incorporating MUSIC attempt to answer these two questions.

1.3 Thesis Overview

Chapter 2 discusses various methods of location finding. It then focuses on triangulation and the means by which AOA may be estimated. The spatial and temporal forms of MUSIC, as well as the relevant theory, are described in detail in chapter 3. Chapters 4, 5, and 6 present and analyze the simulation studies conducted. The case of estimating AOA in the presence of multipath, while moving in a straight line, is addressed in chapters 4 and 5. Chapter 6 deals with estimating the AOA of a single arrival while travelling along a non-linear path.

Finally, the major conclusions of the thesis and recommendations for further work are presented in chapter 7.

CHAPTER 2
VEHICLE LOCATION TECHNIQUES

2.1 Introduction

A wide variety of techniques exist for automatically locating vehicles. These methods have been developed in response to the needs specific to the intended application. Applications for vehicle location systems include fleet management, security for vehicles transporting important cargo or persons, and location of cellular subscribers within the cellular network. Although the intent of this thesis is to investigate technology to apply to the cellular subscriber location problem, this does not exclude the possibility of using this technology for fleet management and security purposes.

2.2 Dead Reckoning

Automatic vehicle location methods can be divided into three general classes: dead reckoning, proximity detection, and radio signal location [1]. Dead reckoning systems continually update a vehicle's location by monitoring changes in direction and distance travelled from a known starting point. This is commonly done with odometers and compasses. After initialization at the known reference point, the vehicle's position is determined by vectorially summing all changes in direction and distances travelled as the vehicle moves away from the reference point. Since all errors in the sensing of direction change or distance travelled will accumulate, periodic reinitialization is necessary at the known reference points dispersed throughout the service area. Error accumulation and the consequential reinitialization are the major drawbacks of this method.

2.3 Proximity Detection

Proximity detection systems use radio signposts installed along likely routes of vehicular traffic. In one of the two types of proximity detection, the radio signposts continually transmit a low power radio signal that identifies the signpost. As a vehicle passes a signpost, it detects the signpost's signal and thereby identifies the signpost. Since the signpost's location is known, the vehicle's location is known. The vehicle may then relay this information to the central processing station. For a mobile cellular system however, this method is impractical due to the high cost of installing and maintaining a signpost network.

2.4 Radio Signal Location Techniques

There are numerous location techniques which fall under the radio signal location category. In all of these methods, radio signals are transmitted between the mobile vehicle and one or more fixed stations. Radio signal location techniques can be divided into three groups: radio ranging or radar, radio trilateration, and radio direction finding or triangulation.

2.4.1 RADAR

Radar determines both the range of a vehicle and its bearing. At the fixed site a directional radio signal is swept through a specified area. Any target within the area covered will reflect the signal back to the fixed site. Because a LOS path between the transceiver and the target is essential, radar is not practical for land mobile applications in which multipath propagation and obstructions are encountered. Reflections from the large number of structures and obstacles found in an urban environment will render useless the information received at the transceiver.

2.4.2 Radio Trilateration

Radio trilateration technologies form a second group of radio location techniques. These methods rely on a quasi-constant wavefront velocity environment. By measuring the propagation time between the mobile and at least three fixed sites, the two dimensional position of the mobile may be determined as the point of intersection of three circles, centered at the sites, and each with radius equal to the propagation distance between the vehicle and the respective fixed site.

A simpler method is to use difference in time of arrival instead of propagation distances. This is termed time difference hyperbolic location. The locus of points for which the difference in propagation time between the mobile and one site and the mobile and another site is equal, will form a hyperbola. Therefore, with three sites, at least two hyperbolas are available and the position of the mobile will be the intersection of the two hyperbolas.

2.4.2.1 LORAN C

The LOng RAnge Navigation system (LORAN C) is a marine navigation system that employs time difference hyperbolic trilateration [1]. It has been in use since 1956 and is operated by the United States and Canadian coast guards. In the 1970's LORAN C was adopted as the primary marine radio navigation system by Canada and the United States [2]. Marine LORAN C coverage includes the Pacific, Atlantic, and gulf coasts of North America, the North

Atlantic, the North Sea, the Mediterranean Sea, as well as parts of the Asian Pacific coast. The system is comprised of chains of land based transmitter sites [3]. A chain consists of a master transmitter and from two to four secondary transmitters. These sites transmit amplitude modulated pulses in the 90 to 110 kHz frequency band. A LORAN C receiver will receive the transmissions from one such chain and calculate time differences between the transmitters in the chain. The navigator may then use these time differences to make a location fix on navigational charts imprinted with LORAN C time difference LOP (Lines Of Position).

Although LORAN C was designed as a marine navigational system, the introduction of two mid-continental chains in the U.S. as of 1991 [4, 5], has resulted in LORAN C signal availability throughout the continental U.S. as well as a marked increase in coverage in western Canada. Now there is almost uninterrupted LORAN C coverage for land vehicle users in southern regions of Canada as shown in Figure 2-1. Transport Canada is presently investigating the potential benefits of adding more LORAN C transmitters in Canada [2].

Therefore, LORAN C may be suited to the problem of land vehicle location. An emergency vehicle location system, using LORAN C has been operational in Detroit since 1989 [6]. The system is able to estimate the location of emergency vehicles with an accuracy of 200 metres. Of prime importance for cellular mobile telephone location are accuracy, reliability, and coverage. The reliability of LORAN C transmitters is greater than 99%, which is more than adequate. However, overall system reliability also depends on adequate coverage or signal availability throughout the travel path of the mobile cellular telephone.

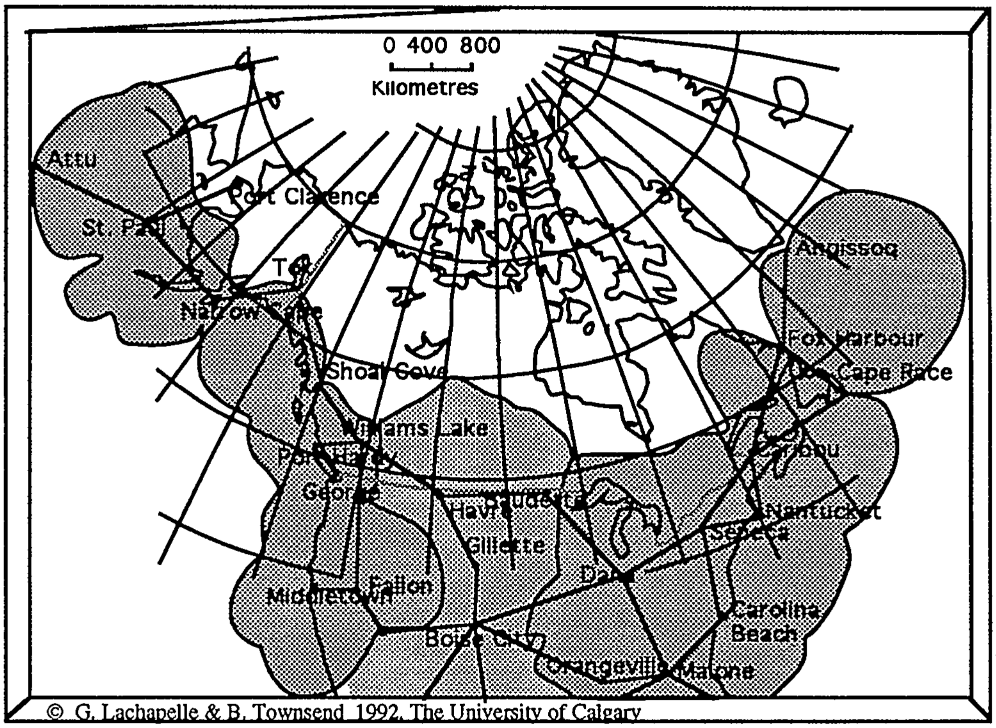

Figure 2-1 North American LORAN C Coverage

Lachapelle has conducted studies on LORAN C signal availability in the urban environment [4], as well as in mountainous regions [7, 8]. His findings indicate that signal availability in both environments is limited due to attenuation of the signal by large buildings and mountains. In the urban setting, signal degradation also results from multipath scattering and interference from power line carriers that transmit in the same frequency band [5]. In the city of Calgary, signal availability was 50% to 60% in the city core and 95% to 100% in residential areas. In mountainous regions the signal availability varied between 65% and 95%.

In Canada, the nominal accuracy of LORAN C is 500 metres or less. Rugged topography causes phase distortion which can result in errors of many hundred metres. Therefore, in most cases a stand alone, single chain, LORAN C system is not able to offer the degree of accuracy required for a microcellular mobile telephone system. However, the phase distortion effect is permanent and may therefore be calibrated for. Lachapelle reports that, with en-route calibration using GPS (Global Positioning System), the accuracy of LORAN C, in mountainous areas is approximately 50 to 100 metres [7]. Using more than one LORAN C chain to make a fix will also improve the accuracy of the location fix.

A further requirement of a mobile cellular telephone location system is small, lightweight, and inexpensive user equipment. Currently a LORAN C receiver may be packaged on a 6 inch square board and costs roughly $500 or less. However, to obtain the necessary accuracy, LORAN C will have to be combined with some other navigational aid (e.g. GPS) and this will increase the equipment cost beyond what is practical for mobile cellular telephony.

2.4.2.2 Spread Spectrum Techniques

Also included in the group of trilateration technologies are those that employ spread spectrum signals. Spread spectrum signals have recently received much attention. CDMA (Code Division Multiple Access), a spread spectrum multiple access technique, has risen to be a serious contender with TDMA (Time Division Multiple Access) and FDMA (Frequency Division Multiple Access). Direct sequence spread spectrum signals are generated by multiplying the signal to be transmitted with a high frequency pseudorandom code. The effect of this is to spread the bandwidth occupied by the signal. Consequently, spread spectrum signals are wideband. At the receiver, the same code is generated and used to despread and thus recover the signal. If each user in a communication system has a different orthogonal code, all users may use the same communication channel at the same time without interference from each other.

2.4.2.2.1 GPS

Spread spectrum signals may also be used for location finding. The GPS (Global Positioning System) is a position fixing spread spectrum based system which uses satellites [9]. It has been under development by the U.S. Department of Defense since 1973. When complete, 21 satellites plus 3 spares will orbit the earth at an altitude of 20,200 km. Presently, 18 satellites are in place. The satellites are distributed in orbit such that 5 to 8 satellites will be "in view" at any point on the earth almost all the time. By simultaneously receiving transmissions from at least 4 satellites, a user may determine his 3 dimensional position as well as the receiver time bias. The GPS constellation is illustrated in Figure 2-2.

The GPS is a ranging system. The pseudo-distance from each satellite to the point to be located is determined from the reception of the satellite transmissions at the receiver. Trilateration is then used to determine the point's location. Satellite transmissions occur on two frequencies; L1 = 1575.42 MHz and L2 = 1227.6 MHz. Each satellite transmits two pseudorandom codes. The P (Precision) code has a chip rate of 10.23 MHz and a code length of 10 days whereas the C/A (Coarse Acquisition) code has a chip rate of 1.023 MHz

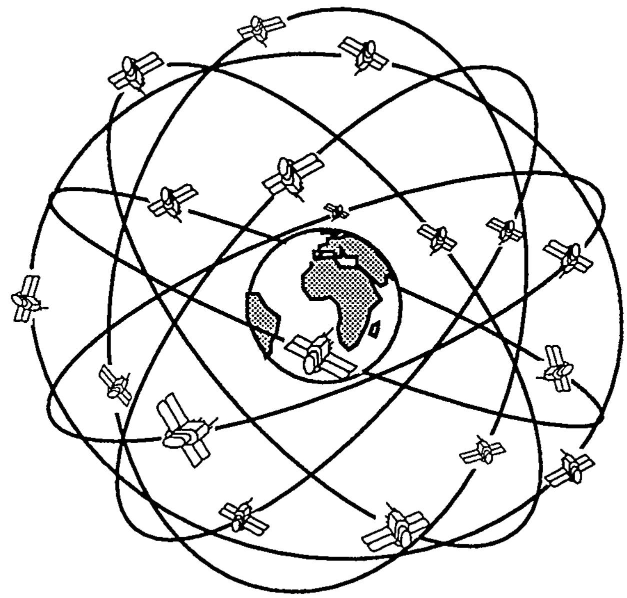

Figure 2-2 The GPS Constellation

and a code length of 1 ms. The P code is transmitted on both the L1 and L2 frequencies while the C/A code is only transmitted on L1. Transmission on two separate frequencies allows for the correction of errors introduced by refraction in the ionosphere. Consequently, position fixes made using the P code will be more accurate than those made with the C/A code. As a result, use of the P code is restricted to the military whereas the C/A code is available for civilian use.

The range between a GPS satellite and a GPS receiver is found from an estimate of the propagation time. The propagation path length is merely the product of the propagation time and the speed of the radio signal which is usually taken to be approximately 300,000 km per second. If the beginning of the C/A code transmitted from the satellite is at time t_o, and the code arrives at the receiver at time t_a, the propagation time is merely $t_a - t_o$. This however requires the receiver to know t_o. Latitude, longitude, elevation, and t_o are the four unknowns to be solved for. Hence the requirement for pseudo-range measurements from four satellites.

To find t_a, the code transmitted from the satellite is synchronized with a replica of the code generated by the receiver. The replica code is shifted from t_o one chip at a time until the correlation between the replica and transmitted code is a maximum indicating synchronization. For the C/A code, each shift corresponds to a distance of 293 metres compared to 29.3 metres for the P code. Therefore, a more precise estimate of the propagation time can be achieved using the P code.

The instantaneous single point accuracy of GPS in the C/A code mode is from 20 to 50 metres [5] which is adequate for the microcellular mobile telephone application. However, the horizontal, 2 dimensional accuracy achievable by civilians has been intentionally degraded to approximately 100 metres by Selective Availability (S.A.). This may not be accurate enough for the location of microcellular telephone subscribers. Accuracy can be significantly improved with DGPS (Differential GPS) [9]. Because the satellites are at such a high altitude (20,200 km), two points on the earth's surface, even 100 km apart, will essentially have the same propagation path to a particular satellite. Hence, the errors caused by refraction in the ionosphere will generally be the same for a transmission from the satellite to each of the two points. If the location of one of the points is known exactly, these errors are known and can be compensated for in making a position fix on the other point. Accuracy of ≤ 5 metres is possible using DGPS [2].

The GPS requires LOS propagation paths between the satellites and the receiver. As a result, GPS signal availability is susceptible to masking by large objects such as high-rise buildings and mountains. Field tests [4] indicate that GPS signal availability in urban residential areas is in the order of 70% and in city cores it is only 40%. In mountain regions [7, 8], the signal availability is 60%. Therefore, GPS signal availability is consistently less than that of LORAN C and far too low for a vehicle location system.

A possible solution is to combine GPS with another system such as LORAN C. In mountain regions, a hybrid GPS/LORAN C system would have a signal availability of approximately 90%, which is an improvement over both LORAN C and GPS individually, but still low for cellular telephone applications [7]. Interoperable GPS/LORAN C receivers are available but still too costly [5].

A land based automatic vehicle location and navigation system based on the integration of dead reckoning, digital map matching, and GPS has been investigated [10]. The system is based on differential odometry with location updates provided by map matching and GPS. The accuracy of the dead reckoning component in stand alone mode was 1% to 2% of the distance travelled. The addition of map matching updates gave negligible improvement. However, with both map matching and GPS coordinate updates, the system accuracy was maintained at approximately 40 metres.

Though able to offer a high degree of accuracy, this system, and in particular the map matching aspect, is complex. It is also hardware and software intensive.

To conclude, GPS has the potential to offer extremely high accuracy should S.A. be removed or DGPS implemented. In addition, reasonably small GPS receivers are available for approximately $600 and the system is operable anywhere in the world. However, GPS signal availability is far too low, even in a hybrid system, for the reliability required by telephony.

2.4.2.2.2 Other Spread Spectrum Methods

Spread spectrum signals may also be used in ground-based location systems. Goud [6] investigated the performance of a spread spectrum radio location technique, applied to the mobile cellular radio location problem. His system uses time difference hyperbolic trilateration of direct sequence spread spectrum signals. The spread spectrum signals would be transmitted by the mobile over a 10 MHz multipath mobile radio channel typical of a dense urban environment. Computer simulations showed that the absolute location error for locating a mobile in a 3 km radius circular service area, with 4 sensors on the circumference, is 50 metres. When the service area was overlaid with a cellular grid of hexagonal cells of radius 500 metres, the system was able to locate the cell in which a mobile was 94.8% of the time.

A spread spectrum, time difference, hyperbolic automatic vehicle location system has been implemented in the Sydney Australia area [11]. It is called QUIKTRAK and has been in operation since 1987. Although the system was developed primarily for fleet management, other applications are possible. The service area is approximately 2000 square kilometres and the system accuracy is roughly 30 metres. The system is capable of supporting thousands of vehicles and making 30,000 location fixes per hour. The direct sequence spread spectrum signals transmitted by the mobiles occupy a bandwidth of 2 MHz.

Although spread spectrum methods have tremendous potential, their processing complexity and need for large amounts of bandwidth continue to be of concern. A narrowband system is far simpler to design and implement.

2.4.3 Radio Direction Finding

The third group of radio location techniques is called radio triangulation or radio direction finding. Unlike the previous two groups, this group is narrowband in nature and therefore does not require any new spectrum other than the allocated cellular bands. This group estimates position by finding the AOA of radio waves. The concept is illustrated in Figure 2-3. Note that in Figure 2-3, it is intended that AOA be measured only in the horizontal plane. Throughout this thesis AOA refers to azimuth not elevation.

Continuous wave (CW) radio signals are transmitted in all directions by transmitters A and B. The signals are received at the mobile and the AOA with reference to some direction (such as North) is determined for each transmitter. The position of the mobile is simply the intersection of the bearings for the two transmitters.

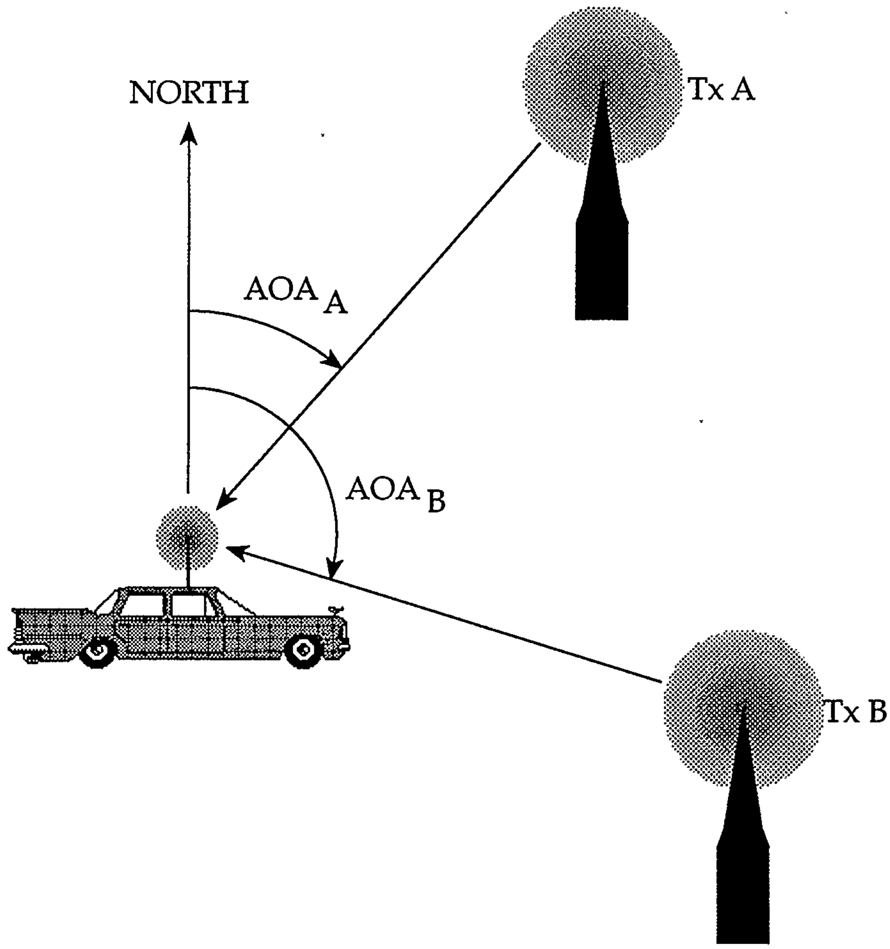

Figure 2-3 Location Finding by Triangulation

Obviously the success of simple triangulation hinges on the existence of one direct propagation path between each transmitter and the mobile. Traditionally, triangulation has not been used for land vehicle location systems because of the multipath nature of the radio propagation environment in urban areas. Reflections of radio signals by obstacles such as large buildings or rugged topography can cause large positional errors. The accuracy of the position fix of course depends on the accuracy of the AOA. As the distance between the transmitters and the mobile increases, so does the error in absolute location due to an error in AOA.

Although much of the arrival direction of radio waves in the multipath environment is still unknown, recent research [12] indicates that radio waves originating from a single source and scattered or reflected, tend to arrive at a distant point in clusters. Indoor measurements reveal that two distinct arrival times generally dominate. The first cluster to arrive generally has the larger magnitude and can be attributed to the direct ray as well as reflections near the antenna. The second cluster is composed of reflected rays which travel longer and more attenuated paths.

The outdoor multipath radio environment is also assumed to consist of two clusters. The digital cellular mobile radio standard IS-54 [13] states that a receiver should expect two separate equal powered arrivals separated by at least 41.2 μs or one symbol interval. The frequency range is 800 to 900 MHz. More likely, if the first cluster to arrive follows a reasonably direct path it will be of greater power than any other cluster which follows an indirect path. Therefore, if a LOS cluster of higher power than any other arriving cluster is assumed to exist, a triangulation method which is able to discern this cluster should be able to make a reasonably accurate estimate of a mobile's position.

AOA may be determined by any one of a host of methods. The application of some of these methods is two-fold. First, these methods can operate in the temporal domain to estimate the frequencies of complex sinusoids in noise. This requires a time series of data. Second, estimating the AOA of planar wavefronts is a spatial problem that requires data from an array of sensors. These methods may be divided into two groups. Parametric methods are those which assume some sort of model for the data process. Nonparametric techniques make no such assumptions about the data. Historically, many of the nonparametric techniques were developed before the parametric methods. That no assumptions about the input process are made and that they are relatively simple, are perhaps the two chief advantages of nonparametric methods. However, the resolution of these methods is limited and it is in this area that under certain conditions, the more modern parametric methods excel. Discussion begins with some nonparametric methods.

2.4.3.1 Interferometry

An interferometer such as an Adcock array is a very simple direction finding system [14]. It consists of two pairs of orthogonally positioned antennas. The antennas are positioned north-south and east-west. If ϕ_{NS} is the signal phase difference between the north and south antennas, and ϕ_{EW} is the same for the east-west antennas, then θ_{AOA} is calculated with the following equations:

$$Adcock(\theta_{AOA}) = \arctan(\phi_{EW}/\phi_{NS}) \tag{2-1}$$

where $\phi_{NS} = (2\pi/\lambda)d_{NS}\cos(\theta_{AOA})$

$\phi_{EW} = (2\pi/\lambda)d_{EW}\sin(\theta_{AOA})$

$d_{NS} =$ separation between north - south elements

$d_{EW} =$ separation between east - west elements

$\lambda =$ signal wavelength.

The Adcock array is able to process only one signal. It is therefore unsuitable for a multiwave environment such as the mobile cellular radio environment. Its other disadvantages include, bearing accuracy as a function of bearing, bearing ambiguity with larger apertures, and poor results when the wavefront is distorted.

2.4.3.2 Beamformers

Beamformers have long been used for AOA determination. The antennas of an array are so phased such as to form a beam in a certain direction. The beam is scanned through the space of interest either mechanically or electronically. The output of the array will peak when the beam is pointing in the direction of an incoming signal. The main drawback of conventional beamformers is resolution which is limited to the beamwidth of the main lobe of the array response. The array response also includes side lobes which can further reduce resolution. A large signal picked up by a side lobe could overshadow a smaller signal received by the main lobe.

A system to estimate vehicle positions using multibeam antennas has recently appeared in the literature [15]. This system was specifically designed and tested to locate subscribers in a mobile cellular radio system. The AOA of signals transmitted by a mobile are estimated at two or more base stations using multibeam antennas.

Each multibeam antenna consists of six beams spread over 120 degrees of azimuthal angle. At any one time, three beams are of interest: the beam with the strongest signal level, the beam on its immediate left, and the beam on its immediate right. The algorithm described in the reference follows a signal level comparison process among the three beams to estimate the AOA. The method assumes that a single wave is received. The authors assume that errors will result with multiple arrivals.

The system was tested with three base stations in a 20 square kilometre area of Tokyo. The free space accuracy of the system, tested with transmitters located on tall buildings, was found to be less than 100 metres. These errors are contributable to sidelobes of the antennas, signal level measurement accuracy, and the performance of the algorithm for AOA estimation.

For a test car driving through the streets, the accuracy was considerably worse. The main reason is of course the multipath propagation environment. Errors were caused by reflections from buildings as well as confusion in AOA by multiple arrivals from various directions even when LOS existed. Errors from these causes were found to be from 500 to 1000 metres.

As a result, the rms position error was found to be over 300 metres. This was reduced to approximately 200 metres by a combination of time series and positional averaging. The authors conclude that this method is only suitable for systems with cell radii in the order of several kilometres.

2.4.3.3 Fourier Transform Methods

The most well known nonparametric methods of determining AOA are those based on the Fourier transform [16, 17]. There are two conventional ways to use the Fourier transform to this end. The first is to perform the Fourier transform, with respect to electrical phase angle, on the spatial series of data (i.e. antenna array output). The power spectrum as a function of electrical phase angle is then calculated as the magnitude squared of the Fourier transform. This is called the periodogram. Peaks in the periodogram correspond to directions of arrivals. The actual AOA, θ_{AOA}, is related to the electrical phase angle, ϕ_e, by the relation,

$$\phi_e = \left(\frac{2\pi d}{\lambda}\right)\sin(\theta_{AOA}) \tag{2-2}$$

where d is the interelement spacing.

The Blackman-Tukey method is the second conventional Fourier method. The autocorrelation function of the data is first found and then the Fourier transform is performed on the autocorrelation function. The autocorrelation and periodogram methods both give similar results.

The major limitation of Fourier based methods is resolution. For a linear array of finite aperture, the radiation pattern due to a single source illuminating the array, consists of a main lobe and a number of sidelobes. As previously mentioned, sidelobes are undesirable. Though their amplitude can be reduced by window functions, this also increases the width of the main lobe. A standard beamwidth (BW) is defined as the angular separation between the dominant peak and the first null in the array response. In equation form,

$$\mathrm{BW} = \frac{2\pi}{N} \tag{2-3}$$

where N is the number of array elements [17].

The larger the array, the narrower the beamwidth of the response resulting in increased resolution. If there are two or more sources illuminating the array, the sources will be resolved if they are separated by two or more beamwidths. If separated by one to two beamwidths, they may be resolvable. Generally, Fourier methods are unable to resolve sources closer than one beamwidth. Therefore, resolution is limited by the width of the main lobe which is limited by the size of the array and any window functions which are employed.

Fourier methods do have advantages which make them attractive for some applications. They are nonparametric and hence no input process model is required. As a result, they can be used with any type of signal. They are robust in that they are insensitive to parameter changes, and they are relatively simple to implement.

2.4.3.4 Maximum Likelihood Estimation

Maximum Likelihood Estimation (MLE) is a parametric technique that may be used to directly estimate AOA [17]. Let **x** be the observed signal vector from the antenna array and θ_{AOA} be the AOA, information of which is contained in **x**. The likelihood function is then defined as the joint probability density function $f(\mathbf{x}|\theta_{AOA})$, of **x** given θ_{AOA}. In other words, the AOA is estimated as that angle which most likely produced the observed signal vector **x**.

MLE overcomes the resolution limitation of the Fourier methods. Also, an MLE estimate of AOA, if unbiased, will achieve the Cramér-Rao lower bound with equality [17]. However, the practicality of MLE is limited by the complexity of maximizing the likelihood function when there are more than two arrivals. This difficulty in multidimensional maximization usually excludes the use of MLE in multipath environments.

2.4.3.5 Superresolution Algorithms

Recently, a number of parametric methods of estimating AOA, which have sub-Rayleigh resolution, have been developed. The Rayleigh resolution criterion was developed by Lord Rayleigh in 1879. For angular spectrum applications it states that two arrivals are considered resolved when the first maximum of one arrival coincides with the first minimum of the second.

The Rayleigh resolution criterion is generally accepted as the criteria by which the resolving ability of AOA estimators is judged. Whereas Fourier based methods cannot exceed the Rayleigh criterion, those parametric methods that do are aptly named superresolution methods. However, to achieve the resolution and accuracy potential of these methods, the models assumed by these methods must closely match the actual physical model from which the data comes. Four fundamental superresolution methods are considered: Autoregressive (AR) modelling, Minimum Variance Distortionless Response (MVDR), Minimum-Norm, and MUSIC.

2.4.3.5.1 AR Spectrum Estimation

AR spectrum estimation is based on linear prediction. A linear predictor or Prediction Error Filter (PEF) may be used to model a stochastic process. For an AR process, the inverse transfer function of the PEF yields a transversal filter or a lattice filter whose tap weights or reflection coefficients have been chosen such that white Gaussian noise at the input of the filter will produce a sample of the process of interest at the output. The spatial power spectrum may be determined from the transfer function of the PEF. Hence, AR spectrum estimation is an indirect method of determining AOA. Unlike Fourier based methods however, the amplitudes of the peaks in an AR spectrum do not correspond to signal powers. The AR spectral estimate is also called the Maximum Entropy (ME) spectral estimate. For a linear array, the AR spectrum is identical to the spectrum obtained by maximizing the entropy of the data process.

In addition to the AR model, which is implemented as an all pole filter, there is also the MA (Moving Average) model and the ARMA (autoregressive - moving average) model. The MA model is implemented as an all zero filter whereas the ARMA model is a combination of the AR and MA models. The AR model is most often used since its coefficients are computed from the Yule - Walker equations which are linear. MA and ARMA coefficients must be computed from non-linear equations which are difficult to solve.

There are three principal algorithms used to determine the AR coefficients. The Levinson-Durbin algorithm indirectly determines the coefficients. First the autocorrelation function of the data is estimated. The Levinson-Durbin algorithm then computes the tap weights for a transversal filter and then the reflection coefficients for a lattice filter. The Burg algorithm is a more direct method which estimates the reflection coefficients directly from the data.

Although it has excellent resolution and is very efficient, the Burg algorithm has three drawbacks [16, 17]. First, depending on the initial phase of the input signal and the length of the data, the Burg algorithm may give biased results. Second, a phenomenon known as line-splitting is known to occur with Burg. Line-splitting is the appearance of two closely spaced signals when there is in fact only one. Lastly, the Burg algorithm may fail to work properly for coherent arrivals such as those encountered in a multipath environment. For two coherent arrivals separated by less than a BW, the Burg algorithm will fail to resolve them properly unless the phase difference between them is an odd multiple of 90°; in that case the process is spatially stationary.

An alternative to the Burg algorithm is the Forward-Backward Linear Prediction method (FBLP). The FBLP method works for both stationary and nonstationary sources. Therefore, it is able to resolve two coherent arrivals which are separated by less than a BW, provided that the SNR is high enough. The FBLP method does have two drawbacks [17]. Unlike MLE, the FBLP method does not achieve the Cramér-Rao bound. Furthermore, below a threshold reached at high SNR, there is a serious degradation in the performance of the method. A modified FBLP method was developed by Tufts and Kumaresan in order to minimize these limitations. The improvement is substantial [19].

For a two ray multipath environment, the modified FBLP method works very well. For diffuse multipath however, it is not able to resolve for the numerous arrivals. If on the other hand, there is a direct arrival accompanied by multipath components, the modified FBLP method should be able to estimate the AOA of that direct arrival.

2.4.3.5.2 MVDR Spectrum Estimation

The MVDR method was developed by Capon and is therefore known as Capon's method. It is also called the Maximum Likelihood Method, not to be confused with Maximum Likelihood Estimation. Like AR modelling, MVDR is based upon the method of least squares. It is different in that it imposes a constraint on the least squares solution. This is best explained by comparison to conventional beamformers [14]. For conventional beamformers, there is interference from signals which are not in the current scan direction. In the case of an MVDR estimator, the average beamformer output power is minimized under the constraint that the beamformer gain in the scan direction is always unity. In other words, the beamformer variance is minimized while the response is not distorted. The effect is to prevent off boresight interference, resulting in improved resolution.

Although the resolution capability of MVDR is better than that of Fourier based methods it is not as good as that of AR modelling [17]. It does, however, determine the relative powers of the signals.

2.4.3.5.3 The Minimum-Norm and MUSIC Algorithms

Minimum-Norm and MUSIC are two superresolution algorithms based on eigenanalysis of the correlation matrix of the data collected by an N element antenna array. Eigendecomposition of the correlation matrix will result in N eigenvalues and their corresponding eigenvectors. If K signals are impinging on the array, the eigenvalues and eigenvectors can be divided into two sets:

1. Those eigenvectors which correspond to the $N-K$ smallest eigenvalues span a space named the noise subspace.
2. Those eigenvectors which correspond to the K largest eigenvalues span a space named the signal subspace.

Therefore, the total space spanned by the eigenvectors of the correlation matrix is partitioned into two subspaces. The signal and noise subspaces are orthogonal complements of each other. Hence the inner product of any signal eigenvector with any noise eigenvector will be null.

The difference between the Minimum-Norm method and the MUSIC method is the processing which occurs after the partition of the eigenspace. In the Minimum-Norm algorithm, a vector **b** of length N is found such that: **b** is orthogonal to each signal eigenvector and hence lies in the noise subspace, the first element of **b** is unity, and its Euclidean norm is a minimum. The vector **b** may be computed from either the noise or signal eigenvectors. Let a steering vector $\mathbf{a}(\theta)$, also of length N, be defined as the complex gain of the antenna array. The Minimum-Norm spectrum is then calculated as,

$$Minimum-Norm\ \ Spectrum(\theta)=\frac{1}{\left|\mathbf{a}^{H}(\theta)\mathbf{b}\right|^{2}} \tag{2-4}$$

where H indicates Hermitian transpose.

The MUSIC algorithm directly exploits the property that the eigenvectors spanning the signal subspace are orthogonal to those spanning the noise subspace. The MUSIC spectrum is created by projecting all possible signal directions, defined by the steering vector $\mathbf{a}(\theta)$, onto the noise subspace, defined by the matrix $\mathbf{V}_N$. $\mathbf{V}_N$ consists of all the noise eigenvectors. For those values of θ which are true signal arrival angles, $\mathbf{a}(\theta)$ will be orthogonal to the noise subspace and will result in a null in the spectrum (or peak in the inverse of the spectrum). Although only one noise eigenvector is theoretically required to make up $\mathbf{V}_N$, all are used in order to suppress spurious peaks. The MUSIC spectrum which yields a peak at the AOA of each signal is calculated as,

$$MUSIC\ \ Spectrum(\theta)=\frac{1}{\mathbf{a}^{H}(\theta)\mathbf{V}_N\mathbf{V}_N^{H}\mathbf{a}(\theta)}. \tag{2-5}$$

For successful operation, both the Minimum-Norm and MUSIC algorithms require knowledge of the number of signals. This may be done by analysis of the eigenvalue magnitudes based on knowledge of the noise power, or by using one of the information-theoretic criteria available [18]. Since the eigenspace must be partitioned into two subspaces, there must be at least one eigenvector spanning the noise subspace. As a result, the antenna array must have at least one more element than the number of signals impinging on the array.

Since the inception of MUSIC, numerous modifications and variations have been developed in order to improve the characteristics and performance of the original algorithm. Perhaps most notable is Root MUSIC [18]. The original MUSIC algorithm must scan through the entire angular interval of 0° to 360° in order to find peaks which correspond to signals. In contrast, Root MUSIC determines the AOA of each signal by finding the signal zeros of a polynomial formed from the noise subspace. The signal zeros are identified as those which lie near the unit circle. Besides having a more direct approach, simulation studies [20] have determined that because of the radial nature of the errors in the estimated signal zeros, Root MUSIC is also superior to original MUSIC in terms of accuracy. Root MUSIC, however, is only suitable for linear, equi-spaced arrays.

Three other variations of MUSIC are worthy of mention. Cyclic MUSIC algorithms eliminate some of the drawbacks of MUSIC when the signals of interest exhibit cyclostationarity [21, 22]. In addition to the spatial coherence properties of the signals, their spectral coherence properties are also used to determine AOA. Oh and Un [23] presented an improved MUSIC algorithm which removed the spatial correlation among sources in the spectral estimator. The result is an improvement in the ability to resolve closely spaced signals. Finally, Brandwood [24] shows that the projection of the array steering vector into the noise subspace can be found without eigenanalysis of the data correlation matrix. His method demonstrates significant savings in computational loads.

2.4.3.5.4 Comparison of Superresolution Algorithms

Haykin [18] gives a detailed discussion regarding the relationships between the four superresolution algorithms which have been discussed. As Haykin points out, these interrelationships are only valid for a wide sense stationary process whose correlation matrix is exactly known. In brief:

1. The MVDR spectrum is a harmonic averaged or smoothed version of the AR spectrum. Hence it contains smaller fluctuations.
2. The MUSIC spectrum corresponds to an MVDR spectrum for a correlation matrix of infinite signal to noise ratio. This means that the resolution of MUSIC is better than that of MVDR.
3. The Minimum-Norm spectrum corresponds to the AR spectrum for infinite signal to noise ratio. Consequently, the resolution of Minimum-Norm is superior to that of AR modelling.
4. The MUSIC spectrum is a weighted harmonic average of the Minimum-Norm spectrum meaning that its spectrum is smoother.

Studies on the performance of MUSIC compared to Minimum-Norm [25] indicate that the resolution threshold of Minimum-Norm is at a lower SNR than that of MUSIC. However, above this threshold the variance in the MUSIC spectrum is smaller than that in the Minimum-Norm spectrum. The superiority of MUSIC over other methods, in regards to detecting the AOA of planar wavefronts, is well documented in the literature. R.O. Schmidt, the inventor of MUSIC, compares the ability of MUSIC to detect the AOA of two signals to that of conventional beamformers, Maximum Likelihood, and Maximum Entropy or AR modelling [26]. MUSIC was found to outperform them all in resolution and accuracy. Haykin [18] compares MUSIC to the Fourier based periodogram method. He displays several cases where the Fourier method has trouble resolving two signals separated by less than two beamwidths, while MUSIC easily and accurately resolves them.

The superiority of MUSIC over the Fourier transform was confirmed by the author. To resolve two 840 MHz signals separated by 10°, with an error less than 0.2°, the Fourier transform required data spread over 7.8 λ. To achieve the same results, MUSIC required data spread over only 0.3 λ.

A further advantage of MUSIC is that it may be used with an array of arbitrary geometry. MUSIC only requires knowledge of the array geometry, which is provided by the steering vector. This is also true of the MVDR method but not AR modelling or Minimum-Norm which require a linear uniform array.

A major drawback of MUSIC and many of the eigendecomposition methods is performance degradation when the signals to be detected are highly or fully correlated (coherent). In the multipath environment signals are in general highly correlated. As a result this issue is of concern when applying MUSIC to the problem of direction finding in the mobile cellular radio environment. Techniques to overcome this are presented in the next chapter.

CHAPTER 3
THE MUSIC ALGORITHM

3.1 Introduction

The previous chapter introduced MUSIC as a superresolution spectral estimation algorithm that may be used to estimate the AOA of radio waves. In this chapter, the theory and application of MUSIC to the mobile cellular telephone location problem is presented. Because MUSIC is able to process both spatial and temporal data, the chapter is divided into two major sections. The focus of the first section is the application of MUSIC in the spatial domain to directly estimate AOA. The second section shows how temporal processing, normally a frequency estimation exercise, may be used to determine AOA when motion is present.

3.2 Spatial MUSIC

3.2.1 Antenna Array

The spatial form of MUSIC estimates the AOA of planar wavefronts impinging on an antenna array. As will be seen later, the number of elements in the array must be at least one greater than the number of signals for which AOA is to be estimated. The requirement of a multi-element array is the first obstacle to be overcome when applying MUSIC to the problem of estimating AOA at the mobile. A multi-element antenna array mounted on an automobile is impractical, costly, and unattractive.

An alternative to a physical antenna array is to use a fewer number of physical antenna elements and exploit the motion of the mobile. A "virtual" antenna array is produced by sampling, in time, one or more mobile mounted antennas as the mobile moves. Figure 3-1 illustrates an automobile with two antennas mounted

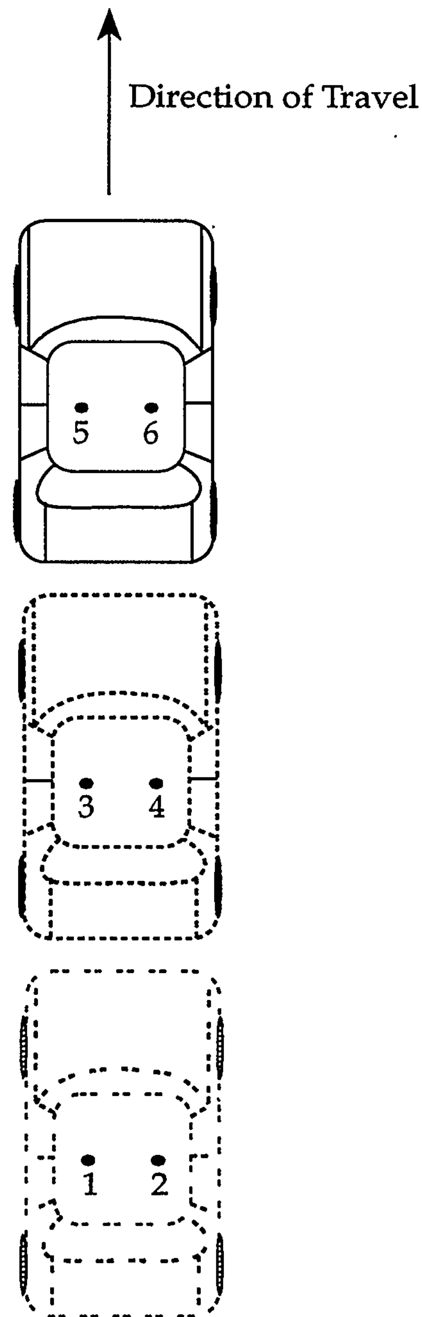

Figure 3-1 Creation of Virtual Antenna Array

on the roof. When the automobile is in motion, sampling the signals received by the antennas at three different times will result in three spatially and temporally separated samples for each antenna. If the sampling rate is sufficiently high, the signal environment will not change significantly from the time of the first sample to the time of the third. Hence, the 6 spatially separated signals simulate a 6 element array.

The geometry of the array is dependent on the spacing of the physical antennas, the sampling frequency, and the velocity of the mobile. The separation of the physical antennas is fixed once the antennas are installed. The separation between virtual elements can be made constant if the sampling frequency is adjusted when the mobile velocity changes or, will vary with velocity if the sampling frequency is fixed. In either case, the separation of virtual elements for the simulations of chapters 4 and 5 was small enough ($\leq$ 30 cm) that a rectangular geometry as in Figure 3-1 can be assumed even if the mobile is not following a straight path.

MUSIC does not restrict the geometry of the array but merely requires knowledge of it. This knowledge is contained in the array gain or steering vector. Consider the virtual array for two physical antennas as shown in Figure 3-2. The direction of motion is in direction Y and the AOA is measured as shown.

The steering vector is defined as

$$\mathbf{a}(\theta) = \left[a_1(\theta)\ a_2(\theta) \cdots a_6(\theta)\right]^T \tag{3-1}$$

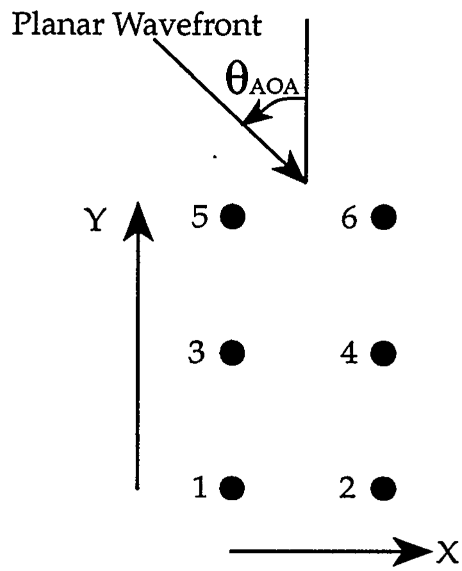

Figure 3-2 Two Dimensional Antenna Array

where $a_n(\theta)$ is the complex gain of the n^{th} element for the direction θ. The angle of $a_n(\theta)$ is the phase at the n^{th} element with respect to the phase at element 1. The phase at element 1 is taken to be zero. For all simulations conducted, the gain amplitude for each element is unity.

Due to the cosine and sine functions used in calculating the angle of $a_n(\theta)$, an ambiguity exists if the array is one dimensional. If a one dimensional array in the Y direction is used, the angular spectrum from 0° through 180° will be repeated from 360° back to 180°. In the case of a one dimensional array in the X direction, the spectrum from 90° to 270° will be a mirror image of the spectrum from 90° back through 0° to 270°. If a two dimensional array is used, the ambiguity is eliminated and the angular spectrum is unique throughout the entire 360°. Therefore, at least two physical antennas, mounted perpendicular to the direction of motion, are required.

3.2.2 Formation of the Correlation Matrix

The ensemble averaged correlation matrix **R** is defined as the expected value of the product of the data matrix **D** with its Hermite [18]. In equation form,

$$\mathbf{R} = E\left[\mathbf{D}^{H}\mathbf{D}\right] \tag{3-2}$$

where the superscript H indicates Hermitian transpose and E the expectation operator. The correlation matrix itself is Hermitian. Although the MUSIC algorithm is based upon the ensemble averaged correlation matrix, in practice it is not known. It is necessary therefore to estimate it with a sample average. Towards this end, two principal methods of organizing the data matrix **D** are considered: temporal smoothing and spatial smoothing. It will later be shown that these two smoothing methods have a significant impact on the performance of eigendecomposition algorithms such as MUSIC [18].

3.2.2.1 Temporal Smoothing

Let an antenna array (not virtual) of arbitrary geometry be composed of N physical antennas or elements. If the signal received at each element is sampled in time, the data matrix can be organized as a sequence of "snapshots". Each snapshot consists of the data received by the array at a specific instant in time. In this way the data is averaged over time and is therefore temporally smoothed. If S samples are collected for each element, the data matrix **D** is defined as,

$$\mathbf{D} = \begin{bmatrix} d_{1,1} & \cdots & d_{1,N} \\ \vdots & \ddots & \vdots \\ d_{S,1} & \cdots & d_{S,N} \end{bmatrix}. \tag{3-3}$$

where N = the number of elements in the antenna array

S = the number of samples per element

$d_{j,k}$ = the j^{th} sample of the k^{th} element.

The estimate of the correlation matrix, $\hat{\mathbf{R}}$, calculated from such a data matrix will be an average over S snapshots of data. Hence,

$$\hat{\mathbf{R}} = \frac{1}{S}\mathbf{D}^H\mathbf{D}. \tag{3-4}$$

In the case of a virtual antenna array, strict temporal smoothing is not possible. Since virtual elements are separated in time as well as space, a snapshot of data as previously defined does not exist. However, again assuming that the signal environment does not change appreciably with time, adjacent samples can be grouped such that each group corresponds to a particular element of the virtual array. In this way, data samples may be grouped into snapshots even though the samples comprising a snapshot of data do not occur at the same instant in time.

For example, in Figure 3-3 an antenna is moving in the direction shown. If the array is to consist of 3 elements with 2 samples each, samples 1 and 2 could correspond to the first element, samples 4 and 5 to the second, and 7 and 8 to the third. One snapshot of data would then consist of samples 1, 4, and 7, and the second snapshot of samples 2, 5, and 8.

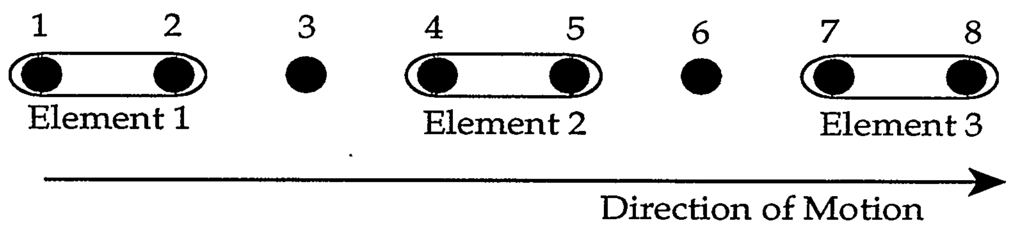

Figure 3-3 Samples in Virtual Antenna Array

The scheme illustrated in Figure 3-3 is in fact temporal smoothing for a virtual array in motion. Sampling an array while it is in motion is an effective way of decorrelating coherent signals impinging on the array [27]. This decorrelation effect will later be shown to be important for the success of MUSIC.

3.2.2.2 Spatial Smoothing

Consider one snapshot of data for an N element array. A subarray is formed consisting of M elements, where $M < N$. Spatial smoothing is then introduced by sliding the subarray across the full array in both the forward and backward directions.

Figure 3-4 illustrates the virtual array created by 2 physical antennas in motion. The antennas are sampled at 4 locations resulting in a 2 by 4, or 8 element array. Also shown are 3 subarrays each consisting of 4 elements. The data matrix is a collection of the 3 data subarrays. The data is thereby spatially averaged. Wherever spatial smoothing is used in this thesis, one dimension of the subarray is always equal to the number of physical antennas. The second dimension is variable. If in Figure 3-4 there were 3 antennas instead of 2, the subarrays could only be of dimension 3 by 2 or 3 by 3.

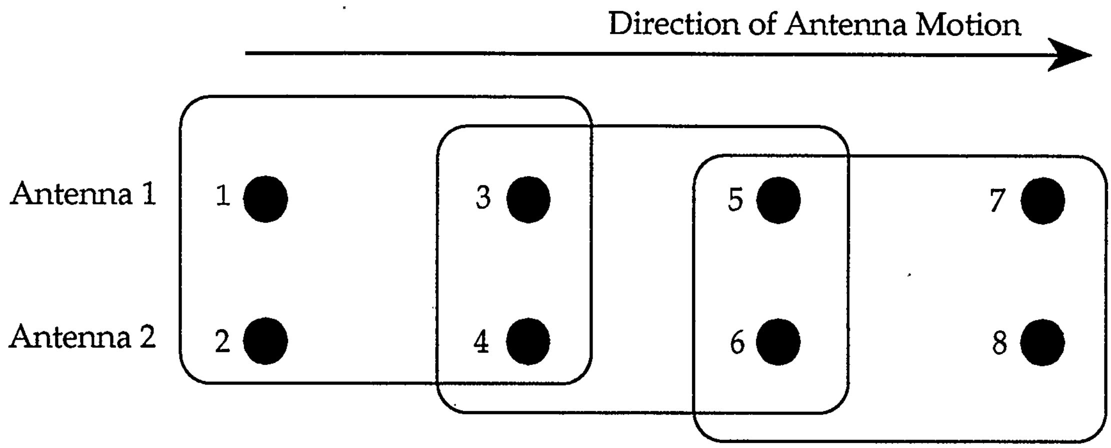

Figure 3-4 A 2 Antenna by 4 Virtual Element Array with 3 Subarrays

The subarray may slide across the full array in the forward direction, backward direction, or both. For the array pictured in Figure 3-4 with both forward and backward spatial smoothing, the data matrix **D** will be,

$$\mathbf{D} = \left[\begin{array}{cccc} d_4 & d_3 & d_2 & d_1 \\ d_6 & d_5 & d_4 & d_3 \\ d_8 & d_7 & d_6 & d_5 \\ \hline d_1^* & d_2^* & d_3^* & d_4^* \\ d_3^* & d_4^* & d_5^* & d_6^* \\ d_5^* & d_6^* & d_7^* & d_8^* \end{array}\right] \tag{3-5}$$

where d_n is the complex data sample of the n^{th} element and * indicates complex conjugate. The matrix elements above the partition line in **D** correspond to forward smoothing whereas the elements below the partition correspond to backward smoothing.

Consider a virtual array of *P* physical antennas and *V* virtual elements yielding a total of $P \times V$ elements. If the subarrays are of dimension *P* by *Sa*, where $Sa < V$, and only forward spatial smoothing is used, the averaging is performed over $V-(Sa-1)$ subarrays. The correlation matrix estimate is then,

$$\hat{\mathbf{R}} = \frac{1}{V-(Sa-1)}\mathbf{D}^H\mathbf{D}. \tag{3-6}$$

If both forward and backward spatial smoothing are used, the averaging is accomplished over a total of $2(V-(Sa-1))$ subarrays and hence,

$$\hat{\mathbf{R}} = \frac{1}{2(V-(Sa-1))}\mathbf{D}^H\mathbf{D}. \tag{3-7}$$

For data which is spatially and temporally separated, temporal and spatial smoothing are somewhat similar. In this context, temporal smoothing can be viewed as one way spatial smoothing in which all possible subarrays are not used. Hence, for the same number of data points, temporal smoothing will not average the data to the extent that spatial smoothing will.

3.2.3 Eigenanalysis of the Correlation Matrix

There are numerous methods which may be used for the eigendecomposition of matrices. For the purpose of this thesis, eigenanalysis by Jacobi transformations for real symmetric matrices was chosen for its simplicity and availability in C code [28]. The simulation data used is complex and hence the complex Hermitian matrix $\hat{\mathbf{R}}$ must be converted to a real symmetrical augmented matrix. Eigenanalysis on the real augmented matrix results in pairs of eigenvalues and eigenvectors. One eigenvalue and eigenvector are then chosen from each pair. Details concerning this method are available in [28].

3.2.4 MUSIC Theory

When planar wavefronts impinging on an array are uncorrelated, or at most partially correlated, MUSIC is able to accurately resolve the arrivals. However, the performance of MUSIC deteriorates when the signals are highly correlated or perfectly correlated (coherent) [29]. The cellular radio environment is of multipath nature due to the many reflections which occur. Hence, the many signal arrivals at a receiving antenna are merely amplitude-weighted, phase-delayed copies of each other and are thus highly correlated. Therefore, the difficulty encountered by MUSIC when processing correlated signals must be resolved for the applications considered.

Various schemes have been devised to overcome the correlated signal problem [29]. Of these, spatial smoothing, temporal smoothing, and an array in motion are of particular interest. For a virtual array, temporal and spatial smoothing are the same in a practical sense. Therefore, a description of how spatial smoothing overcomes the coherent signal problem will suffice. This is followed by a brief discussion of how array motion accomplishes the same thing.

To show how spatial smoothing resolves the coherent signal problem, a description of the underlying theory of MUSIC is in order [29, 30]. For the sake of discussion, consider a linear uniform array consisting of N physical elements (not a virtual array). The K signals impinging on the array are assumed to be narrowband plane waves all centered at frequency ω_o and arriving from directions $\{\theta_1, \theta_2, \cdots, \theta_K\}$ measured with respect to the perpendicular of the array. The signal received at the k^{th} array element is,

$$r_k(t) = \sum_{i=1}^{K} a_i s_i(t) e^{-j\omega_o (k-1)\frac{d}{c}\sin\theta_i} + n_k(t) \tag{3-8}$$

where $s_i(t)$ is the signal of the i^{th} wavefront, a_i is the complex response of the element to the i^{th} wavefront, d is the interelement spacing, c is the propagation speed, and $n_k(t)$ is the noise at the k^{th} element [30].

The noises are assumed to be additive white Gaussian noise with variance σ^2. The signals are stationary, ergodic, complex random processes with zero mean and are uncorrelated with the noises. For convenience, the array elements are assumed to be omnidirectional and hence $a_i = 1$ for all i. Consequently, the array output vector can be written as,

$$\mathbf{r}(t) = \sum_{i=1}^{K} s_i(t)\mathbf{a}(\theta_i) + \mathbf{n}(t) \tag{3-9}$$

where $\mathbf{a}(\theta_i)$ is the steering vector of the array in the direction θ_i and is defined as

$$\mathbf{a}(\theta_i) = \left[1 \; e^{-j\omega_o\tau_i} \; \cdots \; e^{-j\omega_o(N-1)\tau_i}\right]^T, \tag{3-10}$$

$$\tau_i = \frac{d}{c}\sin\theta_i \tag{3-11}$$

and

$$\mathbf{n}(t) = \left[n_1(t) \; n_2(t) \; \cdots \; n_N(t)\right]^T. \tag{3-12}$$

Equation (3-9) can be written in the simpler form,

$$\mathbf{r}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{3-13}$$

where $\mathbf{A}$ is the $N \times K$ matrix

$$\mathbf{A} = \left[\mathbf{a}(\theta_1)\ \mathbf{a}(\theta_2)\ \cdots\ \mathbf{a}(\theta_K)\right] \tag{3-14}$$

and $\mathbf{s}(t)$ is the $K \times 1$ vector,

$$\mathbf{s}(t) = \left[s_1(t)\ s_2(t)\ \cdots\ s_K(t)\right]^T. \tag{3-15}$$

Under the above assumptions, the correlation matrix $\mathbf{R}$ of the array output can be calculated as,

$$\begin{aligned}\mathbf{R} &= E\left[\mathbf{r}(t)\mathbf{r}^H(t)\right] \\ &= E\left[\{\mathbf{A}\mathbf{s}(t) + \mathbf{n}(t)\}\{\mathbf{A}\mathbf{s}(t) + \mathbf{n}(t)\}^H\right] \\ &= \mathbf{A}E\left[\mathbf{s}(t)\mathbf{s}^H(t)\right]\mathbf{A}^H + E\left[\mathbf{n}(t)\mathbf{n}^H(t)\right] \\ &= \mathbf{A}\mathbf{S}\mathbf{A}^H + \sigma^2\mathbf{I}\end{aligned} \tag{3-16}$$

where $\mathbf{S}$ is the signal correlation matrix and $\mathbf{I}$ is the identity matrix.

If the interelement spacing d is less than one half of the signal wavelength, grating lobes will not exist in the array's field intensity pattern [18]. As a result, the columns of $\mathbf{A}$ will all differ from each other. This combined with the Vandermonde structure of $\mathbf{A}$ implies that the columns of $\mathbf{A}$ are linearly independent and the rank of $\mathbf{A}$ is thus K.

When the signals are at most partially correlated, **S** will be nonsingular and also of rank K. Consequently, $\mathbf{ASA}^H$ will be of rank K and have K nonzero eigenvalues. Let $\{\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N\}$ denote the eigenvalues of **R** and $\{\beta_1, \beta_2, \cdots, \beta_N\}$ the corresponding eigenvectors [29]. Then the above rank properties imply that,

$$\lambda_i = \sigma^2, \qquad i \geq K+1, \tag{3-17}$$

$$\beta_i^H \mathbf{a}(\theta_j) = 0, \quad i = K+1, K+2, \cdots, N, \quad j = 1, 2, \cdots, K. \tag{3-18}$$

Those eigenvectors corresponding to the N-K smallest eigenvalues span a subspace called the noise subspace [30]. The subspace spanned by the direction vectors of the signals (columns of **A**) is called the signal subspace and is orthogonal to the noise subspace. MUSIC exploits these facts in the sense that the projection of any steering vector which corresponds to a signal direction onto the noise subspace will result in a null.

3.2.4.1 Decorrelation by Spatial Smoothing

The rank of **S** is dependent on the degree of correlation between the signals. If the signals are uncorrelated, **S** is nonsingular, diagonal, and of rank K. If some of the signals are partially correlated, **S** is no longer diagonal but is still nonsingular and of rank K. When some of the signals are perfectly correlated (coherent), **S** becomes singular and its rank is less than K. In this case, the rank of **R** is also less than K. When this occurs, (3-17) and (3-18) no longer hold and MUSIC breaks down [30]. Therefore, the nonsingularity of **S** must be ensured for the proper operation of MUSIC.

As an example, assume that two of the signals , $s_1(t)$ and $s_2(t)$, are coherent. They are related as,

$$s_2(t) = \alpha s_1(t) \tag{3-19}$$

where α is a complex scalar describing the gain and phase relationship between the two signals. The array output vector can still be written as (3-13). However, $\mathbf{s}(t)$ is now defined as,

$$\mathbf{s}(t) = \left[s_1(t)\ \alpha s_1(t)\ s_3(t)\ \cdots\ s_K(t)\right]^T. \tag{3-20}$$

The correlation matrix of $\mathbf{r}(t)$ is still written as (3-16). However, the rank of $\mathbf{S}$, the signal correlation matrix, is now $K-1$. As a result, the number of eigenvalues of $\mathbf{R}$ with minimum value σ^2 increases from $N-K$ to $N-(K-1)$, meaning that $K-1$ signals are detected instead of K. The problem broadens when $\mathbf{A}$ is considered. Because $\mathbf{A}$ has Vandermonde structure, no linear combination of direction vectors can result in another true signal direction vector. The first column of $\mathbf{A}$ is therefore not a legitimate signal direction vector since it is a linear combination of the direction vectors for signals $s_1(t)$ and $s_2(t)$. The result is an inconsistency. Although $K-1$ signals are detected, only $K-2$ arrival directions are estimated. In general, if m of the K signals are coherent, MUSIC will detect $K-m+1$ signals and resolve $K-m$ arrival angles corresponding to the incoherent signals [30].

Spatial smoothing as described in section 3.2.2.2 will ensure the nonsingularity of the source correlation matrix even when some or all of the signals are coherent.

Again consider a linear uniform array with a total of N elements. The array is divided into overlapping subarrays each consisting of M elements [29]. The first subarray consists of elements $\{1,2,\cdots,M\}$, the second of elements $\{2,3,\cdots,M+1\}$, and so on. There will be a total of L such forward subarrays where

$$L = N - M + 1. \tag{3-21}$$

Let the output of the l^{th} forward subarray be denoted as $\mathbf{r}_l^f(t)$. Following (3-8) to (3-15),

$$\begin{aligned}\mathbf{r}_l^f(t) &= \left[r_l(t)\; r_{l+1}(t)\; \cdots\; r_{l+M-1}(t)\right]^T \\ &= \mathbf{A}_M\mathbf{B}^{l-1}\mathbf{s}(t) + \mathbf{n}_l(t), \;\; 1 \le l \le L\end{aligned} \tag{3-22}$$

where $\mathbf{A}_M$ is the $M \times K$ matrix identical to (3-14) with the last $N-M$ rows removed, $\mathbf{B}^l$ denotes the l^{th} power of the $K \times K$ diagonal matrix

$$\mathbf{B} = diag\left[e^{-j\omega_0\tau_1}, e^{-j\omega_0\tau_2}, \ldots, e^{-j\omega_0\tau_K}\right] \tag{3-23}$$

and

$$\mathbf{n}_l(t) = \left[n_l(t)\; n_{l+1}(t)\; \cdots\; n_{l+M-1}(t)\right]^T. \tag{3-24}$$

Using (3-16), the correlation matrix of the l^{th} subarray is

$$\mathbf{R}_l^f = E\left[\mathbf{r}_l^f(t)\left(\mathbf{r}_l^f(t)\right)^H\right]$$

$$= \mathbf{A}\mathbf{B}^{l-1}\mathbf{S}\left(\mathbf{B}^{l-1}\right)^{H}\mathbf{A}^{H} + \sigma^{2}\mathbf{I}. \tag{3-25}$$

The forward spatially smoothed correlation matrix $\mathbf{R}^{f}$ is defined as the average of all the forward subarray correlation matrices. So,

$$\begin{aligned}\mathbf{R}^{f} &= \frac{1}{L}\sum_{l=1}^{L}\mathbf{R}_{l}^{f} \\ &= \mathbf{A}\left(\frac{1}{L}\sum_{l=1}^{L}\mathbf{B}^{l-1}\mathbf{S}\left(\mathbf{B}^{l-1}\right)^{H}\right)\mathbf{A}^{H} + \sigma^{2}\mathbf{I} \\ &= \mathbf{A}\mathbf{S}^{f}\mathbf{A}^{H} + \sigma^{2}\mathbf{I}\end{aligned} \tag{3-26}$$

where $\mathbf{S}^{f}$ is the forward spatially smoothed signal correlation matrix. If all K signals are coherent, then (3-15) can be written as

$$\mathbf{s}(t) = \left[\alpha_{1}s_{1}(t)\ \alpha_{2}s_{1}(t)\ \cdots\ \alpha_{K}s_{1}(t)\right]^{T} \tag{3-27}$$

where α_{i} relates the i^{th} signal to $s_{1}(t)$. Assuming that $E\left[\left|s_{1}(t)\right|^{2}\right] = 1$, it can be shown that,

$$\begin{aligned}\mathbf{S} &= \boldsymbol{\alpha}\,\boldsymbol{\alpha}^{H} \\ \boldsymbol{\alpha} &= \left[\alpha_{1}\ \alpha_{2}\ \cdots\ \alpha_{K}\right]^{T}.\end{aligned} \tag{3-28}$$

In that case, the forward spatially smoothed signal correlation matrix can be written as,

$$\mathbf{S}^f = \frac{1}{L}\sum_{l=1}^{L}\mathbf{B}^{l-1}\,\boldsymbol{\alpha}\,\boldsymbol{\alpha}^H\left(\mathbf{B}^{l-1}\right)^H$$
$$= \frac{1}{L}\mathbf{C}\mathbf{C}^H \tag{3-29}$$

where

$$\mathbf{C} = \begin{bmatrix} \boldsymbol{\alpha} & \mathbf{B}\boldsymbol{\alpha} & \mathbf{B}^2\boldsymbol{\alpha} & \cdots & \mathbf{B}^{L-1}\boldsymbol{\alpha} \end{bmatrix}$$
$$= \begin{bmatrix} \alpha_1 & & & \mathbf{0} \\ & \alpha_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \alpha_K \end{bmatrix} \begin{bmatrix} 1 & e^{-j\omega_o\tau_1} & e^{-j2\omega_o\tau_1} & \cdots & e^{-j(L-1)\omega_o\tau_1} \\ 1 & e^{-j\omega_o\tau_2} & e^{-j2\omega_o\tau_2} & \cdots & e^{-j(L-1)\omega_o\tau_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-j\omega_o\tau_K} & e^{-j2\omega_o\tau_K} & \cdots & e^{-j(L-1)\omega_o\tau_K} \end{bmatrix}$$
$$= \mathbf{D}\mathbf{T}. \tag{3-30}$$

The rank of $\mathbf{S}^f$ is equal to the rank of **C** which in turn is equal to the rank of **T**, since **D** is of full rank. Matrix **T** is a Vandermonde matrix of dimension $K \times L$ and its rank will be the lesser of K and L. Therefore, **T** will be of rank K if $L \geq K$. In that case $\mathbf{S}^f$, also being of rank K, will be nonsingular and $\mathbf{R}^f$ given in (3-26) will be of the same form as **R** in (3-16). Consequently, MUSIC may be successfully applied even though the signals are all coherent.

The price paid is the total number of sensors N that are required. From (3-18) it is evident that each subarray must have at least $K+1$ elements. Therefore, $M \geq K+1$. The total number of subarrays L, will always be given by

$$L = N - M + 1. \tag{3-31}$$

However, to ensure the nonsingularity of $\mathbf{S}^f$ we require that $L \geq K$. From this it follows that the total number of elements required is at least $2K$. For MUSIC with no spatial smoothing, a total of $K+1$ elements are required.

The number of elements N can be reduced from $2K$ to $3K/2$ by incorporating both forward and backward spatial smoothing. Consider the same array of N elements with L backward subarrays. The first backward subarray of size M consists of elements $\{N, N-1, \cdots, N-M+1\}$, the second consists of $\{N-1, N-2, \cdots, N-M\}$ etc. The complex conjugate of the output vector of the l^{th} backward subarray is

$$\begin{aligned} \mathbf{r}_l^b(t) &= \left[r_{N-l+1}^*(t)\ r_{N-l}^*(t)\ \cdots\ r_{L-l+1}^*(t)\right]^T \\ &= \mathbf{A}_M \mathbf{B}^{l-1}\left(\mathbf{B}^{N-1}\mathbf{s}(t)\right)^* + \tilde{\mathbf{n}}_l^*(t),\ 1 \le l \le L \end{aligned} \tag{3-32}$$

where , $\mathbf{A}_M$ is identical to that of (3-22), and

$$\tilde{\mathbf{n}}_l^*(t) = \left[n_{N-l+1}^*(t)\ n_{N-l}^*(t)\ \cdots\ n_{L-l+1}^*(t)\right]^T, \tag{3-33}$$

and $\mathbf{B}$ is defined in (3-23).

Following a procedure similar to that for forward spatial smoothing, it can be shown that the backward spatially smoothed correlation matrix is

$$\mathbf{R}^b = \mathbf{A}\mathbf{S}^b\mathbf{A}^H + \sigma^2\mathbf{I}. \tag{3-34}$$

If all K signals are coherent, then

$$\mathbf{S}^b = \frac{1}{L}\mathbf{E}\mathbf{E}^H \tag{3-35}$$

where

$$\mathbf{E} = \begin{bmatrix} \boldsymbol{\delta} & \mathbf{B}\boldsymbol{\delta} & \mathbf{B}^2\boldsymbol{\delta} & \cdots & \mathbf{B}^{L-1}\boldsymbol{\delta} \end{bmatrix} = \mathbf{F}\mathbf{T} \tag{3-36}$$

with $\mathbf{T}$ defined in (3-30) and

$$\boldsymbol{\delta} = \begin{bmatrix} \delta_1 & \delta_2 & \cdots & \delta_K \end{bmatrix}^T,$$
$$\delta_k = \alpha_k^* e^{j(N-1)\omega_o\tau_k}, \quad k = 1,2,\cdots,K. \tag{3-37}$$

The matrix $\mathbf{F}$ is the diagonal matrix

$$\mathbf{F} = \begin{bmatrix} \delta_1 & & & \mathbf{0} \\ & \delta_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \delta_K \end{bmatrix} \tag{3-38}$$

and is of full rank. Therefore, if $L \geq K$, $\mathbf{T}$ will be of rank K making $\mathbf{S}^b$ nonsingular and $\mathbf{R}^b$ of full rank. So as in the case of forward spatial smoothing, MUSIC will be able to resolve K coherent signals so long as the number of backward subarrays is at least equal to K. Simulation results demonstrating MUSIC performance for coherent signals, with and without forward/backward spatial smoothing, are given in [29-31].

When both forward and backward smoothing are used, the forward/backward smoothed correlation matrix $\tilde{\mathbf{R}}$ is defined as

$$\tilde{\mathbf{R}} = \frac{\mathbf{R}^f + \mathbf{R}^b}{2}. \tag{3-39}$$

Using the previously given expressions for $\mathbf{R}^f$ and $\mathbf{R}^b$ gives

$$\tilde{\mathbf{R}} = \mathbf{A}\tilde{\mathbf{S}}\mathbf{A}^H + \sigma^2\mathbf{I} \tag{3-40}$$

where

$$\tilde{\mathbf{S}} = \frac{1}{2L}\mathbf{G}\mathbf{G}^H. \tag{3-41}$$

The matrix $\mathbf{G}$ is given by

$$\begin{aligned} \mathbf{G} &= [\ \boldsymbol{\alpha}\ \mathbf{B}\boldsymbol{\alpha}\ \mathbf{B}^2\boldsymbol{\alpha}\ \cdots\ \mathbf{B}^{L-1}\boldsymbol{\alpha} \mid \boldsymbol{\delta}\ \mathbf{B}\boldsymbol{\delta}\ \mathbf{B}^2\boldsymbol{\delta} \cdots \mathbf{B}^{L-1}\boldsymbol{\delta}\ ] \\ &= [\mathbf{DT} | \mathbf{FT}] = \mathbf{D}[\mathbf{T} | \mathbf{HT}] = \mathbf{DG}_o \end{aligned} \tag{3-42}$$

where

$$\mathbf{H} = \begin{bmatrix} \varepsilon_1 & & & \mathbf{0} \\ & \varepsilon_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \varepsilon_K \end{bmatrix}$$

$$\varepsilon_k = \frac{\delta_k}{\alpha_k}, \quad k = 1, 2, \cdots, K. \tag{3-43}$$

If all ε_k in (3-43) are equal, the rank of $\mathbf{G}_o$ will be the lower of K and L. In this case nothing has been achieved by backward smoothing since the condition of $L \geq K$ still holds. However, the likelihood of all ε_k being equal is very small. If at most L of the ε_k are equal and $2L \geq K$, then $\mathbf{G}_o$ will be of rank K and $\tilde{\mathbf{S}}$ will be nonsingular. Using (3-31) in $2L \geq K$ gives

$$2N - 2M + 2 \geq K \tag{3-44}$$

and recalling that $M \geq K + 1$ yields,

$$N \geq \frac{3K}{2}. \tag{3-45}$$

Hence with forward and backward smoothing the minimum number of elements required for complete resolution has been reduced compared to forward or backward smoothing alone.

The smoothed array output correlation matrix for a coherent environment is of the same structure as that for a noncoherent environment allowing the use of MUSIC. Spatial smoothing works equally well for a mixed signal environment consisting of coherent signals and partially correlated signals [29].

3.2.4.2 Decorrelation by Array Motion

Decorrelation of partially correlated or coherent arrivals may also be achieved through motion of the receiving array [27]. In the case of a real array, motion is advantageous to spatial smoothing as a decorrelation scheme because periodic spacing of the array elements is not required. For a virtual array however, this is not significant since the element spacing will be periodic for a constant sampling frequency and mobile velocity.

Consider several signal arrivals all with the same carrier frequency, impinging on an array which is moving. If each of the signals has a unique AOA, then the Doppler shift induced on each signal due to the motion, will also be unique. It can be shown [27] that estimating the correlation matrix **R** by time averaging while the array is in motion, ensures the nonsingularity of the signal correlation matrix **S** and hence the full rank of **R**. If $r_j(t)$ is the output signal of the j^{th} array element and $r_k(t)$ is the same for the k^{th} element, then the time averaged correlation between these two is written as,

$$R_{jk} = \frac{1}{T}\int_0^T r_j(t) r_k^*(t)\, dt \tag{3-46}$$

where T is the duration of time over which the averaging is performed.

To achieve this decorrelation the time averaging must be performed over many differential Doppler cycles for each pair of arrivals. The differential Doppler shift for a pair of arrivals, say the j^{th} and k^{th}, is

$$f_{jk} = \frac{v\left(\cos\theta_j - \cos\theta_k\right)}{\lambda_c} \tag{3-47}$$

where v is the array velocity, λ_c is the carrier wavelength, and θ denotes AOA. The number of differential Doppler cycles is then Tf_{jk}. Hence, there is a minimum distance to be travelled for decorrelation to occur. However, it may not always be desirable to decorrelate closely spaced arrivals. In the case of multipath where the arrivals are expected to be grouped into clusters of closely spaced coherent signals, a single AOA for each cluster rather than the AOA for each individual signal, may be appropriate.

The time averaging of (3-46) is continuous whereas temporal smoothing as given by (3-3) and (3-4) is discrete. Beyond this difference temporal smoothing of a virtual array is identical to time averaging of a real array in motion. Therefore, inherent to the concept of the virtual array is the decorrelation effect produced by motion.

3.3 Temporal MUSIC

Parallel to the spatial problem of estimating AOA is the temporal problem of estimating the frequencies of sinusoids. Whereas spatially spread data collected by an array is required to solve for AOA, estimating the frequencies of sinusoids impinging on a single antenna element requires a time series of data.

The model used for processing the time series data is a transversal FIR filter of length M [18]. The data matrix is constructed using the covariance method which makes no assumptions about the data outside of the data record. Consider the time series of data $\{d_1, d_2, \cdots, d_N\}$ where d_1 occurs before d_2, d_2 occurs before d_3 and so on. Then with forward and backward smoothing

$$\mathbf{D}=\left[\begin{array}{ccc} d_M & \cdots & d_1 \\ d_{M+1} & \cdots & d_2 \\ \vdots & \ddots & \vdots \\ d_N & \cdots & d_{N-M+1} \\ \hline d_1^* & \cdots & d_M^* \\ d_2^* & \cdots & d_{M+1}^* \\ \vdots & \ddots & \vdots \\ d_{N-M+1}^* & \cdots & d_N^* \end{array}\right] \tag{3-48}$$

where the smoothing is temporal for a stationary time series. This temporal smoothing is somewhat different from temporal smoothing for data collected by an array, as described in section 3.2.2.1. Note however that the structure of (3-48) is identical to that of (3-5) which defines the data matrix for a virtual array with spatial smoothing. Indeed if the virtual array of (3-5) is one dimensional in the direction of motion and the time series of (3-48) is, in addition to temporally distributed, spatially distributed, the two data matrices are identical. Hence, for spatially and temporally separated data, temporal smoothing based on the FIR filter model is identical to spatial smoothing. Temporal smoothing as described here will therefore differ from temporal smoothing of array collected data in the same way as spatial smoothing does (see section 3.2.2.2).

As with spatial MUSIC, the correlation matrix is estimated from the smoothed data matrix. For a data series of length N, filter length M, and smoothing in both directions,

$$\hat{\mathbf{R}}=\frac{1}{2(N-(M-1))}\mathbf{D}^H\mathbf{D}. \tag{3-49}$$

Eigenanalysis of $\hat{\mathbf{R}}$ yields M eigenvectors which are divided into the noise and signal subspaces. The matrix $\mathbf{V}_N$ is comprised of the noise space eigenvectors. In contrast to the steering vector of (3-1), the temporal form of MUSIC utilizes a frequency scanning vector defined as

$$\mathbf{f}(\omega) = \left[1 \; e^{-j\omega} \; \cdots \; e^{-j\omega(M-1)}\right]^T,$$
$$-\pi \leq \omega \leq \pi \qquad (3\text{-}50)$$

where ω is angular frequency normalized by the sampling frequency. For those values of ω which correspond to the normalized angular frequencies of the sinusoids contained in the data, $\mathbf{f}(\omega)$ will be orthogonal to the noise subspace spanned by the eigenvectors contained in $\mathbf{V}_N$. Hence, the angular frequencies of the sinusoids will correspond to peaks in the MUSIC spectrum defined in (2-5) with $\mathbf{a}(\theta)$ replaced by $\mathbf{f}(\omega)$.

The AOA and frequency estimation problems become related when motion is involved. The Doppler spread in frequency due to motion, precipitates a relationship between the AOA of an incoming sinusoid and its frequency. This relationship may be exploited such that MUSIC applied in the temporal domain may be used to estimate AOA.

For the purpose of locating a mobile, consider a single tone transmitted by a base station which due to multipath arrives at the mobile at various arrival angles. If the mobile is stationary, the multiple arrivals will all have the same frequency. In that case, the temporal form of MUSIC will provide no direction information. If however, the mobile is in motion, the frequency of each arrival will shift due to Doppler. The Doppler shift experienced by each arrival is dependent on its AOA. As illustrated in Figure 3-5 a signal arriving at 0° with respect to the mobile will experience the maximum Doppler shift of V/λ, where V is the mobile's velocity and λ is the signal wavelength. A signal arriving from 90° will experience no Doppler shift.

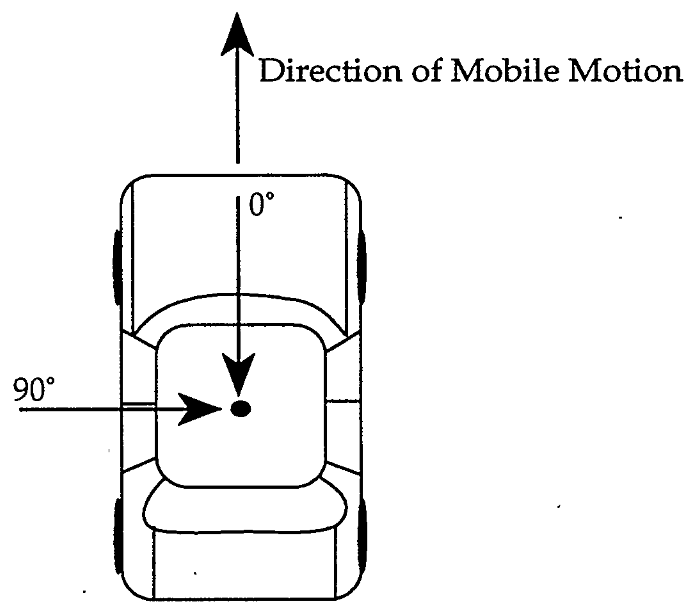

Figure 3-5 Signals Affected by Doppler Shift

Due to the fact that the received antenna signal is sampled, any signal processing will be concerned with the signal's lowpass equivalent rather than the bandpass spectrum. Therefore, applying MUSIC or any other spectral estimation method (i.e. Fourier) will yield the Doppler shift of each sinusoid instead of the absolute frequency.

Let ω_{MUSIC} be the normalized Doppler shift corresponding to a peak in the

MUSIC frequency spectrum. The AOA, θ_{AOA}, of the sinusoid corresponding to that Doppler shift can be determined from the relation

$$2\pi \frac{V}{\lambda} \cos(\theta_{AOA}) = f_s \omega_{MUSIC} \tag{3-51}$$

where f_s is the sampling frequency, and θ_{AOA} is measured as in Figure 3-5.

The cosine function in (3-51) again results in an ambiguity. An arrival from say 20° will experience the same Doppler shift as a signal from 340°. As in the case of spatial MUSIC, two physical antennas will eliminate the ambiguity provided the axis connecting the two antennas is orthogonal to the direction of motion.

Figure 3-6 shows two moving antennas and a planar wavefront arriving from θ. If the antennas are separated by d wavelengths, then the phase difference, ϕ_{12} in radians, between the signals at the two antennas is related to θ by

$$\cos(\theta) = \frac{\phi_{12}}{2\pi d}. \tag{3-52}$$

In this case there is also an ambiguity since the phase difference will be the same for θ and $-\theta$. Note however that the angular reference in Figure 3-6 is 90° from the reference in Figure 3-5. Therefore only one of the two possible solutions for θ using (3-52) will overlap with one of the solutions using (3-51). The ambiguity is thus eliminated.

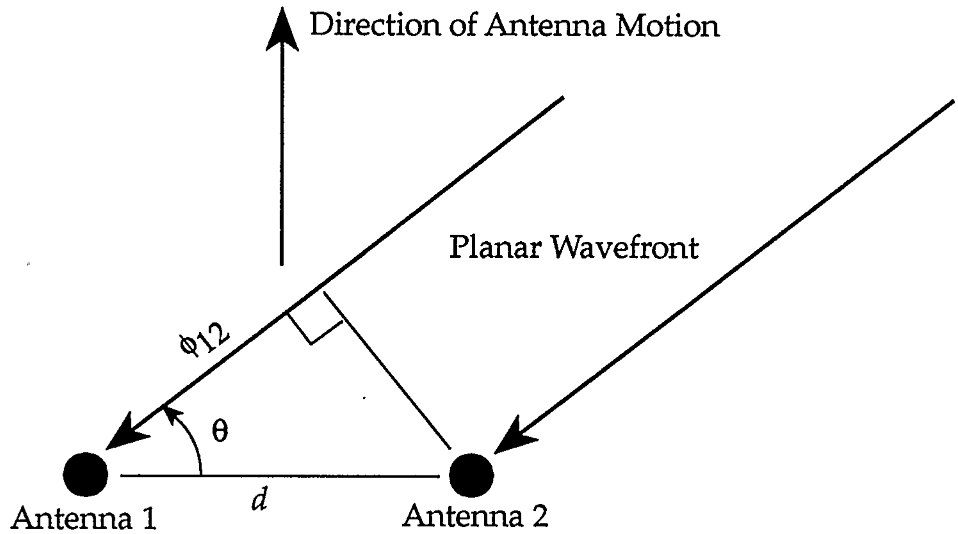

Figure 3-6 Planar Wavefront Approaching Two Antennas

The signal amplitude and phase at each antenna in Figure 3-6 can be determined if the signal frequency is known. Consider K sinusoids impinging on one of the two moving antennas of Figure 3-6. Although the sinusoids all originate from a single transmitted tone, their Doppler shifts will all be unique given that their arrival angles are unique. For N data samples $\{d_1, d_2, \cdots, d_N\}$ collected at a moving antenna, the following set of equations hold;

$$\begin{bmatrix} e^{j\omega_1} & e^{j\omega_2} & \cdots & e^{j\omega_K} \\ e^{j2\omega_1} & e^{j2\omega_2} & \cdots & e^{j2\omega_K} \\ \vdots & \vdots & \ddots & \vdots \\ e^{jN\omega_1} & e^{jN\omega_2} & \cdots & e^{jN\omega_K} \end{bmatrix} \begin{bmatrix} a_1 e^{j\phi_1} \\ a_2 e^{j\phi_2} \\ \vdots \\ a_K e^{j\phi_K} \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix} \tag{3-53}$$

where ω_i is the Doppler frequency, a_i the amplitude, and ϕ_i the phase of the i^{th} sinusoid.

In summary, AOA may be found using Temporal MUSIC and two moving antennas with the following procedure:

1) Calculate the Doppler frequency for each arrival using temporal MUSIC and the data time series for one antenna.

2) Use (3-51) to calculate the pair of possible arrival angles θ and $-\theta$, for each arrival.

3) Repeat 1) for the second antenna.

4) Use the Doppler frequencies found in 1) and 3) in (3-53) to find the amplitude and phase of each arrival at each antenna.

5) Calculate the difference in phase, between the two antennas for each arrival.

6) Use the phase differences in (3-52) to calculate another pair of arrival angles for each arrival.

7) For each arrival find the common angle between the pair found in 2) and the pair found in 6).

CHAPTER 4
TECHNIQUE #1 - PETRA: RESULTS & ANALYSIS

4.1 Introduction

In a multipath environment, it is the AOA of the LOS signal which allows one to triangulate and thereby locate. Therefore, in addition to determining the AOA of the numerous arrivals impinging on the antenna array, one must also determine which of the arrivals is LOS. Towards this end, PETRA (PEak TRacking Algorithm) was developed.

This chapter begins with a description of the three cluster signal environment. This is followed by a more detailed description of PETRA, as well the computer simulation results and an analysis of these results.

4.2 Signal Environment

The IS-54 cellular standard [13] models the mobile cellular radio channel as two independently fading, equal power, Rayleigh channels. The assumed time delay between the two channels is between 0 and 1 symbol intervals, where one symbol interval is 41.6 μs. These two channels can be viewed as two clusters of signals separated by time of arrival.

The IS-54 standard makes no assumptions regarding the direction of arrival of the two clusters or the paths travelled. It is reasonable to assume however that the two clusters are distinguished not only by time of arrival but also by AOA. In this thesis we assume that the first cluster to arrive is composed of signals which approximately follow a direct or LOS path. The signals in this cluster would therefore arrive at nearly the same time and from approximately the same direction. In the same way we assume that the second cluster consists of signals which arrive at a later time and follow a reflected or more indirect path.

It follows that two reasonable assumptions may be made. Firstly, those signals following the reflected path will travel farther and lose power at the reflection. It is therefore probable that the second cluster will be of lesser power than the first. Secondly, signals are more likely to survive a direct path than a reflected path. Hence the LOS cluster will contain more arrivals than the reflected cluster.

The signal environment assumed, for the channel between any one base station and a mobile, is as illustrated in Figure 4-1. Three clusters are assumed instead of two since in an urban environment there are many potential reflectors. The first cluster of signals to arrive corresponds to the LOS path. It contains more arrivals

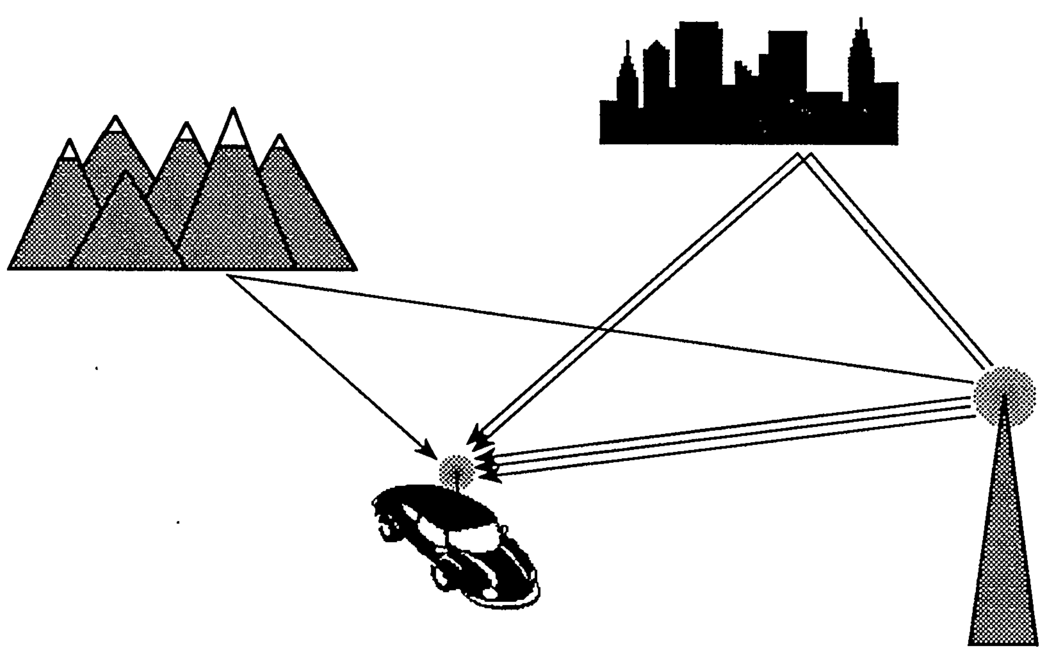

Figure 4-1 Signal Cluster Environment

and has greater power than the other clusters. The second and third clusters to arrive contain progressively fewer arrivals and are of lesser power.

Table 4-1 summarizes the characteristics of the three clusters used in the computer simulations. The mean power is the average power of the individual arrivals which make up each cluster. The total power is the sum of the arrival powers in each cluster. Time spread refers to the time separation between the first and last arrival in a cluster. The time delay is the time span between the beginning of the impulse response and the first arrival in the cluster. Note that the time delay of the 3rd cluster is much shorter than the full symbol interval of 41.6 μs. Because the definable impulse response of the simulated data is only 7.1 μs in length, a time separation between clusters of one symbol interval is not possible.

Table 4-1 Cluster Characteristics

Cluster	No. of Arrivals	Mean Power	Total Power	Time Spread	Time Delay
LOS	13	-15.9 dBW	-4.8 dBW	1.9 μs	0
2	6	-18.2 dBW	-10.4 dBW	2.1 μs	1.9 μs
3	4	-20.5 dBW	-14.5 dBW	0.9 μs	4.4 μs

All simulated data used in this chapter as well as chapter 5, was generated by the modified SURP (Simulation of the Urban Radio Propagation Channel) software package [32]. Hashemi developed SURP using actual measurements in the mobile cellular band, in order to simulate an urban radio propagation channel.

The model used is characterized by the path amplitudes, arrival times, phases, and angles of arrival. Fattouche et al. [32], modified the package so that it simulates short term fading in addition to long term fading.

4.3 PETRA and Spatial MUSIC with Temporal Smoothing

4.3.1 Introduction

PETRA was developed to determine the LOS cluster in a three cluster environment by tracking the dominant peak in the MUSIC spectrum as the number of allocated signal eigenvalues varies from one to three. Recall that the eigenvalues obtained from the eigenanalysis of the data correlation matrix are divided into those attributable to noise and those attributable to signals. The corresponding eigenvectors then span the noise and signal spaces respectively. The user of MUSIC has the freedom to allocate any number of the eigenvalues to signals with the remaining allocated to noise. Allocating only the largest eigenvalue to the signal side, is equivalent to limiting MUSIC to finding only one signal. If the two largest eigenvalues are attributed to signals, then MUSIC is able to resolve two signals, and so on.

PETRA employs spatial MUSIC with temporal smoothing. This form of MUSIC was chosen for no other reason than it was the first studied, and implemented in code, by the author. Temporal smoothing based on array data is used as opposed to temporal smoothing based on the FIR filter model.

4.3.2 Basic Assumptions

PETRA is based on the assumption that when the signal space is defined by only one eigenvector, the dominant peak in the MUSIC spectrum will correspond to that signal (or cluster) with the highest power. If the LOS cluster is of the highest power, then the dominant MUSIC peak should estimate its AOA.

This assumption is based on the results of two sets of simple tests. In the first set, a two ray model was simulated. One signal arrived from 60° and the other from 200°. If two signal eigenvalues were allocated, the peak at 60° in the resulting MUSIC spectrum was of higher amplitude than the peak at 200°. This was true regardless of which signal arrived first or which was of higher power. If however, only one signal eigenvalue was allocated, the dominant peak in the MUSIC spectrum corresponded to the signal of higher power. The results of this test are shown in Table 4-2.

Table 4-2 Two Ray Model Results

1st Arrival	2nd Arrival	1st Arrival Amplitude	2nd Arrival Amplitude	Dominant Peak
200°	60°	**2.5**	1.0	192°
200°	**60°**	1.0	**2.5**	54.5°
60°	**200°**	1.0	**2.5**	192°
60°	200°	**2.5**	1.0	54.5°

The second set of tests used a three ray model. With three rays there are six possible combinations of signal power order. In one particular test, for one allocated signal eigenvalue, the dominant MUSIC peak corresponded to the cluster of highest power in four out of the six combinations. In another test, the association occurred in three out of the six combinations. Therefore, in a more complex signal environment the assumption does not prove to be consistently valid. However, a majority precedence does exist.

4.3.3 The PETRA Process

The process of tracking the dominant peak for 1, 2, and 3 signal eigenvalues is as follows:

1) Find the eigenvalues and corresponding eigenvectors of the data correlation matrix;
2) Define the signal subspace as that space spanned by the eigenvector corresponding to the eigenvalue of highest magnitude; the remaining eigenvectors span the noise subspace;
3) Generate the MUSIC spectrum for the partition specified in 2) and note the AOA of the dominant peak;
4) Repartition the eigenvalues and eigenvectors found in 1) such that the two highest magnitude eigenvalues are attributed to signals and the remaining to noise;
5) Generate another MUSIC spectrum based on the partition in 4) and note the AOA of the peak closest to the AOA found in 3);
6) Repeat step 4) with three signal eigenvalues;
7) Generate another MUSIC spectrum based on the partition in 6);
8) The AOA of the peak closest to the AOA determined in 5) will be the estimate of the AOA of the LOS cluster.

4.3.4 Testing the PETRA Assumption with Clusters

Three clusters of varying widths were used as a first step in determining how well PETRA performs in a cluster environment. The AOA of the LOS, 2nd and 3rd clusters were 240°, 300°, and 90° respectively. The cluster characteristics were as shown in Table 4-1. A 2×2 array, with 50 data samples per element and a spacing of 200 data samples between virtual elements was used. The 2 physical antennas were separated by one half of a wavelength. The speed of the mobile was 50 km/hr. A more detailed explanation for the choice of some of these parameters follows in the next section.

The angular width of the clusters was varied from 2° to 80° (all 3 clusters always having the same width) and PETRA was run on 10 consecutive sets of data for each cluster width. Thus there are 10 trials for each cluster width. For each trial,

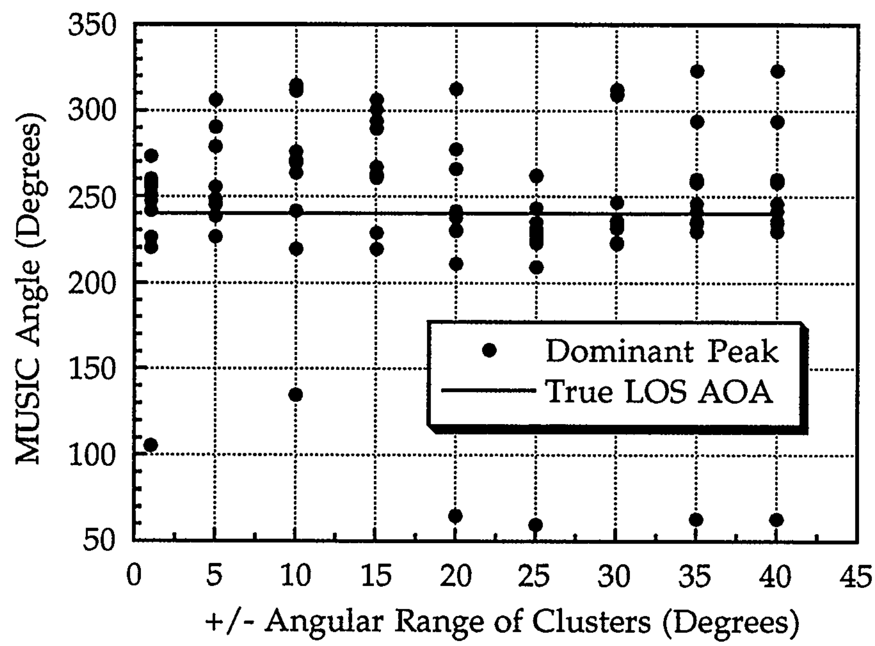

Figure 4-2 Dominant Peak Angle for 1 Signal Eigenvalue

the dominant MUSIC peak for 1 signal eigenvalue and the corresponding peaks for 2 and 3 signal eigenvalues were recorded. The results for the dominant peak for 1 signal eigenvalue are shown in Figure 4-2.

The bold horizontal line in Figure 4-2 corresponds to the true AOA of the LOS cluster at 240° whereas the dots correspond to the dominant MUSIC peaks for the 10 trials at each cluster width. Clearly illustrated is the fact that the majority of dominant peaks correspond to the LOS cluster at 240°. Dots in the vicinity of 300° and 90° demonstrate that occasionally the other clusters will dominate.

Figure 4-3 shows the results for 2 and 3 signal eigenvalues. The average of the 10 trials at each cluster width has been shown instead of the individual results.

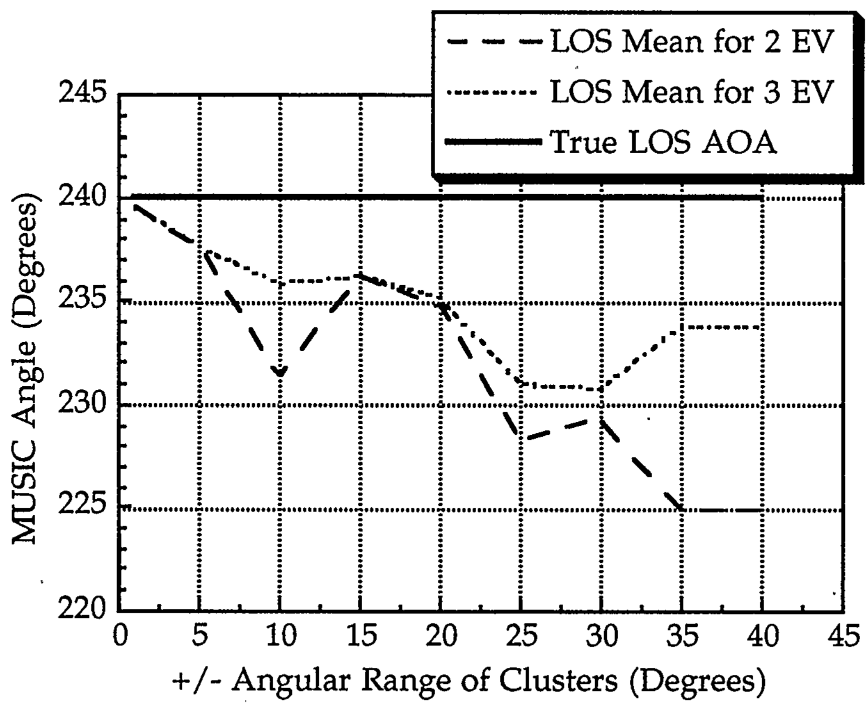

Figure 4-3 Mean Comparisons for 2 and 3 Signal Eigenvalues

Recall that for 1 signal eigenvalue, the AOA of the dominant peak is of interest. For 2 and 3 signal eigenvalues however, the peak closest in AOA to the estimate made with 1 and 2 signal eigenvalues respectively, is the LOS estimate. Figure 4-3 clearly demonstrates two points. First, the LOS AOA estimate becomes more accurate as the number of signal eigenvalues increases; second, as the width of the clusters increases, the LOS estimates tend to drift away from the true value of 240°.

The results in Figure 4-2 and 4-3 demonstrate that PETRA must be run over several data sets in order to be reliable and accurate. Since the LOS cluster is not always properly identified, enough PETRA estimates must be generated to obtain a clear LOS majority. It was also determined that the accuracy of the individual PETRA estimates is very poor. If, however, many LOS PETRA estimates are averaged, the resulting error is much smaller.

4.3.5 Simulation Parameters for Statistical Tests

To obtain a more statistical appreciation for the performance of PETRA, simulations were conducted on a number of different cluster sets. Table 4-3 summarizes the parameters used in these computer simulations.

The number of physical antennas was chosen to be two since this is the minimum number required and hence the most practical and cost efficient. The vehicle speed was chosen to be that typical of urban speed limits. The frequency of transmission is located within the cellular band.

Table 4-3 Simulation Parameters for PETRA

Number of physical antennas	2
Array size	2×2
Physical antenna spacing	0.5 λ
Virtual antenna spacing	0.5 λ
Number of data samples per element	300
Vehicle speed	50 km/hr
Frequency	840 MHz

The other parameters in Table 4-3 were chosen according to tests designed to evaluate their effect on the performance of spatial MUSIC with temporal smoothing. The signal environment for the tests was as given in Table 4-1. The optimal spacing between the physical antennas, as well as the virtual antennas, was found to be in the range 0.4 λ to 0.5 λ. Using as many data samples per element as possible, lowered the variance of the LOS solutions. For a sampling interval of 41.6 μs per sample (data) point and the parameters given in Table 4-3, the virtual elements are separated by 309 data points. Therefore, 300 samples per element will use almost all of the available data.

Regarding array size, a 2×2 array gave comparable results to a 2×5 array for the same amount of total data used. For example, with 10,000 available data points and 200 data samples per array element, a total of 10 data sets were available to the 2×5 array and 25 data sets to the 2×2 array. For the 2×2 array, 23 of the 25 data sets yielded a LOS solution and the mean of these solutions was 5.5° in error. In the case of the 2×5 array, 9 LOS solutions were obtained in 10 trials and the LOS mean was 4° in error.

There are however, two advantages to using a smaller array: processing time and efficient data use. The data processing for the 2×2 array ran almost twice as fast as that of the 2×5 array. With regards to data efficiency, if the amount of available data is fixed, more data samples per element are available for a 2×2 array than for a 2×5 array. If instead the number of data samples per element is fixed, the 2×2 array will require less data and therefore less distance travelled.

4.3.6 Simulation Results for Statistical Tests

Simulations were run for 15 different sets of clusters. The cluster angles were randomly selected and the angular width of all clusters was 50° (± 25°). For each cluster set, PETRA estimated the AOA of the LOS cluster for 200 consecutive data sets. This required a total of 123,600 data points which cover a distance of 200 λ or 71.4 metres at the given frequency. At 50 km/hr, this distance is travelled in 5.1 seconds. A total of 200 simulations was considered adequate to give a LOS majority and to obtain a reasonably accurate LOS AOA estimate.

The results of these simulations are given in Table 4-4. The cluster angles listed in the second column for each particular set, are listed in the order of arrival. Hence, the first angle is the true LOS cluster AOA.

Table 4-4 Simulation Results for PETRA

Set	Cluster Angles	No. of LOS Solutions	Mean of LOS Solutions	Variance of LOS Solutions	LOS Error
1	240° 300° 90°	161/200	236.6°	170.1 $(°)^2$	3.4°
2	120° 40° 200°	130/200	122.8°	286.9 $(°)^2$	2.8°
3	325° 175° 115°	151/200	323.4°	361.1 $(°)^2$	1.6°
4	180° 240° 270°	129/200	174.2°	659.8 $(°)^2$	5.8°
5	328° 271° 295°	108/200	345.2°	1920.5 $(°)^2$	17.2°
6	19° 174° 321°	169/200	19.0°	164.8 $(°)^2$	0.02°
7	250° 275° 79°	124/200	238.0°	276.3 $(°)^2$	12.0°
8	226° 84° 137°	159/200	239.4°	480.8 $(°)^2$	13.4°
9	331° 67° 243°	108/200	329.7°	115.2 $(°)^2$	1.4°
10	319° 300° 45°	90/200	326.7°	131.2 $(°)^2$	7.7°
11	86° 352° 158°	100/200	91.9°	152.9 $(°)^2$	5.9°
12	206° 322° 138°	119/200	193.0°	189.1 $(°)^2$	12.9°
13	180° 235° 0°	146/200	180.2°	72.0 $(°)^2$	0.2°
14	59° 339° 320°	112/200	58.8°	322.9 $(°)^2$	0.2°
15	216° 7° 278°	153/200	215.5°	323.0 $(°)^2$	0.5°
Average	-	**131/200**	-	**375.1** $(°)^2$	**5.7°**

The average difference between the LOS cluster AOA estimated by PETRA over 200 simulations and the true AOA specified in the Hashemi model was 5.7° (LOS Error). For sets 5, 7, 8, and 12, the LOS error was greater than 10°. If these 4 sets are disregarded, the average LOS error becomes 2.7°.

4.3.7 Analysis of Results

Set 5 had the highest LOS error. Upon closer inspection of the 200 trial solutions, the reason became obvious. The AOA solution for each of the 200 trials had to be classified as to which of the three clusters it corresponded to. This classification was done according to boundaries drawn halfway between the actual cluster angles. For example, in set 5 any PETRA solution which fell in the range 311.5° - 119.5° was classified as an AOA estimate of the 328° cluster. If the solutions which fall into this range all lie close to 328°, this method is successful. However, as indicated by the very large variance, some solutions in this range, were not at all close to 328°, but instead congregated around other angles in this range. The mean AOA solution for 16 of the 200 trials was 87° and for another 13 it was 184°. The mean solution of 87° falls into the range corresponding to the LOS cluster and will obviously introduce a large amount of error. If the 16 solutions with mean of 87° are removed from the 108 LOS solutions, the new mean of the remaining 92 solutions is only 0.5° in error.

Sets 7, 8, and 12 also had LOS errors far higher than the average. However, the LOS estimate variance for these three sets was close to, and in two cases even less than the average. Therefore, the reason for their high LOS error is different from that of set 5. In set 7, the span of only 25° between the LOS cluster and 2nd cluster, may be the cause. For set 8 there is no apparent reason; the LOS estimates are just poor. In the case of set 12, a histogram of the 200 trial solutions clearly shows that MUSIC merged the LOS cluster with the third cluster. This resulted in numerous solutions midway between the two clusters.

4.3.8 Critique of PETRA

A total of 200 trials were performed in order to ensure a LOS estimate majority as well as good accuracy. In doing so, it was found that on the average, 66% of the solutions produced by PETRA correspond to the LOS cluster. Therefore, as expected, PETRA must run on several sets of data. The number of data sets need not be as high as 200, but must be large enough to produce a clear majority of LOS solutions.

The number of data sets required to give a reliable LOS AOA is also dependent on the way the non-LOS solutions (those solutions which correspond to the 2nd and 3rd clusters) are distributed amongst the LOS solutions. It was found that the non-LOS solutions often occurred in groups of adjacent trials with the largest group being 9 (sets 4 and 11) and the average being 6. It was also found that blocks of data would yield certain solutions no matter how fine or coarse the block was divided into data sets. The number of data sets would have to be more than twice the largest block of non-LOS solutions to ensure a LOS majority.

As previously mentioned, the accuracy of the individual PETRA LOS AOA estimates is poor. The RMS error for the LOS solutions, averaged over all 15 cluster sets, is 18.4°. This is not sufficient to obtain the positional accuracy required. A far more accurate LOS AOA estimate is obtained if a number of LOS PETRA solutions are averaged (5.7°). Although less than 200 PETRA solutions may give a LOS majority, the accuracy of the LOS AOA estimate will most likely suffer with fewer solutions to average over. The requirement for a large number of solutions to obtain an accurate estimate is PETRA's greatest disadvantage.

Another drawback of PETRA is the processing time. Although the eigenanalysis of the data correlation matrix is performed only once per data set, the MUSIC spectrum is generated three times in order to ensure that the dominant peak is accurately tracked. This is compounded by the requirement of several data sets, in terms of processing time and the time required to collect the necessary data.

The poor accuracy in set 5 was due to the method of classifying the AOA solutions. This method is not practical since it requires apriori knowledge of the AOA of each cluster. A more practical method is to use a sliding window. The concept is to window the sorted solutions such that the largest number of solutions fall into a window of smallest possible width.

A sliding window was tested on the 200 trials of set 1. The solutions were first sorted into ascending order. For the purpose of this exercise alone, a variable width window, centered at the actual LOS cluster AOA, was then used to determine the saturation window width. The saturation window width is that width, which when increased, does not result in a significant increase in the number of window contents.

For the data of set 1, the saturation width was found to be approximately 30°. The 30° window from 222° - 252° contained the maximum number estimates of 132. The mean of the solutions contained in this window is 232.8°. Using this figure as the LOS AOA estimate gives a large error of 7.2°. On the other hand, the centre of the window is 237° which is only 3° different from the true LOS AOA. A sliding window is therefore a promising technique for estimating the LOS AOA from a series of PETRA solutions.

4.3.9 Effect of Mobile Speed

The effect of the distance between data samples was observed by repeating sets 1 through 4 of Table 4-4 at 5 km/hr instead of 50 km/hr. With the data time interval held constant at 41.6 μs, the data samples are 10 times closer than when the speed is 50 km/hr. As a result, the virtual elements are separated by 0.05 λ instead of 0.5 λ. Table 4-5 compares the results at 5 km/hr with the corresponding results for 50 km/hr from Table 4-4.

Table 4-5 Comparison of 50 km/hr and 5 km/hr Results

Set	No. of LOS Solutions in 200 50 / 5 (km/hr)	Mean of LOS Solutions 50 / 5 (km/hr)	Variance of LOS Solutions 50 / 5 (km/hr)	LOS Error 50 / 5 (km/hr)
1	161 / 186	236.6° / 233.6°	$170.1\,(°)^2$ / $95.4\,(°)^2$	3.4° / 6.4°
2	130 / 112	122.8° / 128.7°	$286.9\,(°)^2$ / $322.6\,(°)^2$	2.8° / 1.3°
3	151 / 111	323.4° / 313.5°	$361.1\,(°)^2$ / $779.4\,(°)^2$	1.6° / 11.5°
4	129 / 149	174.2° / 173.5°	$659.8\,(°)^2$ / $990.6\,(°)^2$	5.8° / 6.5°
Average	**143 / 140**	-	**$370\,(°)^2$ / $547\,(°)^2$**	**3.4° / 6.4°**

The results of Table 4-5 indicate that spatial distance between data points has a negligible effect on the number of LOS solutions but a significant effect on the accuracy of these solutions. This decrease in accuracy may be attributed to the smaller distance between virtual elements rather than the smaller distance between data samples allotted to each individual element. The performance of any resolving type of algorithm will suffer as the interelement spacing of the antenna array decreases.

4.3.10 Effect of Noise

To determine the effect of SNR on the performance of MUSIC and PETRA, white Gaussian noise was added to the data produced by Hashemi's model. The powers of all 23 arrivals were summed together to calculate the SNR. As a result, the SNR figures are pessimistic since only the arrivals in the LOS cluster are of concern.

Set 4 of Table 4-4, was chosen for noise tests since the results for that set are very close to the average for the 15 sets. A total of 200 trials were run for each value of SNR tested. The results are shown in Figure 4-4.

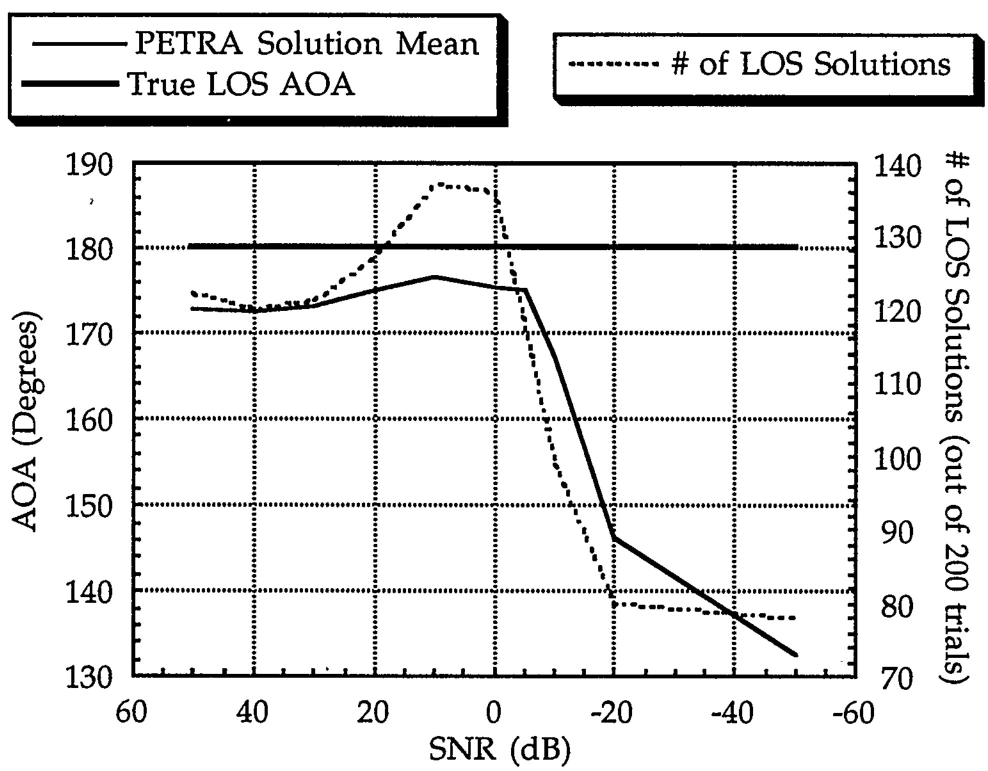

Figure 4-4 Noise Results for Set 4

In Figure 4-4, the thin solid line corresponds to the average of the LOS solutions given by PETRA. Each of the 200 solutions was again classified by the method requiring apriori knowledge of the cluster angles as described in section 4.3.7 . The bold solid line shows the true AOA of the LOS cluster, 180°. Both of the solid lines are referenced to the left-hand vertical axis. Referenced to the right-hand vertical axis, the dashed line is a measure of the number of LOS solutions per 200 trials.

Figure 4-4 illustrates two important points. First, the accuracy of the LOS estimate, given by PETRA, does not begin to deteriorate significantly until the SNR drops below 0 dB. This demonstrates the robustness of MUSIC when the noise power is comparable to the signal power. Second, there is a significant performance improvement as the SNR drops to 0 dB. Both the number of LOS solutions per 200 trials, and the accuracy of the mean of these solutions increase as SNR drops from approximately 30 dB to 10 dB. Recall from Table 4-1 that the second and third clusters are of lesser power than the LOS cluster. As the noise power increases, the second and third clusters become overwhelmed by the noise, leaving only the LOS cluster above the noise floor. Hence it becomes more likely that the PETRA solution will correspond to the LOS cluster since there is only one cluster to choose from. Accordingly, the accuracy of the LOS estimates will also increase since there are no other clusters to skew or affect the estimate.

CHAPTER 5
TECHNIQUES #2 & #3: RESULTS & ANALYSIS

5.1 Introduction

In this chapter, two additional techniques to identify the LOS cluster and estimate its AOA, in a three cluster environment, are considered. If PETRA is called technique #1, then the two approaches considered in this chapter are named technique #2 and technique #3. The primary objective in developing techniques #2 and #3 is to achieve the required AOA accuracy without the need for a large number of estimates to average over.

Technique #2 uses spatial MUSIC with forward and backward spatial smoothing. The LOS cluster is identified by clustering individual arrivals and choosing that group with the most arrivals. Technique #3 is based on temporal MUSIC. Three signals corresponding to the three clusters are resolved and the signal with the highest amplitude is chosen as LOS.

Simulations are carried out for both techniques using the same three cluster signal environment described in section 4.2. Techniques #2 and #3 are first described in more detail and their simulation results presented separately. This is followed by a comparison of the major results for PETRA and techniques #2 and #3.

The chapter concludes with an investigation of temporal MUSIC for the purpose of identifying the LOS arrival in a single cluster environment.

5.2 Technique #2: Resolving for Individual Arrivals

5.2.1 Introduction

Whereas PETRA only attempts to resolve for clusters, technique #2 is based on the ability of MUSIC to resolve for as many of the individual arrivals which make up the clusters as possible. Identification of the LOS cluster is based on one of the assumptions stated in section 4.2; the LOS cluster contains more arrivals than any other cluster. If the individual arrivals in the MUSIC spectrum can be grouped into clusters, the cluster with the most arrivals is chosen as LOS and the AOA of that cluster is estimated by the mean of the arrival angles in that cluster.

If the assumption regarding the relative number of arrivals in the clusters is true, and MUSIC is able to resolve for all of the arrivals, this technique would be fool proof. However, a number of factors exist which prohibit this. The subarray size must be at least one greater than the number of arrivals to be resolved for. Hence, depending on the number of arrivals, practical data limitations may exist. Perhaps of greater significance is the ability of MUSIC to resolve for numerous arrivals. General experience has shown that there is a limit. Spatial MUSIC as implemented by the author was never able to accurately resolve for more than approximately 12 signals no matter how far apart the signals were in terms of AOA. Of course, signals which are very close (approximately 2° and less) may not be distinguishable.

The above drawbacks proved not to be as serious as originally thought. Resolving for fewer than the total number of arrivals still gives an indication of which cluster contains more signals. Consider the MUSIC spectrum shown in

Figure 5-1. Three clusters exist as defined in section 4.2. The three clusters are each 20° (±10°) wide and are centred at 240°, 300°, and 90° in order of time of arrival. The 240° cluster is LOS and therefore contains more arrivals than the other two clusters. The spectrum was obtained from simulated data for two antennas separated by 0.5 λ and moving at 50 km/hr. The antennas were sampled every half wavelength (17.9 cm at 840 MHz), a total of 18 times. The subarrays were of size 2×9 and MUSIC was asked to resolve for 12 arrivals.

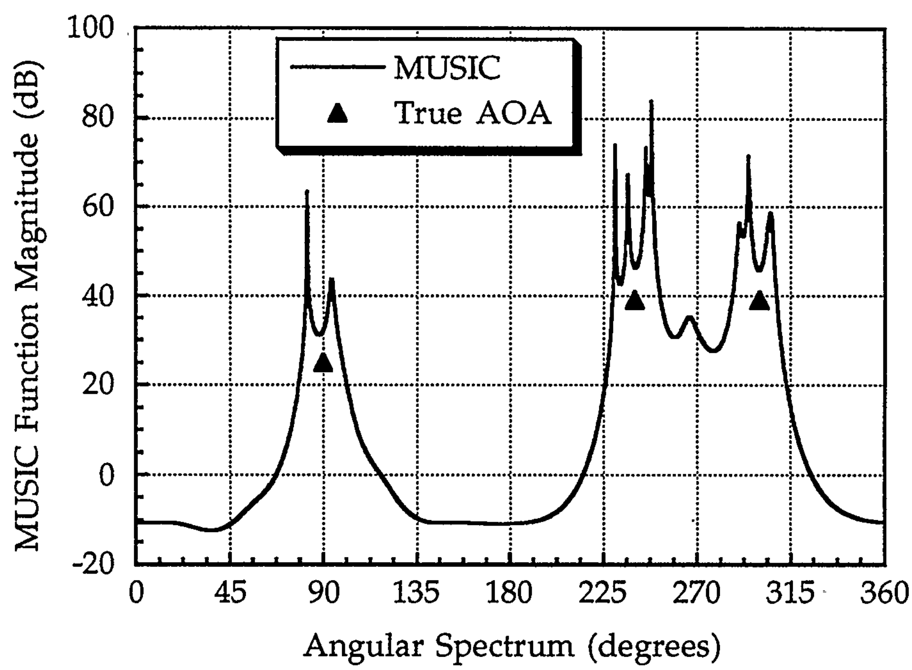

Figure 5-1 Resolving Individual Arrivals with Forward and Backward Smoothing

Clearly seen in Figure 5-1 is that only 9 arrivals of the 23 present are resolved (a maximum of 12 are resolvable since that is all that was asked for). However, it is obvious that 3 clusters exist and that the cluster at 240° contains more arrivals.

Therefore it is not necessary to resolve all of the existing arrivals in order to identify the LOS cluster.

5.2.2 The Process

Unfortunately, not all clusters are as clearly segregated as those in Figure 5-1. Depending on distance between the clusters in terms of AOA , grouping arrivals is a potentially difficult process. When two clusters are close, the arrivals in the spectrum may be such that it is impossible to know that two clusters exist or where one ends and the other begins. This creates a further problem; if the second and third clusters are so close that they are indistinguishable, the total number of arrivals in the combined cluster may outnumber the arrivals in the true LOS cluster. As a result an error will be made in LOS cluster identification.

A sophisticated windowing technique, which takes into account numerous factors, is required for grouping the arrivals in the spectrum. Due to time constraints it was not possible to incorporate such a technique for the results presented. The grouping of the arrivals was therefore done "manually". This approach is of course not entirely objective. It does, however, suffice to give an indication of the performance of technique #2 should a sophisticated, software implemented, grouping process be used.

A number of factors were considered when grouping the arrivals. For spectrums such as Figure 5-1 the clusters are obvious. If the clusters were not as obvious, boundaries halfway between the true cluster angles were used. This required apriori knowledge of the cluster angles and as a result is certainly not practical. The distance, in terms of AOA, between adjacent peaks in the spectrum is an important factor. The primary question is how close must arrivals be in order to be considered as part of the same cluster? Two criteria were used to answer this question. First, the relative heights of the peaks was considered. In Figure 5-1, for example, the small peak at 265° is not considered to be part of the 240° or 300° clusters. A threshold of 0 dB was used to eliminate small spurious peaks. Secondly, 20° was used as a round number for the maximum distance between arrivals of the same cluster. It is expected that a considerable amount of work would be required to investigate the nature of typical MUSIC spectrums in order to develop a suitable grouping technique.

5.2.3 Simulation Parameters

The cluster sets used to test technique #2 are identical to those used to test PETRA. The simulation parameters which determine the data matrix differ however and are summarized in Table 5-1.

Table 5-1 Simulation Parameters for Technique #2

Number of physical antennas	2
Array size	2×24
Subarray size	2×17
Physical antenna spacing	0.5 λ
Virtual antenna spacing	0.32 λ
Signal subspace dimension	10
Vehicle speed	50 km/hr
Frequency	840 MHz

The vehicle speed, frequency of transmission, number of physical antennas and their separation are the same as for the PETRA simulations. The most significant difference is the form of smoothing used. The version of spatial MUSIC used in PETRA, employed temporal smoothing based on snapshots of data from an array. MUSIC as used in technique #2 uses forward and backward spatial smoothing. As previously described, the two forms of smoothing, though similar for spatially and temporally separated data, differ with regards to the extent of smoothing they introduce. Each element in the 2×24 array of Table 5-1 has only one data sample as opposed to the 300 samples per element used in the temporal smoothing of technique #1.

Twenty-four virtual elements were chosen such that either physical antenna could provide enough data to stand alone. Tests to determine the optimum subarray size confirm linear prediction theory which states that the prediction order (subarray size) be 0.75 of the number of array elements [18]. Signal subspace dimension refers to the number of eigenvectors spanning the signal subspace or equivalently, the number of signals MUSIC is asked to resolve. The value 10 is sufficient to identify the LOS cluster. It also satisfies the equations which relate the number of signals to the subarray size as well as to the total number of array elements required for forward and backward spatial smoothing (see section 3.2.4.1). Tests also determined that a virtual element spacing greater than 0.32 λ gained nothing and a spacing less than 0.32 λ reduced resolution.

Although 24 virtual elements per physical antenna are more than enough to solve for 10 signals using one direction of smoothing, it was found that smoothing in both directions is advantageous. The MUSIC spectrum in Figure 5-

1 was created using both forward and backward spatial smoothing. The spectrum of Figure 5-2 was generated in an identical manner except that only forward smoothing was used. For the purpose of clustering, Figure 5-1 is clearly superior to Figure 5-2. In general, it was found that when smoothing in both directions was used, the spectrums were smoother and more representative of the cluster nature of the signal environment. Therefore, both forward and backward smoothing are used.

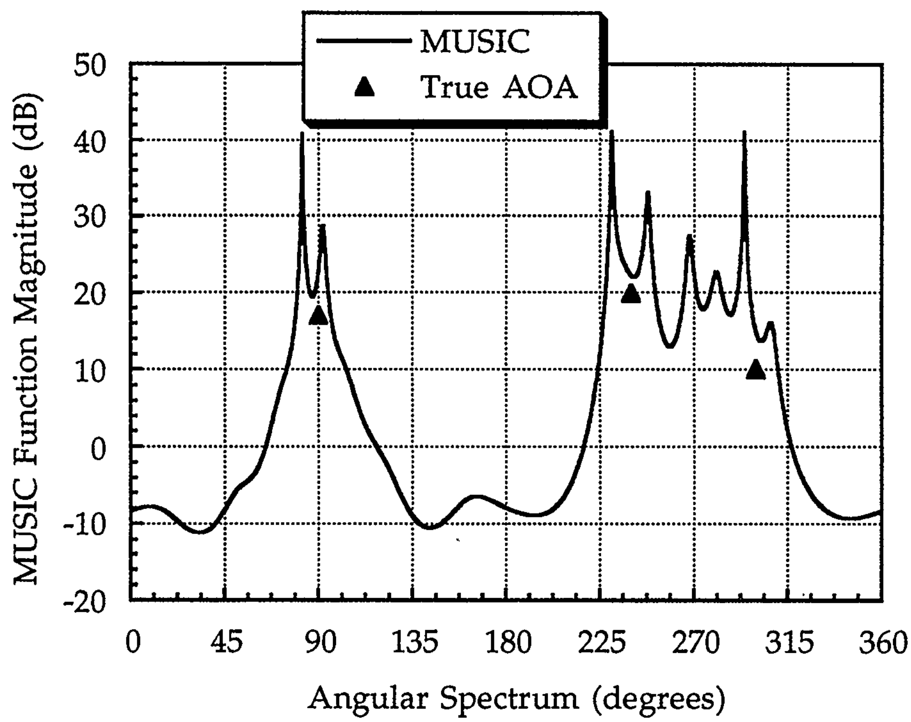

Figure 5-2 Resolving Individual Arrivals with Forward Smoothing Only

5.2.4 Simulation Results

For each of the 15 cluster sets, 20 simulations were run. The distance travelled to collect 24 data points spaced by 0.32 λ with a sampling rate of 41.6 μs/sample, is 2.67 metres (7.5 λ). Therefore, enough data is collected for 20 simulations in 55.6

metres. This distance is covered in 4 seconds at 50 km/hr. The simulation results are shown in Table 5-2. The cluster angles listed in the second column for each particular set, are listed in the order of arrival. Consequently, the first angle is the true LOS cluster AOA.

Table 5-2 Simulation Results for Technique #2

Set	Cluster Angles	No. of LOS Solutions	Percent by Peak Count	Percent by Peak Magnitude	RMS LOS Error with respect to:	
					centre	average
1	240° 300° 90°	19/20	84.2%	15.8%	4.5°	5.0°
2	120° 40° 200°	16/20	81.3%	18.7%	2.9°	3.0°
3	325° 175° 115°	17/20	76.5%	23.5%	4.1°	5.4°
4	180° 240° 270°	0/20	-	-	-	-
5	328° 271° 295°	3/20	66.7%	33.3%	1.9°	4.0°
6	19° 174° 321°	11/20	63.6%	36.4%	5.2°	6.2°
7	250° 275° 79°	5/20	20%	80%	5.2°	4.1°
8	226° 84° 137°	16/20	93.8%	6.2%	19.5°	11.4°
9	331° 67° 243°	7/20	57.1%	42.9%	2.8°	3.9°
10	319° 300° 45°	20/20	100%	0%	13.2°	13.8°
11	86° 352° 158°	19/20	94.7%	5.3%	3.6°	2.9°
12	206° 322° 138°	1/20	100%	0%	0.5°	7.3°
13	180° 235° 0°	0/20	-	-	-	-
14	59° 339° 320°	18/20	61.1%	38.9%	5.0°	5.9°
15	216° 7° 278°	20/20	80%	20%	4.1°	3.6°
Mean	-	**11.5/20**	**75.3%**	**24.7%**	**5.6°**	**5.9°**

The "No. of LOS solutions" refers to the number of simulations out of 20 in which the cluster chosen as LOS was indeed the true LOS cluster. This number is further broken down in the next two columns of Table 5-2. In the event that two clusters had an equal number of arrivals, the magnitude of the MUSIC peaks were used to make the decision. In such a case, the cluster with the highest average magnitude of peaks was chosen as LOS. Hence, "Percent by Peak Count" refers to the percentage of LOS solutions which were chosen because that cluster contained the most peaks. "Percent by Peak Magnitude" corresponds to the cases in which there was a tie and the cluster with the highest average peak magnitude was chosen. Take set 1 as an example. In 19 out of the 20 simulations, the cluster chosen as LOS was indeed the LOS cluster. In 16 out of those 19 simulations (84.2%), the cluster was chosen because it contained the most arrivals. In the other 3 of the 19 simulations (15.8%), there was a tie and the correct cluster was chosen because of peak magnitude.

The LOS cluster AOA was estimated by averaging the AOA of each of the peaks (arrivals) in the correctly identified LOS cluster. The RMS error of these estimates was then found with respect to the spatial centre of the Hashemi LOS cluster (second last column of Table 5-2) as well as the average arrival angle of the Hashemi LOS cluster (last column).

5.2.5 Analysis of Results

The true LOS cluster was identified by a slim majority of 57.5%. For 5 of the 15 cluster sets, the number of correct identifications was significantly less than 50%. In the case of set 4, the LOS cluster was never correctly identified because the 240° and 270° are close enough that they appear as one. As a result, their combined peaks always outnumbered those of the 180° cluster. For sets 12 and 13 there is no apparent reason for the dismal success rate. In both cases the peak count of the second cluster outnumbered the peak count of the LOS cluster, two to one. Two or more of the clusters in sets 5 and 7 were sufficiently close that it was very difficult to distinguish between clusters. As a result, errors in the grouping process are likely to be numerous.

Of the solutions that correctly identified the LOS cluster, a sizable majority (75.3%) were so reached by peak count as opposed to relative peak magnitude. Not appearing in Table 5-2 is the number of correct solutions arrived at by peak count as a percentage of the total number of simulations. On the average, for 46% of the simulations, the LOS cluster was correctly identified by peak count. Breaking ties with relative peak magnitude increases the success rate to 57.5%. However, since the height of a peak is not an indicator of the power of the signal it represents, this process is questionable.

The average accuracy of the AOA estimate was 5.6° RMS. For sets 8 and 10 the error was much larger. Again the small angular separation between clusters is to blame in the case of set 10. The 319° and 300° clusters appear as one and the effect of the 300° cluster is to skew the AOA estimate down. For set 8, spurious peaks close enough to the LOS cluster that they must be considered as part of the cluster by an objective grouping technique, have again skewed the LOS AOA estimate.

Bias in the AOA estimates was observed in the results of technique #2 as it was in those of PETRA. For example, bias is apparent in Figure 4-4. In the results for technique #2, a general trend regarding bias was observed. For larger absolute errors (i.e. > 5°), the errors tended to be one-sided and hence the AOA estimates skewed in one direction. When the absolute error was smaller no bias was evident.

The cause of the bias is not due to a bias within the LOS cluster generated by Hashemi. This is demonstrated by the results of Table 5-2. The RMS error does not significantly change when it is calculated with respect to the average arrival angle of the cluster as opposed to the spatial centre. It is likely that the presence of the two other clusters affects and biases the AOA estimate for the LOS cluster.

5.3 Technique #3: Resolving Cluster Angles with Temporal MUSIC

5.3.1 Introduction

For cases in which the amplitude of the arriving signals is desired, temporal MUSIC is convenient to use since it directly yields frequency. As described in section 3.3, the frequencies of the arrivals can be used to estimate the amplitude and phase of the arrivals. Using data from two physical antennas, the phase information may be used to eliminate the cosine and sine ambiguities. In this section, the amplitude information is used to identify the LOS cluster. That the signals comprising the LOS cluster are of higher power than those signals in the other clusters, is an assumption made in this thesis. Therefore, the LOS cluster should be recognizable by determining the amplitude of the individual clusters.

5.3.2 The Process

In concept, the process of technique #3 is very simple. As in PETRA, MUSIC is set up to resolve only 3 signals and the assumption is made that an average of each cluster is obtained rather than individual arrivals. In conjunction with the original data collected, the frequencies estimated by MUSIC are used to estimate the corresponding amplitudes and phases. The frequencies are converted to AOA and the phases are used to eliminate the ambiguities. The signal of highest amplitude is chosen as LOS and its AOA is the LOS AOA estimate.

5.3.3 Simulation Parameters

Temporal MUSIC can be thought of as processing a time series of data with an FIR transversal filter. Two time series are required in order to ensure a unique spectrum. The length of the filter is comparable to the subarray dimension in spatial MUSIC. Hence, the length of the filter must be at least one greater than the number of signals to be resolved. In order to facilitate some smoothing, the data record length must be at least one greater than the filter length.

In contrast to spatial MUSIC, in which the subarrays consist of data from each of the two physical antennas (see Figure 3-4), temporal MUSIC processes the two data series separately. Forward and backward smoothing was used and therefore the data matrix for each of the two antennas was constructed as equation (3-48). Thus the temporal smoothing used here differs from that used in technique #1. The simulation parameters are summarized in Table 5-3.

Table 5-3 Simulation Parameters for Technique #3

Number of physical antennas	2
Data series length	5
Filter length	4
Physical antenna spacing	0.25 λ
Data point spacing	0.0647 λ
Signal subspace dimension	3
Vehicle speed	50 km/hr
Frequency	840 MHz

An advantage of temporal MUSIC over the spatial form is immediately obvious upon comparison of Table 5-3 with Tables 4-3 and 5-1. Temporal MUSIC requires far fewer points to make an AOA estimate. In addition the points need not be spaced as far apart as with spatial MUSIC.

A physical antenna spacing of 0.25 λ was used as opposed to the 0.5 λ used with techniques #1 and #2. To a certain extent, eliminating the angular ambiguity with phase was more successful when the two time series were spatially closer.

5.3.4 Simulation Results

Simulations were run using the same 15 cluster sets used for techniques #1 and #2. The distance travelled to collect 5 data points spaced by 0.0647 λ, with a sampling rate of 41.6 μs/sample, is 0.095 metres. Fifty simulations were conducted for each cluster set. For 50 simulations, enough data is collected in 5.9 metres. This distance is covered in 0.42 seconds at 50 km/hr.

Although technique #3 is simple in concept, the implementation was not. A significant problem encountered was the pairing up of signals obtained from the two time series for the purpose of eliminating the angular ambiguity. Although only three signals were to be resolved, there were often more than three or less than three peaks in the MUSIC spectrum. As a result, the number of peaks produced by one time series was not always equal to the same produced by the second time series. When this was the case, pairing of the signals was done manually. This was further complicated by the fact that the amplitudes, phases, and AOA of signals resolved from the two time series were often different enough to cast doubt on which signals should be paired up.

Another major problem was eliminating the angular ambiguity. This had two possible causes. In those cases in which the pairing of signals was at best a guess, errors most probably existed. The angular ambiguity elimination process was in those cases doomed to fail. At other times, when there was no doubt regarding the pairing of signals, the phases themselves had to be blamed. The frequency matrix of equation (3-53) must be inverted to determine the amplitudes and phases. It is possible that this matrix was at times close to being nonsingular. This would cast suspicion on the resulting phase values.

Due to the above problems, not all simulations were successful. A decision process was required to determine whether any particular simulation was valid or not. Two criteria were used: signal amplitude and elimination of the angular ambiguity. For a simulation to be considered valid, the angular ambiguity of the signal of highest amplitude had to be eliminated successfully.

Table 5-4 Simulation Results for Technique #3

Set	Cluster Angles	% of Valid Solutions	RMS Valid Error w.r.t.		% of LOS Solutions	RMS LOS Error with respect to:	
			centre	average		centre	average
1	240° 300° 90°	66%	11.1°	11.0°	66%	11.1°	11.0°
2	120° 40° 200°	38%	32.7°	28.8°	30%	11.1°	8.5°
3	325° 175° 115°	42%	66.6°	67.2°	34%	11.0°	11.2°
4	180° 240° 270°	58%	32.0°	32.0°	40%	19.2°	19.2°
5	328° 271° 295°	58%	25.8°	27.5°	38%	12.6°	13.9°
6	19° 174° 321°	64%	30.3°	31.8°	42%	12.9°	8.4°
7	250° 275° 79°	54%	15.8°	14.3°	54%	15.8°	14.3°
8	226° 84° 137°	70%	29.8°	33.6°	64%	10.3°	14.7°
9	331° 67° 243°	52%	11.8°	11.8°	52%	11.8°	11.8°
10	319° 300° 45°	54%	27.1°	27.0°	44%	9.8°	9.7°
11	86° 352° 158°	16%	30.1°	30.6°	10%	8.4°	7.6°
12	206° 322° 138°	48%	36.7°	36.4°	34%	8.8°	13.9°
13	180° 235° 0°	58%	31.0°	31.0°	28%	18.6°	18.6°
14	59° 339° 320°	56%	38.4°	38.0°	46%	14.5°	14.4°
15	216° 7° 278°	58%	32.3°	31.2°	52%	9.0°	7.7°
Mean	-	**52.8%**	**30.1°**	**30.1°**	**42.3%**	**12.3°**	**12.3°**

The simulation results are presented in Table 5-4. The "% of Valid Solutions", is the number of simulations, as a percentage of 50, which are valid in the sense described above. The "RMS Valid Error" is the RMS error of the valid solutions calculated with respect to the spatial centre of the Hashemi LOS cluster and with respect to the average arrival angle of the Hashemi LOS cluster. The next three columns contain the same type of information but only for the LOS solutions within the valid solutions. LOS solutions are those for which the absolute error is less than 25°. Any solution with an error greater than 25° is considered to correspond to a non-LOS cluster. The "% of LOS Solutions" is also represented as a percentage of 50.

5.3.5 Analysis of Results

The results of Table 5-4 demonstrate that in only a slim majority of simulations was the angular ambiguity of the signal of highest amplitude eliminated successfully. As discussed in the previous section, eliminating the angular ambiguity with phase is very problematic. However, these problems seem to be particular to the cluster environment. When temporal MUSIC was used to estimate the AOA of a few individual arrivals, eliminating the ambiguity by phase worked much better; indeed the case of unsuccessful ambiguity elimination was very rare. When resolving for clusters as opposed to individual arrivals, it is likely that the signals resolved from the two individual data series did not correspond close enough. This was also seen when comparing the AOA and amplitude of signals resolved from consecutive data sets of the same data series. Often the MUSIC spectrum and signal amplitudes were quite different. These inconsistencies are not a major problem in spatial MUSIC where the angular ambiguity is inherently eliminated.

A much more promising result is the number of valid solutions which correspond to the true LOS cluster. Expressed as a percentage of the total number of simulations, this number is less than a majority (42.3%). However, 80.1% of the valid solutions corresponded to the LOS cluster. The other 19.9% represented one of the other clusters or spurious peaks. Hence, of the 3 techniques for estimating the LOS AOA, technique #3 has the highest success rate in terms of identifying the LOS cluster.

Those valid solutions which were non-LOS, of course had high AOA errors and therefore inflated the RMS error. When these solutions were removed, the average RMS error decreased from 30.1° to 12.3°. How much the error decreased for specific cluster sets, of course depended on how many non-LOS solutions existed and how far they were from the true LOS AOA.

An RMS error of 12.3° is roughly twice as large as the errors obtained with techniques #1 and #2. Evidently, temporal MUSIC with the number of data points used, does not estimate the "average" AOA of a cluster as well as spatial MUSIC does in technique #1. For one cluster set, the cluster width was decreased to 10° (±5°). This resulted in an absolute error of 2.3°. The number of valid solutions, and the percentage of those which were LOS however, did not increase. As expected, cluster width has a significant effect on accuracy.

The RMS errors in Table 5-4 again confirm that there is little if any bias in the arrival angles of the Hashemi LOS cluster. The RMS error is the same regardless of whether the spatial centre or the average arrival angle of the Hashemi LOS cluster is used to calculate it. Any bias in the AOA estimates of the valid solutions is most likely due to the non-LOS solutions. This will however, depend on whether more than one non-LOS cluster has the tendency to be chosen and whether these non-LOS solutions lie on either side of the LOS AOA. When considering only the valid LOS solutions, large biases were evident in some cases and small biases in others. Again it is likely that the non-LOS clusters have an effect in those cases which exhibit a bias.

5.4 Comparison of Techniques #1, #2, and #3

Table 5-5 summarizes and compares the significant results for all three techniques.

Table 5-5 Summary of Results for Techniques #1, #2, and #3

Technique #	Mean Percentage of Success	AOA Error
1	65.5%	5.7°
2	57.5%	5.6°
3	80.1%	12.3°

Mean percentage of success refers to the number of times the method was able to identify the LOS cluster and estimate its AOA, expressed as a percentage of the simulations conducted or, in the case of technique #3, the valid simulations. For technique #1, every simulation resulted in a solution but not all solutions corresponded to LOS. A successful solution is defined as one which does correspond to LOS. Technique #2 is similar in that each simulation resulted in a spectrum from which a solution could be obtained. Solutions for which the largest group of peaks were clustered around the LOS AOA are deemed successful. Technique #3 is somewhat different in that the angular ambiguity elimination process did not work for every simulation. Practically, it would be simple to throw away the invalid solutions. For the purposes of comparison, the ability of technique #3 to produce a solution for which the angular ambiguity is eliminated, is separated from performance in terms of LOS cluster identification and AOA estimation. The results in Table 5-5 are based on the latter.

In addition to the results of Table 5-5, the three techniques may also be compared in terms of the amount of data required and complexity. For the results presented in chapters 4 and 5, the following simulated travel distances were required to collect enough properly spaced data, at 50 km/hr, to run one simulation (that is make one AOA estimate):

- technique #1 - 0.36 metres
- technique #2 - 2.67 metres
- technique #3 - 0.095 metres.

By far, temporal MUSIC requires the smallest interpoint spacing and therefore the shortest travel distance. Technique #2 requires many virtual array elements in order to resolve closely spaced individual arrivals. Techniques #1 and #3 do not require arrays as large since they only attempt to resolve for three clusters.

The number of simulations required to obtain a majority of LOS solutions and an accurate AOA estimate must also be considered. The results for technique #1 were obtained with 200 simulations. This required a travel distance of 72 metres. Techniques #2 and #3 required far fewer simulations and hence shorter travel distances, 53.4 metres and 4.75 metres respectively. When the transmitter and receiver are very close, the AOA will change quickly as the receiver moves. As a result, the shorter the travel distance to make an AOA estimate, the better. Therefore, in terms of interpoint spacing, the number of points, and the number of solutions required, technique #3 is clearly the best and technique #1 the poorest.

The actual number of data points used affects complexity as well as the distance travelled. Along with smoothing, the number of data points determines the size of the data matrix which must be multiplied by its Hermite to calculate the correlation matrix. The data matrix for technique #3 was the smallest (4×4) and even though 2 data matrices must be generated and processed, the number of computations involved will be by far the fewest. Technique #1 required a 300×4 data matrix whereas Technique #2 required a 16×34 data matrix.

Complexity is further affected by the size of the correlation matrix generated from the data matrix. The larger the correlation matrix, the larger the number of eigenvalues and eigenvectors to be calculated. Both techniques #1 and #3 generate a 4×4 correlation matrix whereas technique #2 generates a 34×34 correlation matrix. Obviously the eigenanalysis for technique #2 involves the most computations.

A third complexity factor is the type and amount of processing which needs to be done after a MUSIC spectrum has been generated. It is in this area that technique #3 suffers the most. To begin, two separate MUSIC spectrums must be generated from the data series of the two antennas. Secondly, the phase and amplitude of each resolved arrival must be computed. This calculation involves the inversion of a complex matrix. After this, the phase values are used to calculate AOA for an orthogonal reference and this AOA must be compared to the original AOA calculated from the frequencies resolved by MUSIC.

Techniques #1 and #2 are roughly equivalent in terms of post-processing. Technique #1 requires the generation of three MUSIC spectrums as well as the peak tracking process. In addition, a windowing or grouping process is required to distinguish the LOS solutions from those which are not. Technique #2 is most straightforward in the generation of the MUSIC spectrum but requires a sophisticated technique for clustering the peaks in the spectrum. Division of the solutions into two groups, those which are LOS and those which are not, is also required.

In conclusion, although technique #3 has many advantages, it does not achieve the required accuracy for the intended application. Techniques #1 and #2 have identical accuracy and comparable success rates. Although more complex, technique #2 requires a shorter travel distance for data collection than does technique #1. Therefore, technique #2 is the best choice for accurately estimating AOA in the shortest possible distance.

5.5 The Single Cluster Environment

5.5.1 Introduction

To this point, the assumed signal environment has consisted of three clusters. This chapter concludes with an investigation of a single cluster which contains a

LOS component. It is assumed that the LOS arrival is of significantly greater power than all others. Temporal MUSIC is therefore used to resolve for as many of the individual arrivals as possible. The arrival of highest amplitude is then identified as the LOS component.

A potential AOA estimation system for the multipath environment could use one of techniques #1, #2, or #3 to identify the LOS cluster and estimate its AOA. Using this knowledge, the other clusters may be spatially filtered out of the data. It is likely that a more accurate AOA estimate for LOS could then be achieved [33].

5.5.2 Cluster Configuration

The structure of the clusters used for simulation purposes is identical to that of the LOS cluster in the three cluster environment (see Table 4-1) except in one respect. That arrival closest to the centre of the cluster, in terms of AOA, is given a power 10 dB greater than the next most powerful arrival. With this exception, all other cluster parameters are identical, including the number of arrivals in the cluster (13), the cluster width (± 25°), the impulse response, and the power of each arrival.

The single clusters used here are identical to the LOS clusters used previously in terms of AOA as well. Table 5-7 contains the AOA of the LOS arrival in each cluster. Cluster #13 has not been included since it is identical to cluster #4.

5.5.3 Simulation Parameters

Temporal MUSIC, with forward and backward smoothing as per technique #3, is used. The simulation parameters are presented in Table 5-6.

Table 5-6 Simulation Parameters for Single Cluster Tests

Number of physical antennas	2
Data series length	25
Filter length	19
Physical antenna spacing	0.25 λ
Data point spacing	0.1 λ
Signal subspace dimension	13
Vehicle speed	50 km/hr
Frequency	840 MHz

Compared to the parameters used for technique #3 (Table 5-3), more points and more widely spaced points are used here. This is necessary since in the single cluster case, up to 13 individual, and relatively closely spaced arrivals are to be resolved. Technique #3 attempted to resolve only 3 widely spaced signals and therefore required fewer data points which could be more closely spaced.

5.5.4 Simulation Results

Simulations were conducted using the LOS clusters of the 15 cluster sets used in simulations for techniques #1, #2, and #3. Twenty simulations were run for each cluster. A distance of 0.85 metres was travelled to collect the 25 data points necessary for 1 simulation. Table 5-7 summarizes the results.

Table 5-7 Single Cluster Simulation Results

Cluster #	LOS AOA	% of Valid Solutions	RMS Error
1	237.1°	70%	5.2°
2	120.6°	85%	1.3°
3	325.5°	75%	2.5°
4	177.5°	60%	13.5°
5	328.7°	80%	3.3°
6	18.6°	90%	2.1°
7	247.3°	75%	2.3°
8	221.5°	80%	3.2°
9	330.6°	90%	2.3°
10	318.0°	80%	4.3°
11	79.8°	75%	0.9°
12	207.0°	100%	2.5°
14	58.6°	95%	2.2°
15	214.1°	90%	2.3°
Mean	-	**82%**	**3.4°**

In Table 5-7, LOS AOA refers to the AOA of the most powerful arrival in the cluster. In some cases, this LOS arrival was not very close to the spatial centre of the cluster. For example, the centre of cluster #11 is 86°. It is important to note that the errors reported in the results for techniques #1, #2, and #3 were with reference to the centre of the LOS cluster or the average arrival angle of the cluster. The RMS errors of Table 5-7 are referenced to the LOS AOA values shown.

Because temporal MUSIC was used, not all simulations yielded a solution. As in the case of technique #3 which also used temporal MUSIC, the angular ambiguity of the signal of highest amplitude was not always successfully eliminated. However, unlike technique #3, all solutions are assumed to be LOS. Therefore, percentage of valid solutions refers to the number of simulations, expressed as a percentage of 20, for which the angular ambiguity of the signal of highest amplitude (LOS) was successfully eliminated.

5.5.5 Analysis of Results

The success rates presented in Table 5-7 are far higher than those for technique #3 in Table 5-4. This is attributed to the fact that in the single cluster case, individual arrivals are resolved as opposed to clusters. Elimination of the angular ambiguity is far more successful when dealing with individually resolved arrivals than with signals which represent entire clusters. Put simply, angular ambiguity elimination with individual arrivals does not suffer from many of the problems which plague the process when dealing with whole clusters.

The RMS errors in Table 5-7 are lower than those for techniques #1, #2, and #3. Improved accuracy was predicted because of fewer potentially interfering signals from other clusters. More likely, better accuracy has been achieved because the

AOA of a specific individual arrival has been estimated whereas techniques #1, #2, and #3 all estimate an average AOA for a cluster.

Cluster #4 is the only cluster which had a significantly higher than average error. It also has the lowest success rate. Because of the cosine ambiguity, an arrival from say 165° will have the same Doppler frequency shift as a signal from 195°. Temporal MUSIC is therefore unable to distinguish the two. A cluster centred at 180° will have arrivals both above and below 180° and in the case of cluster #4, there are numerous pairs of independent arrivals which are reflections of one another around 180°. These pairs of arrivals may add constructively or destructively and as a result MUSIC will see arrivals spread from 155° to 180° instead of from 155° to 205°.

Since all solutions for cluster #4 were less than 180°, the mean signed error is -12.6°. Hence the average solution was 167.4° which is almost exactly the centre of the new half size cluster. Considering this, as well as the fact that the range of errors was from -5.6° to -21.2°, it appears that the AOA estimate is more of an average in this case. It is expected that a cluster centred at 0° would experience the same problem.

CHAPTER 6
ESTIMATING AOA WITH CURVED ARRAYS

6.1 Introduction

In all previous chapters, the simulated data has been for antennas travelling in a straight line. As a result, at least two antennas were necessary in order to create a two dimensional virtual array. An advantage of MUSIC is its ability to work for an array of any geometry. Thus the array need not be rectangular, nor the array elements uniformly spaced. We may therefore eliminate the need for two antennas, but still obtain a two dimensional virtual array, by sampling an antenna as it moves along any non-linear path.

The advantage of sampling along a non-linear trajectory is obvious; a MUSIC spectrum unique through 360° is obtained with only one antenna. There is a price to pay however, in the form of increased complexity. Since MUSIC requires knowledge of the array geometry, the trajectory of the antenna and relative positions of the data samples must be known. This would require additional instrumentation onboard the vehicle to detect the amount of curvature in the vehicle's travel path. In addition, smoothing and definition of the steering vector must be re-examined.

The primary focus of this chapter is not to determine whether a non-linear virtual array suppresses all cosine ambiguities. That the ambiguities are suppressed is demonstrated in Figure 6-1.

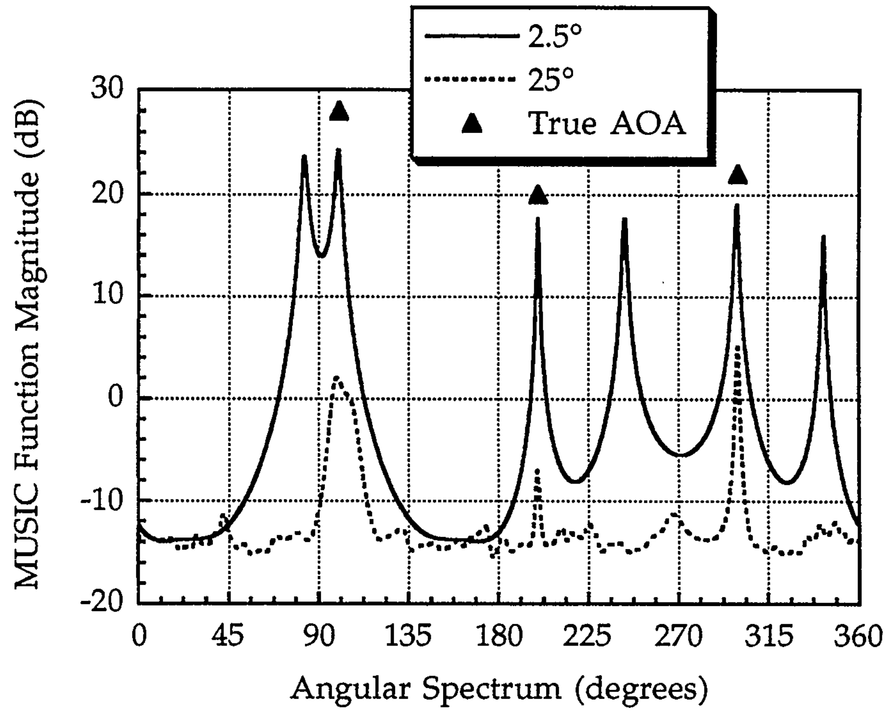

Figure 6-1 Suppression of Ambiguities

Three signals are present and the pointers in Figure 6-1 show their angular locations. The solid line represents the MUSIC spectrum for data collected over 2.5° of an arc of radius 20 metres. The three signals and their ambiguities are clearly evident. When the same number of data samples are collected over 25°, the resulting spectrum is indicated by the dashed line.

The primary question then is how much curvature is required in order to adequately suppress the cosine ambiguity? In this chapter an answer is sought for the case of one arrival only. Simulated data for path radii of 20 metres and 90 metres are used to this end.

6.2 MUSIC and the Curved Array

Spatial MUSIC with forward spatial smoothing was used to investigate the ambiguity suppression for a single arrival. An arc was chosen as the antenna trajectory because it is the simplest non-linear trajectory to deal with. For data lying on an arc, smoothing is possible in only one direction. Recall that spatial smoothing uses a subarray which slides across the data points of the virtual array. In the case of a linear array, the forward and backward subarrays will always have the same orientation, with respect to the arriving signals, as they are slid across the array data. When the array data lies on an arc however, the orientation of the subarrays changes as they are moved across the array. Moreover, the orientation change for the forward subarray will be different from that of the backward subarray. As a result, the AOA determined, will differ between the two directions of smoothing.

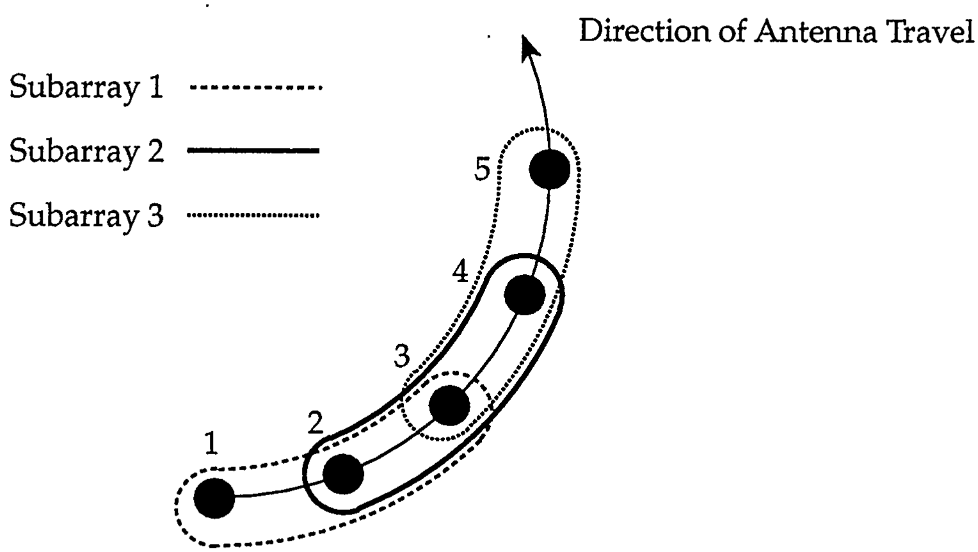

Figure 6-2 Smoothing Along a Curved Array

The orientation of the steering vector will determine which direction of smoothing gives the correct AOA and which gives a false AOA. Figure 6-2 illustrates the case of a virtual array consisting of five data points and subarrays of dimension three. Let the forward subarrays be defined as subarrays 1, 2, and 3 where the first element of subarray 1 is element 1, the first element of subarray 2 is element 2 and so on. If the steering vector of the array is defined as subarray 2, with element 2 as the first element, then forward smoothing as defined will give the proper AOA. The first backward subarray would be subarray 3 and its first element would be element 5. The second backward subarray would be subarray 2 and it would begin with element 4. The third backward subarray would be subarray 1 and its first element would be element 3. For the steering vector defined as above, backward smoothing will produce an AOA which is a reflection of the correct AOA with the line connecting elements 1 and 5 as the axis of reflection. Therefore, if both forward and backward smoothing are used, the correct AOA as well as an ambiguity (not to be confused with the cosine ambiguity) will result.

When smoothing with subarrays is used, the steering vector must be of the same dimension as the subarrays. Hence the question arises as to which of the subarrays is to be used as the steering vector. Obviously the middle subarray should best represent the entire array; hence the choice of subarray 2 as the steering vector for the case illustrated in Figure 6-2. In the case of an even number of subarrays, a middle subarray does not exist. In that case one of the two subarrays adjacent to the centre of the array must suffice.

For the simulation results to follow, forward smoothing as defined above for Figure 6-2, is used. The AOA is referenced as shown in Figure 6-3. The first point of the virtual array is assumed to be positioned at the origin of the vertical and horizontal axes and subsequent points follow on a circle in a counter-clockwise direction.

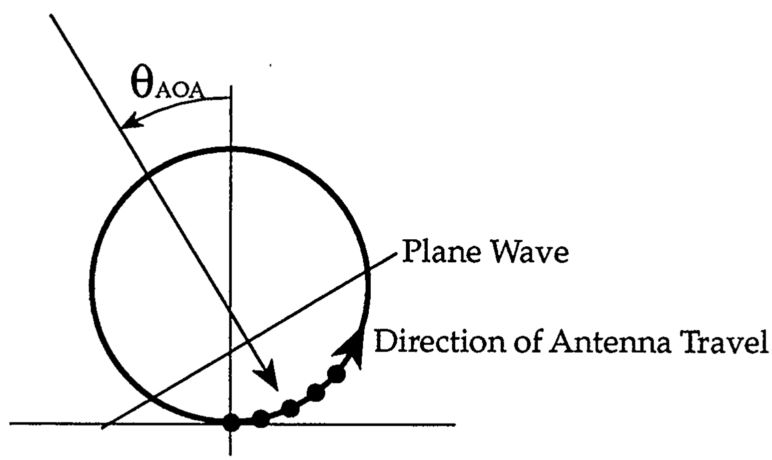

Figure 6-3 AOA Reference for Curved Arrays

6.3 Simulations for 20 metre Radius Arc

6.3.1 Introduction

The simulation data used in this section consisted of data points collected by an antenna travelling along an arc of radius 20 metres. The antenna velocity was 20 km/hr and the sampling rate 41.6 μs/sample. The frequency of transmission was 840 MHz. The signal environment consisted of only one signal with an AOA of 246.7°.

Two groups of simulations were conducted. In the first group, the distance between data points was fixed at 0.5 λ. The number of points (i.e. number of elements in the virtual array) was increased from 3 to 40 in steps of 1. The dimension of the subarrays was always made to be the integer value of 0.75 of the array dimension. Hence, the subarrays became longer and more numerous as the number of elements in the virtual array increased. For each array dimension, the magnitudes of each of the two peaks in the MUSIC spectrum, as well as AOA, were recorded. The same data was then generated for a straight line and the same results recorded for comparison.

For the second group of simulations, the number of virtual array elements was held fixed and the spacing between points was varied from 0.026 λ to 0.53 λ. For each interpoint spacing, the magnitude and AOA of each of the two peaks in the resulting MUSIC spectrum were recorded. Simulations in this manner were conducted for 17, 20, and 21 points.

Through these two groups of simulations, the extent of suppression of the ambiguity is measured as a function of the degrees of arc over which the data is collected. In addition, the nature of the suppression when more data points are added is compared to that achieved when the spacing between a fixed number of data points is increased.

6.3.2 Simulations for Fixed Interpoint Spacing of 0.5 λ

The peak magnitude results for the curved array are shown in Figure 6-4. The length of the virtual array in degrees of arc, is measured along the horizontal axis. The individual measurement points are shown. The first point corresponds to the array consisting of 3 elements whereas the last point corresponds to the array of 40 elements. In addition to the peak magnitudes, the magnitude of the spectrum floor is also shown.

As the virtual array becomes longer, it eventually becomes sufficiently two dimensional and the ambiguity altogether disappears. The peak corresponding to the correct AOA also decreases in magnitude; but since it starts at a higher value and does not fall off at the same initial rate, it eventually dominates the spectrum.

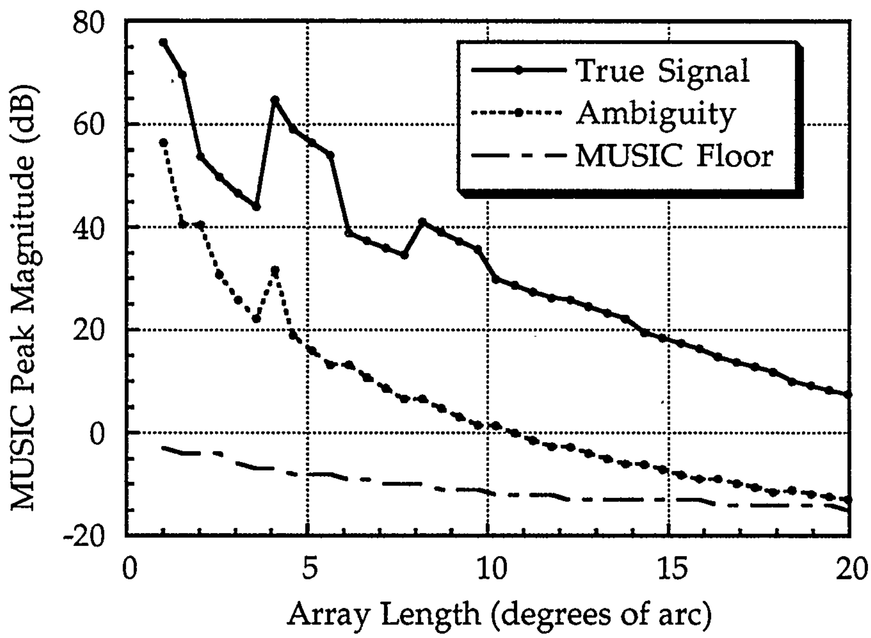

Figure 6-4 Peak Magnitude Comparison for Radius of 20 metres and Fixed Interpoint Spacing of 0.5 λ

A very consistent square wave type pattern is evident in the true signal curve. The levels are defined by groups of 4 points. This pattern is particular to the curved array since it does not arise for a linear array. Figure 6-5 illustrates the same information as Figure 6-4 but for linear arrays up to a total of 28 elements. Because the data falls on a straight line, the length of the array is measured in wavelengths and not degrees of arc. The square wave type pattern does not appear in Figure 6-5. Also clear is that the ambiguity does not disappear with increased array length.

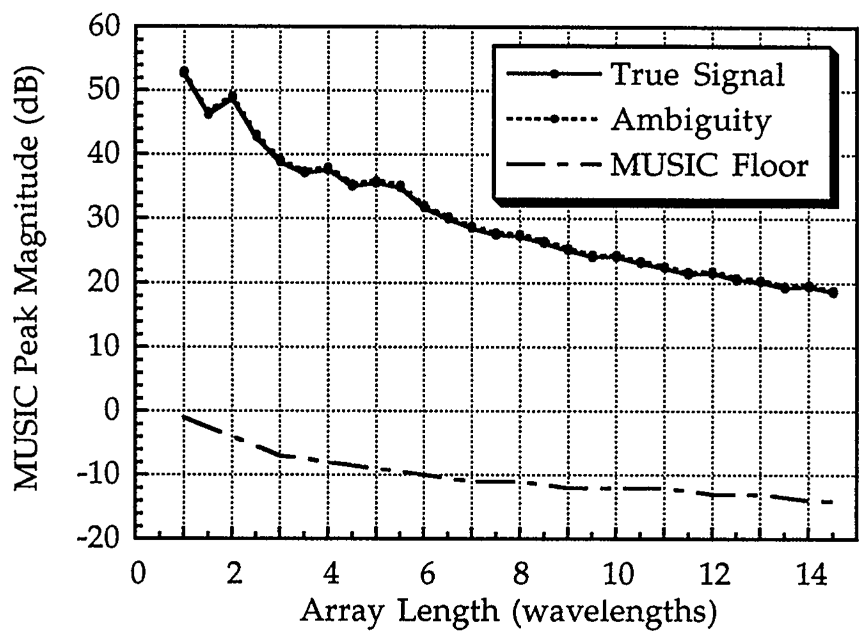

Figure 6-5 Peak Magnitude Comparison for Linear Arrays

The square wave pattern of Figure 6-4 is a function of whether the number of subarrays used is even or odd. For arrays with 3 through 10 elements, Table 6-1 contains information regarding the number of subarrays, their dimension, and the subarrays used as steering vectors. Also presented are the true signal peak magnitudes and the corresponding arrival angles.

Table 6-1 Subarray and Steering Vector Information for Various Arrays

No. of Array Elements	No. of Subarrays	Subarray Dimension	Steering Vector Subarray	AOA	Peak Magnitude (dB)
3	2	2	1	246.4°	78.5
4	2	3	1	246.4°	72.3
5	3	3	1 2	246.1° 246.7°	56.0 58.2
6	3	4	1 2	246.1° 246.7°	51.9 54.1
7	3	5	1 2	246.1° 246.7°	48.8 51.0
8	3	6	1 2	246.1° 246.7°	46.3 48.5
9	4	6	2	246.4°	69.2
10	4	7	2	246.4°	60.5

One sees from Table 6-1 that the peak magnitudes are dependent on the number of subarrays. Arrays with three or four elements, have two subarrays. When the number of elements increases to five, the number of subarrays increases to three and remains at three until there are nine array elements. The number of subarrays then becomes four. The sudden changes in the peak magnitudes as one changes from four elements to five, and from eight elements to nine are clearly seen.

Whether the number of subarrays is even or odd will also effect the choice of steering vector. When there is an odd number of subarrays, the middle one is the best choice for the steering vector. When the number of subarrays is even, one of the two subarrays slightly off centre must be chosen. The choice of steering vector will affect the degree to which it is orthogonal to the noise subspace and consequently will affect the peak magnitude. The square wave pattern in Figure 6-4 is the result.

Table 6-1 also shows the effect of the steering vector on the accuracy of the AOA estimate. Recall that the true AOA is 246.7°. When there is an odd number of subarrays and the middle one is chosen as the steering vector (subarray 2 for the case of three subarrays), the error in the AOA estimate is nil. If the middle subarray is not chosen or does not exist, as in the case of an even number of subarrays, error in the AOA estimate appears.

A comparison of the AOA estimate between the true peak and the ambiguity is presented in Figure 6-6. Except for the slight shift in AOA when the number of subarrays changes from even to odd and vice versa, the AOA estimate for the true peak is constant as the array length increases. The ambiguity AOA estimate, however, changes in a consistent manner as the array increases in length. This is to be expected since the axis of reflection, which can be thought of as the line connecting the first and last points of the array, changes orientation as additional array points are added.

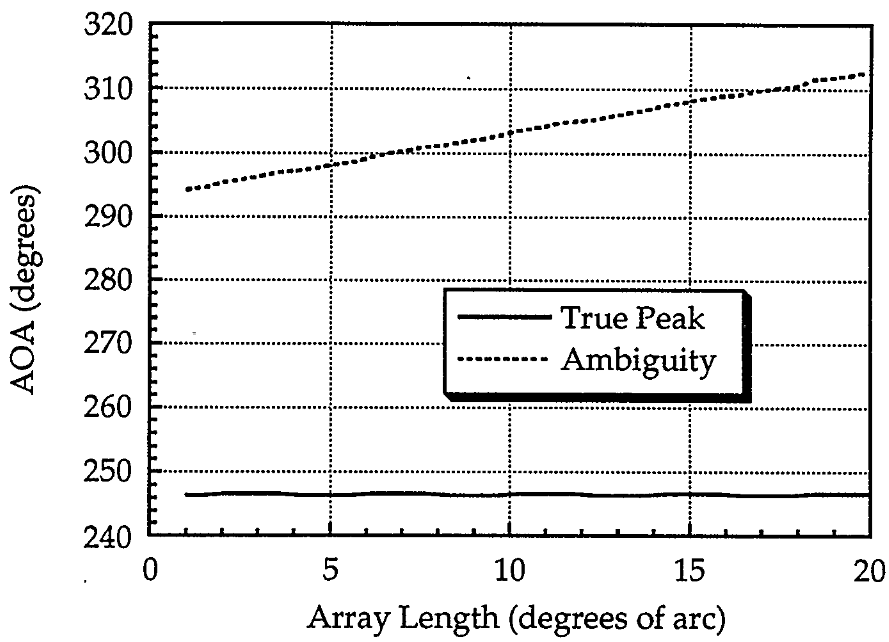

Figure 6-6 AOA Comparison for Fixed Interpoint Spacing of 0.5 λ

6.3.3 Simulations for Fixed Number of Array Elements

The purpose of this group of simulations is to determine the extent of ambiguity suppression as the array is made longer by increasing the spacing between data points. Figures 6-7 and 6-8 compare the true and ambiguous peak magnitudes for curved arrays of 20 and 21 data points respectively. In both cases the interpoint spacing is varied from 0.026 λ to 0.53 λ. The horizontal axes again measure the length of the arrays in degrees of arc.

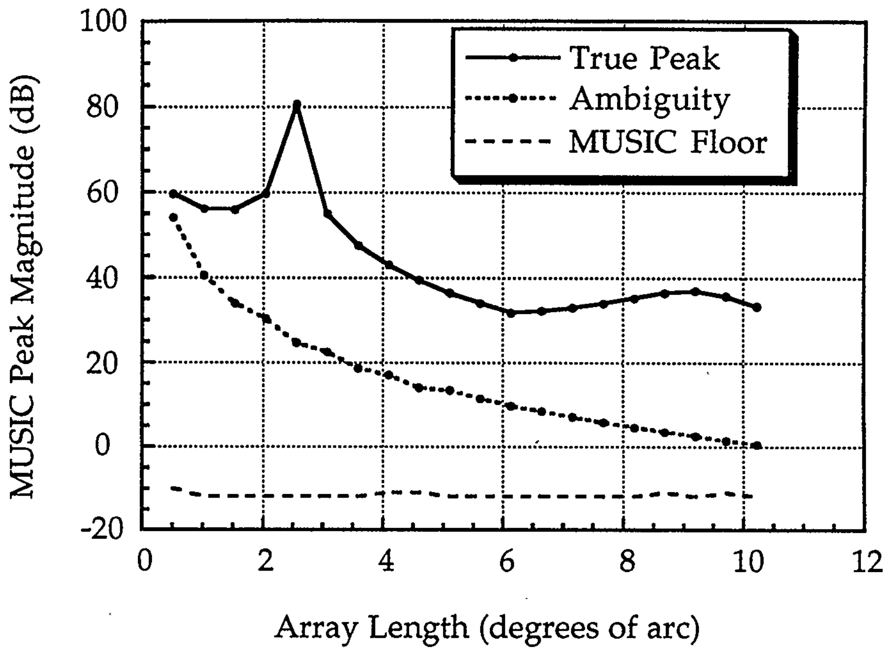

Figure 6-7 Peak Magnitude Comparison for a Curved Array of 20 Points and Radius 20 metres

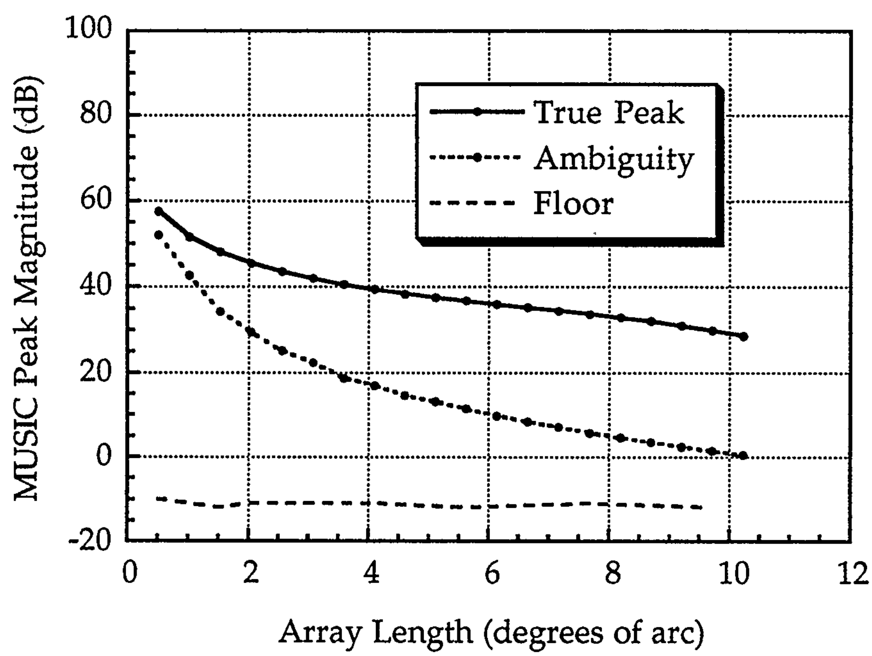

Figure 6-8 Peak Magnitude Comparison for a Curved Array of 21 Points and Radius 20 metres

Since for a fixed number of array points, the number of subarrays is also fixed, the square wave type pattern of Figure 6-4 does not appear in Figure 6-7 or 6-8. There is however, a difference in the true peak curves of Figures 6-7 and 6-8. The curve in Figure 6-7 exhibits a peak at 2.6° of arc length and does not drop consistently as the curve in Figure 6-8. Again this difference is attributable to the number of subarrays. For the array of 20 points the number of subarrays is 6; an even number. For the 21 point array, there are 7 subarrays. This is confirmed with an array of 17 points. It too has 6 subarrays and exhibits a peak at 2.6° of arc length. In the case of an even number of subarrays, one of the two innermost subarrays must be used to define the steering vector. How well this off-centre steering vector approximates the ideal one depends on the number of subarrays. In larger arrays, which contain many subarrays, the two innermost subarrays will better approximate the middle of the array than in the case of smaller arrays with fewer subarrays. In any case, the relationship between the noise subspace and the steering vector corresponding to the signal AOA, will not vary with array length in the smooth consistent manner of Figure 6-8.

The peak at 2.6° in Figure 6-7 allows for a 20 dB difference in peak magnitude between the true signal and the ambiguity, with an array arc length as small as 1.3°. In the case of Figure 6-8, the same separation is obtained with an array arc length of 3.3°. These results are comparable to those for Figure 6-4. In that case the separation was already 20 dB with the smallest array of 3 points, which has an arc length of approximately 1°. However, in Figure 6-4, the separation tends to decrease with increased array length whereas in Figure 6-7 and 6-8 the separation tends to increase.

A comparison between the AOA estimates of the true signal and the ambiguity, for an array of 20 points and increasing interpoint spacing, results in a figure virtually identical to Figure 6-6.

6.4 Simulations for 90 metre Radius Arc

6.4.1 Introduction

In order to observe the effect of arc radius, simulations were conducted for data lying on an arc of radius 90 metres. The vehicle velocity was 50 km/hr, the sampling rate 41.6 μs/sample, and the frequency of transmission 840 MHz. Again the data was simulated for only one arrival and its AOA was 130°. Measurements of the true signal peak magnitude and ambiguity peak magnitude were made as the array size was increased by adding additional points to the array. The interpoint spacing was fixed at 0.4 λ.

6.4.2 Simulations for Fixed Interpoint Spacing of 0.4 λ

Simulations were conducted for array sizes of 3 elements up to 50 elements. Once again the dimension of the subarrays was 0.75 of the total number of elements in the array. Figure 6-9 shows the results of the simulations.

The true signal curve of Figure 6-9 also exhibits the square wave pattern seen in Figure 6-4. Again there are two different levels in the curve corresponding to odd and even numbers of subarrays. However, in contrast to the 20 metre radius case, here, arrays with an even number of subarrays have lower peak magnitudes than do arrays with an odd number of subarrays. This is not a function of the

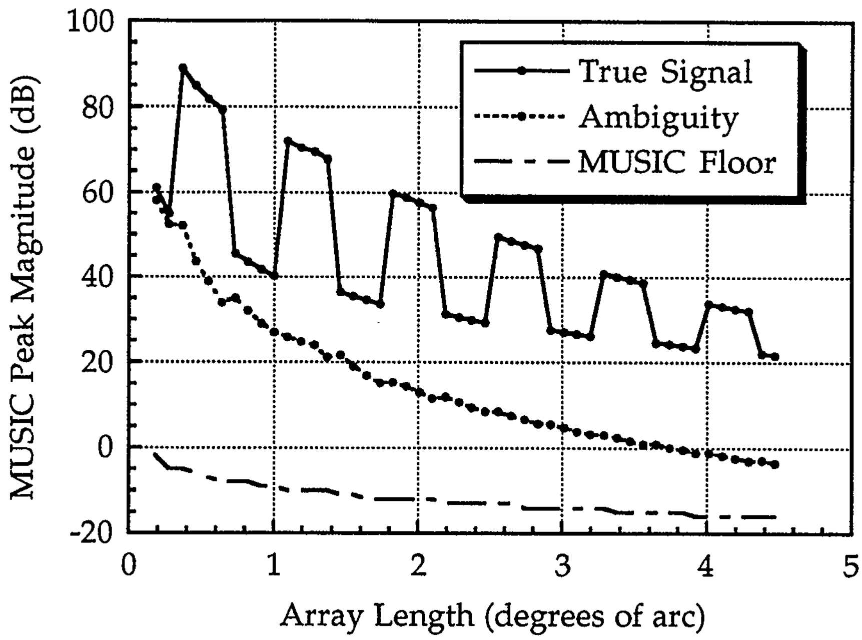

Figure 6-9 Peak Magnitude Comparison for Radius of 90 metres and Fixed Interpoint Spacing of 0.4 λ

interpoint spacing since the same pattern was observed for spacings other than 0.4 λ. From these results it is impossible to predict whether an even number of subarrays or an odd number of subarrays will give the larger peak magnitudes.

In Figure 6.9, a worst case separation of 20 dB is achieved for arrays greater in length than 2.3°; the worst case being the lower level of the decaying square wave pattern. For the 20 metre radius case of Figure 6-4, a worst case separation of 20 dB is obtained with array lengths of approximately 3° of arc length. Of course 3° of arc length corresponds to a longer travel distance for the 90 metre radius arc than for the 20 metre radius arc.

6.5 Conclusions

In conclusion, it may be stated that an odd number of subarrays will give the best AOA accuracy since there is a middle subarray available for definition of the steering vector. In addition, it was found that the AOA estimate for the true signal remained constant with varying array length whereas the AOA estimate of the ambiguity changed in a consistent manner.

The ambiguous peak was totally suppressed with an array length of 20 arc degrees for a 20 metre radius arc and with an array length of approximately 10 arc degrees for a 90 metre arc. To obtain a 20 dB separation between the true signal and the ambiguity required arrays of length 1 to 3 degrees of arc. This depended on the combination of interpoint spacing and the number of array elements (Figures 6-7 and 6-8), as well as whether the number of subarrays was even or odd (Figures 6-4 and 6-9).

CHAPTER 7
CONCLUSIONS

7.1 The MUSIC Algorithm

The MUSIC algorithm, as implemented, was found to be very capable of accurately estimating the AOA or frequency of individual arrivals using simulated data. In particular, the resolution possible with MUSIC exceeds that of more traditional methods, such as the Fourier transform, by many times. The temporal form of MUSIC is especially impressive since it is able to resolve signals with very few, and closely spaced points. It was found that methods such as temporal or spatial smoothing, as well as array motion, effectively decorrelate coherent arrivals. Therefore, MUSIC is not excluded from use in a coherent signal environment such as the multipath radio channel.

7.2 Three Techniques for the Three Cluster Environment

Three different techniques for estimating the LOS AOA in a three cluster environment were developed and tested. Simulations were conducted in order to evaluate the performance of each technique in the areas of identifying the LOS cluster and estimating the corresponding AOA. Technique #3, which bases the identification process on cluster amplitude, was found to have the best success rate in terms of LOS cluster identification. This assumed, however, that simulations for which the cosine ambiguity could not be eliminated are not included in the success rate calculation. If all simulations were included, the success rate of technique #3 fell below 50% and technique #1 became superior in this regard.

Both techniques #1 and #2 estimated the LOS cluster AOA with an accuracy in the order of 5.5°. The accuracy of technique #3 was only 12°. Both techniques #1 and #3 attempted to estimate an average AOA for the LOS cluster. Spatial MUSIC as used in technique #1, was far more successful in this venture than temporal MUSIC in technique #3. A significant reason for this was found to be the cosine ambiguity elimination process used with temporal MUSIC. For individually resolved arrivals, the ambiguity of each arrival was very successfully eliminated using the phase data of the two antennas. For resolved clusters however, signal matching between the two antennas as well as phase errors, severely reduced the effectiveness of the process. However, even in those cases where the cluster ambiguity could be eliminated, temporal MUSIC was not able to determine an average cluster AOA with the same degree of accuracy as spatial MUSIC in technique #1.

In comparing the three techniques, other factors were considered as well. Technique #1 required a large number of solutions in order to obtain an AOA estimate of reasonable accuracy. Technique #2, although requiring more array points, required the fewest number of solutions. Since technique #3 used temporal MUSIC, it required fewer data points to make an AOA estimate than either of techniques #1 or #2. In addition, the required spacing between data points for temporal MUSIC is much less than for spatial MUSIC. As a result, the AOA update rate would be much higher for a system using technique #3.

Complexity in terms of eigenanalysis and additional processing are also considerations. Due to the relative sizes of the correlation matrices, technique #2 required many more computations to perform the eigenanalysis than techniques

#1 or #3. Technique #1 required a large amount of additional processing albeit relatively simple. In addition to the generation of three MUSIC spectra, the peak tracking algorithm and a windowing technique which groups the solutions such that a LOS majority can be found, are required. Technique #2 also requires such a windowing process as well as a sophisticated algorithm for clustering the peaks in the MUSIC spectrum. Elimination of the cosine ambiguity with signal phase incurred a relatively large processing burden on technique #3.

Considering all of the above factors, technique #2 was found to be the best overall method for estimating the LOS AOA in a three cluster environment.

An AOA accuracy of 5.6° (as achieved by technique #2) should be adequate to obtain a positional accuracy of 100 metres or less in the microcellular radio environment. If 2 transmitters situated on a circle are 90° apart, and the mobile receiver is located at the centre of the circle, the minimum and maximum positional errors for AOA errors of 5° will be 11% and 14% of the circle radius respectively. Although a transmitter separation of 90° is perhaps the ideal case, using AOA estimates from multiple transmitters will give a positional accuracy better than 10% of the circle radius. Therefore, if the above mentioned circles correspond to cells with radii in the order of 1 km, the desired positional accuracy will be obtained with AOA estimates made by techniques #1 and #2.

7.3 The Single Cluster Environment

Temporal MUSIC was used to estimate the AOA of the LOS component in a single cluster. Because individual arrivals were resolved and not clusters, the resulting accuracy was better than 4°. For the same reason, elimination of the angular ambiguity was not a significant problem as it was with technique #3. This was reflected in the very high success rate.

A more accurate system could use a two stage process. In the first stage, technique #1 would be used to estimate the AOA of the LOS cluster. The antenna data is then spatially filtered in order to remove the non-LOS clusters. Temporal MUSIC could then resolve for individual arrivals, identify the LOS arrival by its amplitude, and provide a more accurate estimate of its AOA.

7.4 Curved Arrays

Simulations confirmed that the angular ambiguity can be eliminated by using a two dimensional virtual array created with one antenna moving along a non-linear path. A circular arc was chosen as the antenna path since it is the least complicated non-linear trajectory for which to define an array steering vector. The definition of the steering vector was determined to have a significant effect on AOA accuracy. When spatial smoothing using subarrays was used, the best accuracy was achieved when the number of subarrays was odd and the middle subarray was used to define the steering vector.

To suppress the MUSIC peak of the angular ambiguity by 20 dB with respect to the true signal peak, the virtual array length was required to be at least 1° to 3° of arc. This was affected by the number of array elements, their spacing, and whether the number of subarrays was even or odd. Arc radius, however, did not have a significant effect on the required array length in degrees of arc.

As the arrays were made longer, the ambiguous peak eventually dropped to the spectrum floor and disappeared. This occurred at 20 degrees of arc for a radius of 20 metres and approximately 10 degrees of arc for a 90 metre radius. When converted to circumference, the antenna had to travel twice as far on the arc of 90 metre radius in order to generate a sufficiently two dimensional array and eliminate the ambiguity.

7.5 Recommendations for Further Work

It is possible that technique #1 may be improved in the areas of efficient data use and perhaps accuracy by employing spatial smoothing in the form of subarrays rather than temporal smoothing. Computational complexity would be reduced since fewer data points in total would be required. The additional smoothing obtained with forward and backward subarrays may improve decorrelation and accuracy.

As seen in the single cluster results, temporal MUSIC has the potential to make very accurate AOA estimates for individual arrivals. Technique #3 would be the preferred technique if this same degree of accuracy could be achieved for clusters, perhaps by improving the ambiguity elimination process. This may be possible by using SVD to calculate the signal phases.

Throughout this thesis the direction of transmission has been from the cell site to the mobile. Consideration should be given to the reverse case. The availability of arrays at the cell sites as well as no signal processing required at the mobile may make the reverse case more desirable.

The value of this thesis work is dependent upon the actual AOA nature of the outdoor urban radio channel. Tests to determine whether arrivals actually do cluster, as well as what the configuration of these clusters might be, would determine whether the techniques presented would provide the accuracy required. In addition, the percentage of time a LOS path exists between a cell site and a mobile should be determined.

REFERENCES

[1] E. N. Skomal, *Automatic Vehicle Locating Systems.* New York: Van Norstrand Reinhold, chapters 1, 3, 4, and 5, 1981.

[2] G. Lachapelle and B. Townsend, "En Route Coverage Validation and Calibration of LORAN-C with GPS," *Proceedings of GPS -90*, Institute of Navigation, Washington, D.C., Sept. 19-21, 1990, pp. 199-205.

[3] L. Tetley and D. Calcutt, *Electronic Aids to Navigation.* London: Edward Arnold, chapter 4, 1986.

[4] G. Lachapelle, B. Townsend, M.E. Cannon, and D. Halayko, "GPS and LORAN-C Signal Availabilty for Urban Navigation: A Case Study," *presented at IEEE International Conference on Vehicle Navigation and Information Systems*, Dearborn, MI, October 20-23, 1991.

[5] G. Lachapelle and B. Townsend, "A Comparative Analysis of LORAN-C and GPS for Land Vehicle Navigation in Canada," *Canadian Journal of Aeronautics*, Canadian Aeronautics and Space Institute, Ottawa, Vol. 36, No. 1, pp. 29-35, Mar. 1990.

[6] P. A. Goud, "A Spread Spectrum Radiolocation Technique and its Application to Cellular Radio," M.Sc. Thesis, Dept. of Elec. & Comp. Engineering, University of Calgary, Aug. 1991.

[7] G. Lachapelle and B. Townsend, "Hybrid GPS/LORAN-C: An Effective Radionavigation System Mix for Vehicular Navigation in Mountainous Areas," *Proceedings of 48th Annual Meeting*, Institute of Navigation, Washington, D.C., June 28 - July 1, 1992, pp. 3-11.

[8] G. Lachapelle, B. Townsend, and D. Halayko, "Analysis of GPS LORAN-C Performance for Land Vehicle Navigation in the Canadian Rockies," *IEEE AES Magazine*, pp. 24-28, May 1992.

[9] G. J. Sonnenberg, *Radar and Electronic Navigation*, Sixth Edition. London: Butterworths, chapter 7, 1988.

[10] C. B. Harris, "Prototype For a Land Based Automatic Vehicle Location and Navigation System," M.Engg. Thesis, Dept. of Surveying Engineering, University of Calgary, Dec. 1989.

[11] G. C. Hurst, "QUIKTRAK: A Unique New AVL System," *Proceedings of the First Vehicle Navigation and Information Systems Conference*, 1989, pp. A60-A62.

[12] G. Morrison, M. Fattouche, and D. Tholl, "Parametric Modelling and Spectral Estimation of Indoor Radio Propagation Data," *Proceedings of Wireless '92*, Calgary, Alberta, July 8-10, 1992.

[13] Electronic Industries Association, "Cellular System Dual-Mode Mobile Station - Base Station Compatibility Standard," EIA Project Number 2215, *IS-54*, May 1990.

[14] D. S. Hill, "Multiple signal DF using superresolution: a practical assessment," *IEE Electronics & Communication Engineering Journal*, Vol. 2, No. 6, pp. 221-232, Dec. 1990.

[15] S. Sakagami, S. Aoyama, K. Kuboi, S. Shirota, and A. Akeyama, "Vehicle Position Estimates by Multibeam Antennas in Multipath Environments," *IEEE Trans. on Veh. Tech.*, Vol. 41, No. 1, pp. 63-68, Feb. 1992.

[16] J. B-Y. Tsui, *Digital Microwave Receivers Theory and Concepts*. Norwood, MA: Artech House, Inc., chapters 3, 4, and 5, 1989.

[17] S. Haykin, J. H. Justice, N. L. Owsley, J. L. Yen, and A. C. Kak, *Array Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, chapter 4, 1985.

[18] S. Haykin, *Adaptive Filter Theory*, Second Edition. Englewood Cliffs, NJ: Prentice-Hall, chapter 12, 1991.

[19] D. W. Tufts and R. Kumaresan, "Estimation of Frequencies of Multiple Sinusoids: Making Linear Prediction Perform Like Maximum Likelihood," *Proc. IEEE*, Vol. 70, No. 9, pp. 975-989, Sep. 1982.

[20] B. D. Rao and K. V. S. Hari, "Performance Analysis of Root MUSIC," *IEEE Trans. on Acoustics, Speech, and Sig. Proc.*, Vol. 37, No. 12, pp. 1939-1949, Dec. 1989.

[21] S. V. Schell, R. A. Calabretta, W. A. Gardner, and B. G. Agee, "Cyclic MUSIC Algorithms for Signal-Selective Direction Esitmation," *Proceedings of ICASSP 89, IEEE Conference on Acoustics, Speech, and Sig. Proc.,* Vol. 4, Glascow, Scotland, May 23-26, 1989.

[22] G. Xu and T. Kailath, "Direction-of-Arrival Estimation via Exploitation of Cyclostationarity - A Combination of Temporal and Spatial Processing," *IEEE Trans. on Signal Processing,* Vol. 40, No. 7, pp. 1775-1786, Jul. 1992.

[23] S. K. Oh and C. K. Un, "Improved MUSIC Algorithm for High-Resolution Array Processing," *Electronics Letters*, Vol. 25, No. 22, pp. 1523-1525, Oct. 26, 1989.

[24] D. H. Brandwood, "Noise-space Projection: MUSIC without eigenvectors," *IEE Proceedings*, Vol. 134, Pt. H, No. 3, pp. 303-309, Jun. 1987.

[25] M. Kaveh and A. J. Barabell, "The Statistical Performance of the MUSIC and the Minimum-Norm Algorithms in Resolving Plane Waves in Noise," *IEEE Trans. on Acoustics, Speech, and Sig. Proc.*, Vol. ASSP-34, No. 2, pp. 331-341, Apr. 1986.

[26] R. O. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," *IEEE Trans. on Ant. and Propag.*, Vol. AP-34, No. 3, pp. 276-280, Mar. 1986.

[27] F. Haber and M. Zoltowski, "Spatial Spectrum Estimation in a Coherent Signal Environment Using an Array in Motion," *IEEE Trans. on Ant. and Propag.*, Vol. AP-34, No. 3, pp. 301-310, Mar. 1986.

[28] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipies in C.* Cambridge: Cambridge University Press, chapter 11, 1988.

[29] S. U. Pillai, B. H. Kwon, "Forward/Backward Spatial Smoothing Techniques for Coherent Signal Identification," *IEEE Trans. on Acoustics, Speech, and Sig. Proc.*, Vol. 37, No. 1, pp. 8-14, Jan. 1989.

[30] T-J. Shan, M. Wax, T. Kailath, "On Spatial Smoothing for Direction-of-Arrival Estimation of Coherent Signals," *IEEE Trans. on Acoustics, Speech, and Sig. Proc.*, Vol. ASSP-33, No. 4, pp. 806-811, Aug. 1985.

[31] R. T. Williams, S. Prasad, A. K. Mahalanabis, and L. H. Sibul, "An Improved Spatial Smoothing Technique for Bearing Estimation in a Multipath Environment," *IEEE Trans. on Acoustics, Speech, and Sig. Proc.*, Vol. 36, No. 4, pp. 425-432, Apr. 1988.

[32] M. Fattouche, E. B. Olasz, K. E. Scott, and H. Zaghloul, "Modified Simulation of the Urban Radio Propagation Channel (SURP)," *submitted to IEEE Trans. on Veh. Tech.*, 1993.

[33] C. Zhou, F. Haber, and Q. Shi, "On Spatial Filtering for Angle-of-Arrival Estimation in a Scattering Environment by Eigendecomposition-Based Methods," *IEEE Trans. on Ant. and Propag.*, Vol. 38, No. 6, pp. 931-934, Jun. 1990.