NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, tests publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

Canada

COMPUTING NUMERICAL SOLUTIONS TO THE ELECTROMAGNETIC TWO-DIMENSIONAL SCALAR INVERSE SCATTERING PROBLEM

by

Andrew James Kozlowski

A THESIS SUBMITTED TO THE
SCHOOL OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Ottawa-Carleton Institute for Electrical Engineering
Department of Electrical Engineering
Faculty of Engineering
University of Ottawa
Ottawa, Ontario K1N 6N5, CANADA

ISBN 0-315-46791-6

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

Acknowledgements

I would like to express my gratitude to my supervisors Dr. S. S. Stuchly and Dr. M. M. Ney for their guidance during the course of this research, and for their valuable comments concerning this thesis. I would also like to express my thanks to the technical staff in the Dept. of Electrical Engineering (Mr. G. Hartsgrove, Mr. M. Master, Mr. B. Carraro, and Mr. S. Symons) and in the Physics workshop for their help toward this work. The discussions with Mr. A. Delecki are greatly appreciated. The Computing Center at the National Research Council (CANADA) generously supplied a number of FORTRAN subroutines needed for the numerical work. The help supplied by the User Services group at the University of Ottawa Computing Center in obtaining some numerical code is also appreciated. Finally, I would like to thank my friend the VaxStation II/RC which struggled with me through the computations.

This thesis was typeset using LaTeX in 11pt report format on a Macintosh II computer. Drawings were done a Macintosh computer. Graphs and two-dimensional plots were done on a VAX station II/RC workstation. Printing was done using QMS Lasergraphix, DEC LN03 Plus, and APPLE LaserWriter printers.

In order to obtain a satisfactory solution, one should first ask the proper question.

Digital computers have one tremendous advantage over analytical methods. It is that if you do not make too many programming errors and the machine does not break down, it will always produce numbers.

D.S. Jones, Proc. IEE, vol. 121, no. 7, Jul 1972, p. 574.

Moving *fast* is not going somewhere.

Abstract

This thesis examines the reconstruction of low-loss two-dimensional dielectric objects with electrical parameters $(\kappa_D(\underline{x}), \mu_0)$ which are imbedded in a low loss ambient medium (κ_A, μ_0). Electromagnetic energy with an incident wave vector $\underline{k}$ at an angle θ with respect to the +y-axis and time-harmonic $(e^{j\omega t})$ dependence interacts with the object, producing a scattered field $E_s(\underline{x};\theta)$ which is measured along some observation contour O outside the object. Using Maxwell's equations and Green's identities, the effective current $J_{eff}(\underline{x}) \triangleq (k_A^2 - k_D^2(\underline{x}))E(\underline{x};\theta) = \omega^2\mu_0\epsilon_0(\kappa_D(\underline{x}) - \kappa_A)E(\underline{x};\theta)$ is related to the scattered field $E_s(\underline{x};\theta)$ through a Fredholm integral equation of the first kind, which in operator form can be expressed as $KJ_{eff} = E_s$. This represent a well-known ill-posed (i.e. unstable) problem. The difficulties involved in computing solutions to this problem are examined, and a two-step reconstruction process is proposed. First, a set of observation points is chosen which reduces the instability of the problem after discretization by the method of projections. This instability is measured by the condition number (κ_2) of the resulting linear system. If the condition number is sufficiently low with respect to machine precision, then the second step is to use bounds on the real and imaginary parts of the complex relative permittivity $\kappa_D(\underline{x})$ as *physically meaningful* constraints in a computational procedure. This procedure is based on the optimization problem of minimizing the difference between the reconstructed and measured scattered fields. The conjugate gradient method permits the constraints to be incorporated with some success.

List of Symbols and Abbreviations

a	radius of a circular cell [m]
ANA	automated network analyzer
B	magnetic flux density [Wb/m^2]
C	set of complex numbers
C^n	set of $n-$ tuples of complex numbers
$C^s(\Omega)$	set of functions with continuous derivatives $\leq s$ in a region Ω
d	edge size of a square cell [m]
D	domain of the object
D	electric displacement [coul/m^2]
dp	decimal places (of precision)
DP	direct problem
e_i	unit vector with a 1 in the $i-$th position : $(0, 0, \ldots, 1, \ldots, 0)$
e_{NMS}	normalized mean-square error
e_{RMS}	average root mean-square error (per cell)
E	total E-field $= E_i + E_s$ [V/m]
E_i	incident E-field [V/m]
E_s	scattered E-field [V/m]
f	(i) frequency [Hz]
	(ii) source function
F	the set $E-$ field data needed for reconstruction
g	data function
$G()$	Green's function
H	magnetic intensity [A/m]
$H_n^{(1)}()$	$= J_n() + jY_n()$ Hankel function of the first kind of order n

$H_n^{(2)}()$	$= J_n() - jY_n()$ Hankel function of the second kind of order n
IP	inverse problem
J	volumetric electric current [A/m^2]
J_m	volumetric magnetic current [V/m^2]
$J_n()$	Bessel function of the first kind of order n
$J_{eff}(\underline{x})$	$= (k_A^2 - k_D^2(\underline{x}))E(\underline{x})$ [A/m]
J_p	polarization current [A/m^2]
k	wavenumber [m^{-1}]
$\underline{k}$	wavevector [m^{-1}]
$k_D(\underline{x})$	$\sqrt{\omega^2\mu\epsilon(\underline{x}) - j\omega\mu\sigma(\underline{x})}$ wavenumber in the object, D [m^{-1}]
k_A	$\sqrt{\omega^2\mu\epsilon_A - j\omega\mu\sigma_A}$ wavenumber in the object, D [m^{-1}]
K	a linear operator
K	surface current [A/m]
K_B	integral scalar Helmholtz operator (Bojarski formulation)
K_H	integral scalar Helmholtz operator
K_R	electromagnetic scattering equation
K_{LS}	operator giving the least-squares solution of K
L	the underflow limit.
$L^n(\Omega)$	the set of functions $f(\underline{x})$ such that $\int_\Omega \|f(\underline{x})\|^n d\underline{x} < \infty$
L_H	differential scalar Helmholtz operator
LSMN	least squares minimum norm
M	number of observation points taken
MGS	modified Gram-Schmidt orthogonalization
MoP	Method of Projections
N	number of cells in the discretized object, D
n/c	not computed
NMS	normalized mean square (error)
O	the set of position (x,y) data needed for reconstruction
O	set of points on the observation contour
$O(\beta_n)$	order of, i.e. $\alpha_n = O(\beta_n)$ if $\|\alpha_n\| \leq K\|\beta_n\|$ for some constant K
p	number of passes
P	(i) number of projections (ii) a projection operator

R	spatial resolution [m]
r_0	nominal radius of a circular observation contour [m]
R^n	set of n-tuples of real numbers
RMS	root mean square (error)
S	the bounding surface of an object
SIE	Surface Integral Equation (method)
SVD	Singular Value Decomposition
t	(i) time [s]
	(ii) precision (number of places used to represent a number)
TE	transverse electric
TM	transverse magnetic
$\underline{u}$	$= (u, v)$ spatial frequency coordinates [/m]
U	the overflow limit
$U(r_0, \delta r)$	a uniform probability distribution from $r - \delta r$ to $r + \delta r$
VIE	Volume Integral Equation (method)
v_p	velocity of propagation
$\underline{x}$	position vector outside $D = (x, y)$ [m]
$\underline{x}'$	position vector inside $D = (x', y')$ [m]
$Y_n()$	Bessel function of the second kind of order n
z	(i) a spatial coordinate
	(ii) a complex number

α	attenuation constant [Np/m] or [dB/cm]
β	(i) base (radix) for (machine) computation
	(ii) phase constant
ϵ	permittivity [F/m]
ε_g	error in the data g, $\|g - g_\epsilon\| \leq \varepsilon_g$, where g_ϵ is the measured value
$\kappa(A)$	$= \|A\|\|A^\dagger\|$ condition number of an operator or matrix
$\kappa(\underline{x})$	$= \kappa'(\underline{x}) - j\kappa''(\underline{x})$ complex relative permittivity

λ_A	wavelength in the ambient medium [m]
λ_0	wavelength in free space [m]
μ	(i) permeability [H/m]
	(ii) measurement average
$\mu(\underline{x})$	x-ray attenuation [Np/m]
ρ	(i) distance [m]
	(ii) volumetric electric charge density [coul/m^3]
ρ_s	surface charge density [coul/m^2]
σ	(i) conductivity [S/m]
	(ii) measurement standard deviation
ω	angular frequency [rad/s]

A^*	adjoint of A
$\overline{A}$	complex conjugate of A
$\widetilde{A}$	Fourier transform of A
$\widehat{A}$	estimate of A
A^{-1}	classical inverse of A
$A^\dagger$	generalized inverse of A
A^T	transpose of A
A^H	(complex) conjugate transpose of A
$[a]$	vector a
$[A]$	matrix A
$[A]^{(n)}$	order n matrix approximation of a continuous operator
$\|z\|$	magnitude of a complex number, z
$/z$	phase of a complex number, z
$\|\|A\|\|_p^m$	The Holder $p-$norm of A raised to the $m-$th power
$\int$	integral
$CPV\int$	Cauchy principal integral with singularity removed
$f^{(i)}$	$i-$th derivative of f

∇	del operator
$\triangleq$	equal to by definition
$\partial\Omega$	the boundary of the domain Ω

Contents

List of Tables

List of Figures

Chapter 1

INTRODUCTION

The inverse scattering problem is one of object reconstruction, i.e. to find the spatial dependence (and in a more general setting, the temporal dependence as well) of a property of a scattering object. In turn, object reconstruction is a specific type of imaging in which the exact values of some physical parameter are sought, rather than just an "image". The IEEE dictionary [1977] defines an *image* as

> "a spatial distribution of a physical property such as radiation,...,or reflectivity, mapped from another distribution of either the same or another physical property."

but it is more comprehensive to look at an image as a property which can be displayed with the values of a property modified in some way, perhaps scaled by an unknown factor or by filtering out noise. The fundamental problem still remains – to understand the underlying mechanics of the reconstruction process in order that the limitations of the procedure be understood.

An example of an important imaging system which was made practical by computer technology, is the CAT (computer aided tomography) scanner. These imaging systems find the X-ray attenuation $\mu(\underline{x})$ of an object by measuring the X-ray intensity, I, with a set of detectors.

Electromagnetic radiation of frequency less than that of X-rays has also been used for imaging. In 1886, Heinrich Hertz discovered that electromagnetic waves can be reflected by metallic and dielectric objects. The microwave frequencies ($f \approx 1$GHz) have been used since the 1920's - 1930's when RADAR was developed for the detection of ships and aircraft.

During World War II, the use of RADAR was expanded rapidly to include target location and the technology was also used for navigation. During the 1950's, synthetic aperture radar (SAR) was developed for the imaging of the complex reflectivity of terrain in order to acquire reconnaissance while flying over unfriendly regions.

1.1 The Problem

The propagation of electromagnetic waves is often considered in the mathematically simpler scalar (rather than vector) form. In this thesis, a study is made of the following scalar electromagnetic inverse problem. Consider a low-loss two-dimensional (i.e. cylindrical) object D of finite support whose cross-section lies in the x-y plane (see Figure 1). The object has electrical parameters $(\kappa_D(\underline{x}) \in C, \mu_0)$ and is imbedded in a low-loss isotropic homogeneous ambient medium $(\kappa_A \in C, \mu_0)$. All media are assumed to be non-magnetic, i.e. $\mu = \mu_0$. For computational simplicity, consider a transverse magnetic (TM) field incident on the two-dimensional object at an angle θ with respect to the y-axis (see Figure 2.)

Let u represent either the total $E-$field $E(\underline{x};\theta)$ or the scattered $E-$field $E_s(\underline{x};\theta)$, with $e^{j\omega t}$ time dependence. This leads to the *scalar Helmholtz equation*

$$\nabla^2 u(\underline{x}) + k_A^2 u(\underline{x}) = \begin{cases} J_{eff}(\underline{x};\theta) & \underline{x} \in D \\ 0 & \underline{x} \notin D \end{cases} \tag{1}$$

where $J_{eff}(\underline{x};\theta)$ is the effective current in the object which produces the scattered field, $E_s(\underline{x};\theta)$, D is the domain of the object, and

$$k_A \triangleq \omega\sqrt{\mu\epsilon_0\kappa_A} \in C \tag{2}$$

is the wavenumber of the ambient medium. u must also satisfy the Sommerfeld radiation condition in two-dimensions (see Jones [1964, pp. 56-57])

$$\lim_{r\to\infty} \sqrt{r}u \text{ is bounded} \tag{3}$$

$$\lim_{r\to\infty} \sqrt{r}(\frac{\partial u}{\partial n} + jk_A u) = 0$$

where $r = |\underline{x}| = \sqrt{x^2+y^2}$ and $\frac{\partial u}{\partial n} = \underline{n}\cdot\nabla$ with $\underline{n}$ the unit outward normal to S (see Williams [1980, p. 285]). Essentially these conditions state that as $r \to \infty$, the Poynting vector $\mathbf{E}\times\mathbf{H}$ must be directed radially outward and decreasing as r^{-2}, while the radial components E_r and H_r must be decreasing faster than r^{-1}. In other words, the solution must represent an

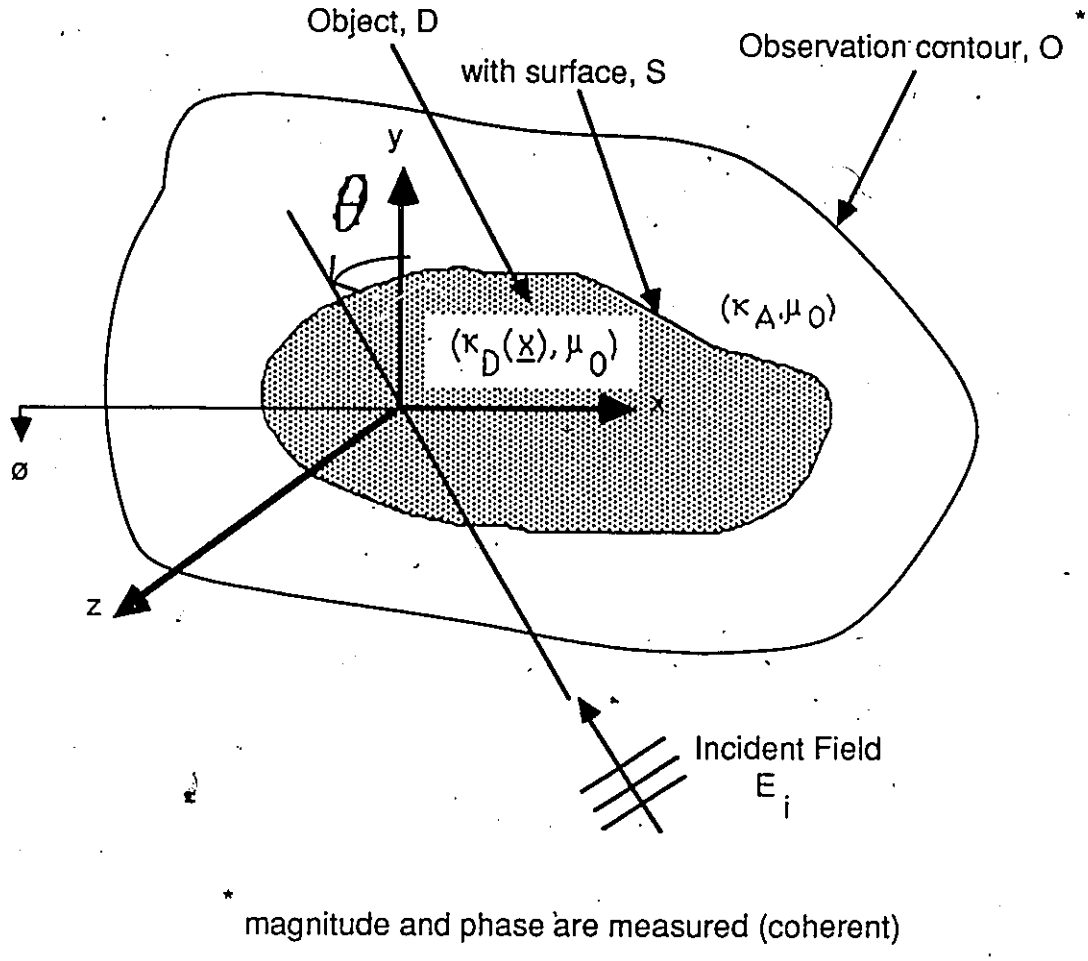


Figure 1: Reconstruction Model for a Cylindrical Object D
The object is a low-loss dielectric $(\kappa_D(\underline{x}), \mu_0)$ and is imbedded in a low-loss ambient medium (κ_A, μ_0). The $E-$field is measured on observation contour O.

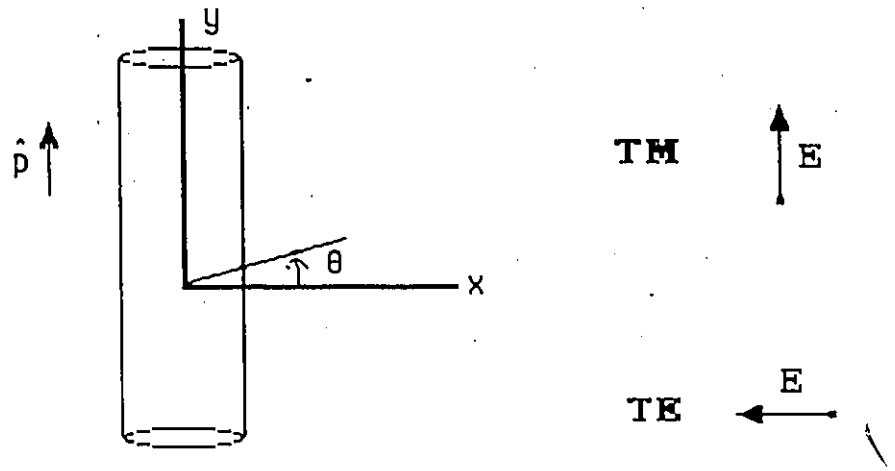


Figure 2: TM- and TE- polarized plane waves
Plane waves are incident on a cylindrical object. In a TM- (TE-) wave, there is no H (E) component along the direction of propagation, $\hat{p}$. (Based on Mentzer [1955, p.57].)

outward propagating wave at infinity and not an inward propagating one. This additional condition leads to a unique solution for the problem.

This partial differential equation invites two problems. First, consider the problem of determining the scattered field $E_s(\underline{x};\theta)$ on a surface O from the complex relative permittivity of the object $\kappa_D(\underline{x})$, i.e.

$$\mathrm{DP}(\kappa_D) : \kappa_D \rightarrow E_s \tag{4}$$

This is called the *direct problem* for the complex relative permittivity κ_D, and is formulated from elementary principles. Physically, the energy in the incident field interacts with the object and produces a scattered field. Direct problems generally are mathematically stable.

Next, consider the problem of determining the complex relative permittivity of the object from its scattered field, i.e.

$$\mathrm{IP}(\kappa_D) : E_s \rightarrow \kappa_D \tag{5}$$

where the scattered field is measured along some (not necessarily closed) *observation contour*, O. This *inverse problem* has application in a number of imaging problems (e.g. biomedical, air pockets in pyramids, detection of plastic weapons on personnel). The objective of this thesis is to investigate this *inverse scattering problem* by using the two-dimensional scalar model previously described. The integral equation which describes the

relation between the known scattered field and the unknown complex relative permittivity is a *Fredholm integral equation of the first kind*, whose solution leads to an ill-posed problem. Specifically, this thesis deals with the following problems which have not been considered by previous researchers.

- a fundamental understanding of the instability of the problem and especially of the computational problems involved
- object reconstruction in a low-loss ambient medium, e.g. water, which can provide a good "match" to objects of biomedical interest
- to avoid the assumption of *weakly scattering objects* which does not accurately model objects with large contrasts in permittivity values, and leads to a resolution limitation of $\lambda_A/2$
- to construct a new algorithm to solve highly ill-conditioned problems by imposing *physical* rather than *mathematical* constraints

(The inverse problem for perfectly conducting objects was considered with some success by Imbriale and Mittra [1970].)

Inverse problems are often based on inverting the mathematical expression for the direct problem and tend to be mathematically unstable, i.e. sensitive to the input data. Such problems have been intensively studied since the turn of the century. In 1905-1907, Herglotz and Wiechart (see Romanov [1974]) solved the first important inverse problem. They considered the following physical setting: A region D is bounded by a surface S, Waves are generated on S, and their travel times are recorded on S. They solved the problem of finding the propagation velocities v_p of the disturbances inside D under the assumptions of a spherical earth model with propagation velocity increasing monotonically with depth. For the first time, this permitted inferences to be made about the Earth's internal structure from seismological data. Electromagnetic inverse problems have been widely studied. A good review of these is given in the IEEE Antennas and Propagation Society Special Issue on "Inverse Methods in Electromagnetics" [1981].

Let us now consider the inverse problem IP(κ_D) more precisely. A low-loss dielectric material can be characterized by its *complex permittivity* $\epsilon \triangleq \epsilon' - j\epsilon''$ where $\epsilon' > 0$ is the *dielectric constant* of the material, and $\epsilon'' = \frac{\sigma}{\omega} \geq 0$ its *dielectric loss factor* which includes

all dissipative effects. We define the *complex relative permittivity* as

$$\kappa \triangleq \frac{\epsilon}{\epsilon_0} = \kappa' - j\kappa'', \quad \kappa' > 0, \quad \kappa'' \geq 0 \tag{6}$$

where ϵ_0 is the permittivity of free space.

1.2 The Application

As an application of this inverse problem, this thesis uses numerical values appropriate for the reconstruction needed for biomedical imaging at the nominal operating condition of

$$f = 3 \text{ GHz}, \text{ T} = 25\text{C} \tag{7}$$

This point is a good trade-off among three factors (see Lin [1985] and Guerquin-Kern [1985]):

1. resolution
2. attenuation
3. data acquisition time.

A simple, but useful, criterion for determining when two monochromatic spectral components of equal intensity are resolved was proposed by Rayleigh (see Born and Wolf [1983, p. 333].) Such components are "just resolved" when the principal intensity maximum of one coincides with the first intensity minimum of the other. Using this classical criterion originally stated for incoherent stellar sources, the spatial resolution R for a circular aperture of radius a is $0.61\lambda_A/a$ radians. Hence this criterion says that smaller wavelengths give better spatial resolutions. For coherent sources, however, where both magnitude and phase are important (such as the systems considered here), no figure for R has been determined (see Goodman [1968, p. 130+] for a discussion.) It should be noted that a number of authors attempting to solve the IP(κ_D) problem under the assumption of weakly scattering objects have obtained theoretical resolution limits of the order of $\lambda_A/2$ (Wolf [1969, p. 155], Kom [1981, p. 20+], and Pichot et al. [1985, p. 420].)

Increasing frequency (smaller wavelength), however, leads to increased attenuation in the materials considered. This leads to weaker received signals and hence increased receiver data acquisition time (to average out the noise).

Material	T [C]	f [GHz]	Measured κ $\approx$ 3GHz	Max. cell size (1) [mm]	α [Np/m]	β [°/mm]	Reference
Free space	all	all	1.0 - j0.0(t)	20.0	0.000	3.6	
Acrylic	20	3.00	2.6 - j0.0	12.4	0.000	5.8	von Hippel [1954]
Fat	22	2.450	4.0 - j0.8(p)	10.0	0.011	7.7	Stuchly&Stuchly, [1980, p.24]
Muscle	22	2.450	50.0 - j16.5(p)	2.8	0.070	25.8	Stuchly &Stuchly [1980,p.24]
Sea Water (0.1M)	?	3.000	75.5 - j18.1	2.3	0.065	31.5	Shen&Kong [1983,p.55]
Distilled Water	25	3.000	76.5 - j11.3 (t)	2.3	0.040	31.6	Kraszewski [1986]
Untreated Water	22	3.150	77.5 - j13.0	2.3	0.046	31.8	(2)

Legend:

(t) = theoretical value

(p) = value for phantom material

(1) = Richmond [1965] $0.2\lambda_0/\sqrt{\kappa'} \approx 20/\sqrt{\kappa'}$ [mm] at 3 GHz

(2) = measured by J. Dadkhah at the University of Ottawa with program PERMRUN on the HP8410B network analyzer

Table 1: Complex Relative Permittivities of various materials of interest
The nominal operating condition is f = 3.000 GHz, T = 25C. α and β are computed from Lin [1985].

Table 1 lists the complex relative permittivities at our nominal operating condition. We shall be working with a distilled water ambient, hence

$$\kappa_A = 76.5 - j11.3 \tag{8}$$

and the wavelength is

$$\lambda_A = \frac{\lambda_0}{\sqrt{\kappa'_A}} = 11.43\text{mm} \tag{9}$$

Referring to Table 1, appropriate bounds on the real and imaginary parts of the complex permittivities are

$$\begin{aligned} l_R = 1 \quad &\leq \kappa'_D(\underline{x}) \leq \quad 76.5 = u_R \\ l_I = -20 \quad &\leq \kappa''_D(\underline{x}) \leq \quad 0 = u_I \end{aligned} \tag{10}$$

1.3 An Overview of the Thesis

In this thesis, the relevant equation for the inverse problem is derived from Maxwell's equations, its mathematical properties are discussed, and an algorithm is developed to compute solutions to that problem. Because in the frequency range considered, the objects are of the order of a wavelength, the object scatters incident energy in all directions. *Thus, neither the straight-line propagation assumption of ray optics, nor the weakly-scattering assumption of the Born approximation are valid* (see Figure 3). The resulting formulation leads to the problem of solving an integral equation which is ill-posed (i.e. unstable). Upon discretization of this integral equation by the method of projections, the resulting linear system of equations becomes rapidly ill-conditioned as the number of "cells" increases. Hence, it is difficult to solve this problem even in the case of exact input data given the realistic constraint of finite machine precision and "reasonable" computation time.

In Chapter 2, two formulations for the two-dimensional reconstruction problem with monochromatic ($e^{j\omega t}$) incident electromagnetic radiation are derived using Maxwell's Equations. The relationship between the scattered field E_s and the unknown effective current J_{eff} from which the complex relative permittivity κ_D is computed leads to an ill-posed problem. Such problems, in general, *do not have unique solutions if the input data E_s is not known exactly.*

In Chapter 3, key concepts from functional analysis and linear algebra are reviewed. Different types of "solutions", discretizations of the operator suitable for numerical computation, and a quantification of the stability of the resulting linear system are discussed.

In Chapter 4, the method of projections (a special case is the method of moments) is applied to the most promising formulation described in Chapter 2 and its properties are discussed.

With computer simulations, the problem of reducing the instability of the computed results is discussed in Chapter 5. Object reconstructions with the algorithms described in Chapter 4 are presented. The reconstructions demonstrate the problems in using the method of projections in solving inverse problems.

A summary of the results of this thesis is presented in Chapter 6, and the existing problems are highlighted.

In Appendix A, the analytical solution to $E-$field scattering by a circular low-loss dielectric cylinder is given. Since the scattered field for this object is known exactly, it

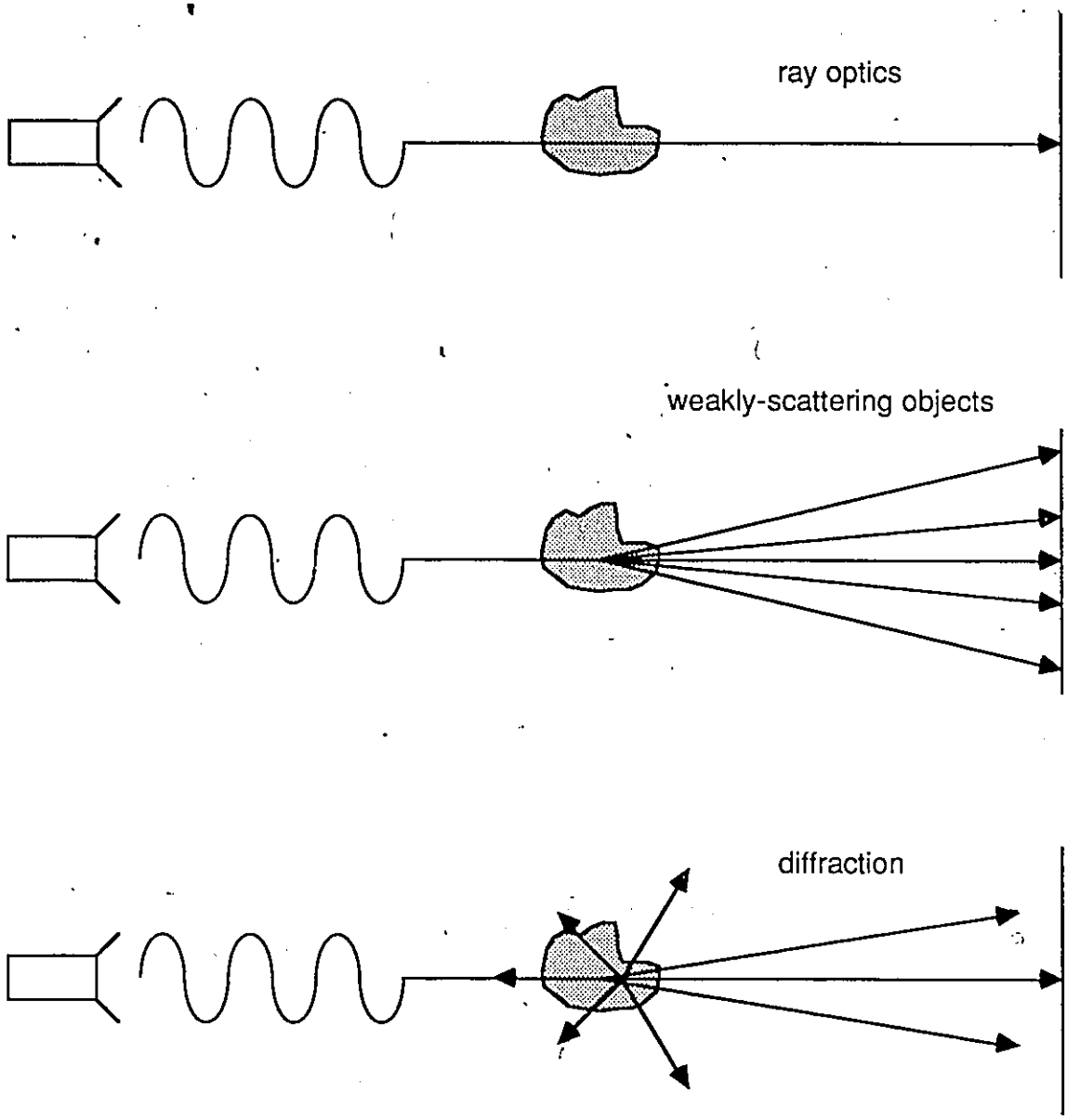


Figure 3: Models for scattering by objects

provides a useful reference for experimental calibration of a measurement system.

A measurement system (MSCAN) which was built to aid in the construction of a numerical reconstruction algorithm is discussed in Appendix B. MSCAN measures the $E-$field scattered by objects in a tank filled with water. The system operates at a frequency of 3.156 GIIz.

In Appendix C, a summary of the computer software used in the reconstructions is given.

In Appendix D, some of the VAX/FORTRAN-77 code used to reconstruct objects is listed.

Mathematical preliminaries which are necessary to discuss the inherent difficulties in the problem are given in Appendix E.

In this thesis, an attempt was made to maintain notation consistent with the literature. However, it is inevitable that the notation is inconsistent at points or that the same symbol will have several meanings which are sensitive to the context. Hopefully, these problems have been minimized.

Chapter 2

SCATTERING and the TWO-DIMENSIONAL SCALAR HELMHOLTZ EQUATION

The direct scattering problem DP(κ_D) described in Chapter 1 was first treated in simplified form by Descartes using ray optics. Subsequently, the problem was formulated as a boundary value problem using the classical method of separation of variables. Analytically exact solutions have been obtained for simple scattering objects whose boundary surface coincides with the eleven coordinate systems to which this method can be applied (see Morse and Feshbach [1953, pp. 655-664]). Among these geometries applied to *lossless objects*, i. e. ($\kappa_D''(\underline{x}) = 0$), were circular cylinders (Rayleigh in 1881), spheres (Mie in 1908) and elliptical cylinders (Yeh in 1957).

2.1 Formulation

For the general direct scattering problem, there are two principal formulations: (1) the Volume Integral Equation formulation (VIE) and (2) the Surface Integral Formulation (SIE).

2.1.1 The Volume Integral Equation Method

The volume integral equation method is also called the polarization source formulation method (see Bates [1972]) for a discussion). It models the scattering object as the source of a field which travels undisturbed through space, and is applicable to *material objects*, i.e. $\sigma_D < \infty$, only. This approach leads to a set of coupled vectorial equations which are very difficult to solve. However, the assumption that the incident $E-$field is linearly polarized and that polarization transformation (e.g. depolarization , rotation) are negligible is often made, leading to a single scalar equation.

With this approach, scattering problems are formulated in terms of integral equations and solved by numerical methods. Using the VIE formulation, Richmond [1965] considered lossless two-dimensional dielectric objects of arbitrary cross-section and computed the scattered field by using the method of moments (see Chapter 4) with pulse expansion and point matching. He first considered incident TM-polarized plane waves [1965], and later extended the results to incident TE-polarized plane waves [1966]. These formulations essentially represent the source inside the cylinder by an equivalent 'polarization' current

$$J_p = j\omega\epsilon_0(\kappa_D - \kappa_A)E \tag{11}$$

2.1.2 The Surface Integral Equation Method

Later, Wu and Tsai [1974] considered a surface integral equation formulation for the problem of a homogeneous low-loss dielectric cylinder of arbitrary cross-section in a homogeneous ambient medium. They derived a system of two coupled integral equations for the surface field ($E(\underline{x})$ on S) using Maxwell's equations, Green's theorems, and boundary conditions (continuous tangential E and H on the surface S):

$$\begin{aligned} E_i(\underline{x}) &= E(\underline{x})/2 + (CPV)\int_S [G_D(\underline{x},\underline{x}')\frac{\partial E(\underline{x}')}{\partial n'} - E(\underline{x}')\frac{\partial G_D(\underline{x},\underline{x}')}{\partial n'}]dS' \\ 0 &= E(\underline{x})/2 - (CPV)\int_S [G(\underline{x},\underline{x}')\frac{\partial E(\underline{x}')}{\partial n'} - E(\underline{x}')\frac{\partial G(\underline{x},\underline{x}')}{\partial n'}]dS' \end{aligned} \tag{12}$$

where $(CPV)\int$ denotes the Cauchy principal value of the integral

$$(CPV)\int_a^b f(x)dx \triangleq \lim_{\epsilon\to+0}[\int_a^{c-\epsilon} f(x)dx + \int_{c+\epsilon}^b f(x)dx] \tag{13}$$

where $G(\underline{x},\underline{x}')$ and $G_D(\underline{x},\underline{x}')$ are the Green's functions outside and inside the object, respectively. The Cauchy principal integrals are required to integrate around the singularities.

2.1.3 Comparison

The SIE formulation requires fewer unknowns than is necessary for the VIE formulation, and is applicable to both material ($\sigma_D < \infty$) and perfectly conducting ($\sigma_D = \infty$) objects. However, finding the Green's function $G_D(\underline{x}, \underline{x}')$ inside a homogeneous object is a very difficult analytic problem. Further, for high loss objects the boundary conditions become more complicated (see, for example, Senior [1981]).

2.2 The Scalar Helmholtz Equation

Let the electromagnetic fields be represented by the electric $\mathbf{E} \in C$ and magnetic $\mathbf{H} \in C$ field vectors. At every point in whose neighbourhood the physical properties of the medium are continuous, i.e. in open regions, these fields obey Maxwell's Equations

$$\begin{aligned} \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{H} &= \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \\ \nabla \cdot \mathbf{D} &= \rho \\ \nabla \cdot \mathbf{B} &= 0 \end{aligned} \tag{14}$$

In addition, the constitutive relations $\mathbf{D} = \epsilon\mathbf{E}$, $\mathbf{B} = \mu\mathbf{H}$, $\mathbf{J} = \sigma\mathbf{E}$, and ϵ, μ, and $\sigma \in C$ are needed to allow unique determination of the field vectors from known current and charge distributions. If the fields are exceptionally strong, then the right hand sides must be supplemented by higher powers of the field vectors (Born and Wolf [1983, p. 3].) If the fields are time-harmonic, i.e. have ($e^{j\omega t}$) time dependence, then Maxwell's Equations have the form

$$\begin{aligned} \nabla \times \mathbf{E} &= -j\omega\mathbf{B} \\ \nabla \times \mathbf{H} &= j\omega\mathbf{D} + \mathbf{J} \\ \nabla \cdot \mathbf{D} &= \rho \\ \nabla \cdot \mathbf{B} &= 0 \end{aligned}$$

At the boundaries, the following conditions hold (the brackets [ ] denote discontinuities in the field components on crossing the boundary)

$$[\hat{n} \times \mathbf{E}] = \mathbf{0} \tag{15}$$

$$
\begin{aligned}
[\hat{n} \times \mathbf{H}] &= \mathbf{K} \\
[\hat{n} \cdot \mathbf{D}] &= \rho_s \\
[\hat{n} \cdot \mathbf{B}] &= 0
\end{aligned}
$$

where $\mathbf{K}$ is the surface current distribution, and ρ_s is the surface charge distribution. When we speak of a *low-loss material*, we mean one where the surface current $\mathbf{K}$ can be considered negligible, i.e.

$$[\hat{n} \times \mathbf{H}] = \mathbf{K} = 0 \tag{16}$$

The continuity of charge relation is also required, which for time harmonic field dependence is

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t} = -j\omega\rho \tag{17}$$

2.2.1 Spatial Domain - Differential Formulation

The following development of the time-reduced wave propagation ("Helmholtz") equation is based on van Bladel [1961], Livesay and Chen [1974], Chen [1981] and Kom [1981].

Let V_A be a low-loss isotropic, homogeneous ambient region with constitutive parameters $(\epsilon_A, \mu_A, \sigma_A)$ and fields (E_A, H_A). V_A contains a penetratable isotropic inhomogeneous object of volume V_D with constitutive parameters $(\epsilon_D(\underline{x}), \mu_D(\underline{x}), \sigma(\underline{x}))$ and a set of fields (E_D, H_D). The main assumptions are

- time-harmonic ($e^{j\omega t}$) field variation, thus avoiding assumptions about the dispersiveness of the media
- non-magnetic media ($\mu_A = \mu_D = \mu_0$)
- J_{eff} and E_s are $\in C^2(V_D)$

(A function is of class $C^m(\Omega)$ if all its partial derivatives through order m are continuous in some region Ω.) Additional assumptions which are later needed are

- The $E-$fields are linearly polarized in the $+z$ direction.
- There is no polarization transformation.
- There are no non-linear effects, e.g. there is no frequency conversion between the incident field and the scattered field (as occurs in laser action based on pumping).

The *incident fields* (E_A, H_A) interact with the object and give rise to *scattered fields* (E_s, H_s). Using the principle of superposition, the scattered field is defined by (see Thiele [1973, p.8])

$$E_s = E_D - E_A \tag{18}$$
$$H_s = H_D - H_A$$

Using Maxwell's curl equations

$$\begin{aligned}
\nabla \times H_s \quad = \nabla \times (H_D - H_A) \quad &= (\sigma_D + j\omega\epsilon_D)E_D - (\sigma_A + j\omega\epsilon_A)E_A \\
&= (\sigma_D + j\omega\epsilon_D)E_D - (\sigma_A + j\omega\epsilon_A)(E_D - E_s) \\
&= ((\sigma_D - \sigma_A)E_D - (\sigma_A + j\omega\epsilon_A))(E_D - E_s) \\
\nabla \times E_s \qquad = -\tfrac{\partial \mu H_s}{\partial t} \quad &= -j\omega(\mu_D - \mu_A)H_D - j\omega\mu_A H_s \\
&= -j\omega\mu_A H_s
\end{aligned}$$

Identifying the electric and magnetic currents

$$J_s \triangleq ((\sigma_D - \sigma_A) + j\omega(\epsilon_D - \epsilon_A))E_D = j\omega\epsilon_0(\kappa_D - \kappa_A)E_D$$

$$K_s \triangleq -j\omega(\mu_D - \mu_A)H_D = 0$$

and proceeding in standard way to find the wave equation

$$\begin{aligned}
\nabla \times \nabla \times E_s \quad &= \quad -j\omega\mu_0 \nabla \times H_s \\
&= \quad -j\omega\mu_0(J_s + j\omega(\epsilon_A - j\sigma_A/\omega)E_s)
\end{aligned}$$

Hence,

$$-\nabla^2 E_s + \nabla\nabla \cdot E_s = -j\omega\mu_0 J_s + k_A^2 E_s \tag{19}$$

where $k_A^2 = \omega^2\mu_0(\epsilon_A - j\sigma_A/\omega)$ is the square of the wavenumber in the ambient medium.

Outside the object, by applying the continuity of current equation

$$\nabla \cdot E_s = \frac{1}{\epsilon_A}\nabla \cdot D_s = \rho_s = -(\nabla \cdot J_s)/(j\omega\epsilon_A) \tag{20}$$

Substituting eq. (20) into eq. (19), and rearranging we have the differential form of the Helmholtz equation

$$\nabla^2 E_s + k_A^2 E_s = J_{eff} \tag{21}$$

where

$$J_{eff} \triangleq j\omega\mu_0(J_s + (\nabla\nabla \cdot J_s)/k_A^2) \tag{22}$$

is the *effective current* of the object. If the ambient medium is free space, i.e. $\epsilon_A = \epsilon_0$, then J_{eff} is called the *polarization current*, J_p.

Further, assume that the incident field E_i satisfies

$$\nabla^2 E_i + k_A^2 E_i = 0 \tag{23}$$

e.g. a plane wave, then the total field is given by adding eq. (21) and eq. (23)

$$\nabla^2 E + k_A^2 E = J_{eff} \tag{24}$$

where $E(\underline{x};\theta) = E_i(\underline{x};\theta) + E_s(\underline{x};\theta)$ represents the total field. Hence, the Helmholtz equation

$$L_H u \triangleq \nabla^2 u + k_A^2 u = J_{eff} \tag{25}$$

is satisfied by $u = E_s$ as well as any $u = E_s + E_i$ where E_i satisfies eq. (23), i.e. E_i is in the null space of the operator L_H.

2.2.2 Spatial Domain - Integral Formulation

Formulations based on differential operators such as eq.(25) which as unbounded operators can be troublesome, and are best avoided. Hence, we turn to integral formulations of eq. (21) which can be derived using the Green's function method. As the integral formulation and the differential formulation eq. (21) will be inverses, the steps must be carefully examined. First, some definitions.

Let

$$Lu = f(x) \text{ in } \Omega \tag{26}$$

and

$$Bu = 0 \text{ on } \overline{\Omega}$$

where L is a linear partial differential equation and B is a set of linear boundary-initial conditions such that for each continuous f the problem (1) has a unique solution. We call the union of Ω and its limit points the closure of Ω and denote it as $\overline{\Omega}$.

SINGULARITY SOLUTION. A *singularity solution* σ for a linear differential operator L with a pole at ξ is a solution of the equation

$$L\sigma = \delta(x - \xi)$$

where ξ is considered as a parameter.

GREEN'S FUNCTION. $G(\underline{x}, \xi)$ is the *Green's function* for eq. (26) if its unique solution is given by

$$u(\underline{x}) = \int_\Omega G(\underline{x}; \xi) f(\xi) d\Omega$$

Essentially, this forms the inverse (L^{-1}) for the operator (L). When no boundaries are present, this function is called the free-space Green's function.

Refer again to Figure 1. The incident field $E_i(\underline{x}; \theta)$ propagates at an angle θ with respect to the +y-axis. Because $E_i(\underline{x}; \theta)$ is linearly polarized in the +z-direction, we have

$$\nabla(\nabla \cdot J_s) = 0 \tag{27}$$

and hence $\rho_s = 0$ (no scattering charge). Further, we now have a scalar equation with

$$J_{eff}(\underline{x}; \theta) = j\omega\mu_0(j\omega\epsilon_0(\kappa_D - \kappa_A)E_D) = k_0^2(\kappa_A - \kappa_D)E_D = (k_A^2 - k_D^2)E_D \tag{28}$$

where

$$F(\underline{x}) \triangleq (k_A^2 - k_D^2(\underline{x}))E(\underline{x}) \tag{29}$$

is sometimes called the *scattering potential* (see Wolf [1969]).

GREEN'S SECOND INTEGRAL IDENTITY. Let u, v be scalar functions that are $C^2(\Omega)$ and $C^1(\overline{\Omega})$. Then (see, for example, John [1982, p.79+])

$$\int_\Omega (u\nabla^2 v - v\nabla^2 u) d\Omega = \int_{\partial\Omega} (u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n}) dS$$

Let $E_s(\underline{x}; \theta)$ be a scalar function satisfying the scalar Helmholtz equation (25), and ψ be the singularity solution of the same equation, i.e.

$$\nabla^2\psi(\underline{x}) + k_A^2\psi(\underline{x}) = \delta(\underline{x} - \underline{x}')$$

Applying Green's second integral identity to $E_s(\underline{x};\theta)$ and $\psi(\underline{x})$ to a region $\Omega \supset D$ we have

$$\begin{aligned}\int_{\partial\Omega}(\psi\frac{\partial E_s}{\partial n} - E_s\frac{\partial\psi}{\partial n})dS &= \int_\Omega(\psi\nabla^2 E_s - E_s\nabla^2\psi)d\Omega \\ &= \int_D(\psi(\nabla^2 E_s + k_A^2 E_s) - E_s(\nabla^2\psi + k_A^2\psi))d\Omega \\ &= \int_D(\psi J_{eff} - E_s\delta(\underline{x}-\underline{x}'))d\Omega\end{aligned}$$

Hence,

$$K_B Jeff \triangleq \int_D \psi J_{eff}d\Omega - \int_C(\psi\frac{\partial E_s}{\partial n} - E_s\frac{\partial\psi}{\partial n})dS = E_s \tag{30}$$

This is the inverse scattering problem formulation of Bojarski [1981]. It requires knowledge of the scattered field E_s and its normal derivative $\frac{\partial E_s}{\partial n}$ over a closed contour C. Essentially, it is a boundary value problem. Computationally, approximating a derivative such as $\frac{\partial E_s}{\partial n}$ by finite differences is very difficult when the data is noisy (see Conte and de Boor [1980, pp. 294-303]), making this formulation sensitive to measurement error. Hence, we do not pursue this approach further.

By taking the region to be all of space $\Omega = R^2$, the Sommerfeld radiation condition eq. (3) can be applied to the scattered field as well as the stronger condition

$$\lim_{r\to\infty} E_s = 0 \tag{31}$$

which guarantees that E_s will have the required regularity properties at infinity (see Bleistein and Handelsman [1986, p.57]).

Assuming radial symmetry, because the contour integrals will vanish, the Green's function for the problem can easily be found for the two-dimensional problem with position vector $\underline{x} \triangleq (x,y)$. (see, for example, Jones [1964, section 1.34]).

$$\psi(\underline{x},\underline{x}') = \frac{j}{4}H_0^{(2)}(k_A|\underline{x}-\underline{x}'|) \tag{32}$$

Hence, we have the result obtained by Richmond [1965], and applied to the inverse problem considered here (for a lossless ambient medium) by Ney [1983].

$$K_H J_{eff} \triangleq \int_D \frac{j}{4}H_0^{(2)}(k_A|\underline{x}-\underline{x}'|)J_{eff}(\underline{x}';\theta)d\underline{x}' = E_s(\underline{x};\theta), \quad \underline{x}\in R^2\setminus D \tag{33}$$

where we have called K_H the inverse operator of L_H (eq. (25)). Müller [1969, pp. 287-288] has shown that (1) there exists a solution of this equation, and (2) its solution represents the

solution of the given scattering problem. The vectorial differential form of the Helmholtz equation can also be put into integral form with somewhat more labor (see van Bladel [1961] for details).

The real and imaginary parts of the Hankel function $H_0^{(2)}(z) = x + jy$ are plotted in Figure 4 for the cases of free-space and distilled water propagation at the nominal operating condition, (7). Note how much faster the kernel decays in water ambient compared with propagation in free space.

Once eq. (11) has been solved for $J_{eff}(\underline{x};\theta)$, the complex relative permittivity $\kappa_D(\underline{x})$ can be computed from knowledge of the incident field inside the object $E_i(\underline{x}), \underline{x} \in D$ by using the relation

$$\begin{aligned} J_{eff}(\underline{x};\theta) &= (k_A^2 - k_D^2(\underline{x};\theta))E(\underline{x};\theta) \\ &= (k_A^2 - k_D^2(\underline{x};\theta))(E_i(\underline{x};\theta) + E_s(\underline{x};\theta)) \\ &= \omega^2\mu_0\epsilon_0(\kappa_A - \kappa_D(\underline{x}))(E_i(\underline{x};\theta) - E_s(\underline{x};\theta)) \qquad (34) \\ &= k_0^2(\kappa_A - \kappa_D(\underline{x}))(E_i(\underline{x};\theta) + E_s(\underline{x};\theta)) \qquad (35) \end{aligned}$$

where $k_0 = \frac{\omega}{c} = \omega\sqrt{\mu_0\epsilon_0}$ is the wavenumber in free space. Rearranging and noting that inside D eq. (33) also applies,

$$\begin{aligned} \kappa_D(\underline{x}) &= \kappa_A - \frac{J_{eff}(\underline{x};\theta)}{k_0^2 E(\underline{x};\theta)} \\ &= \kappa_A - \frac{J_{eff}(\underline{x};\theta)}{k_0^2(E_i(\underline{x};\theta) + \int_D \frac{j}{4}H_0^{(2)}(k_A|\underline{x} - \underline{x}'|)J_{eff}(\underline{x};\theta)d\underline{x}')} \qquad (36) \end{aligned}$$

Thus, finding the complex relative permittivity $\kappa_D(\underline{x})$ from the effective current $J_{eff}(\underline{x})$ is straightforward, and therefore this thesis concentrates on the solving for the latter.

2.2.3 Fourier Domain - Differential Formulation

Before discussing the applicability of the Fourier Domain techniques to our problem, we review some Fourier Theory.

FOURIER TRANSFORM. The two-dimensional *Fourier Transform* $\tilde{f}$ of a function f is defined by the relation

$$\tilde{f}(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-j2\pi(ux+vy)}dx\,dy \qquad (37)$$

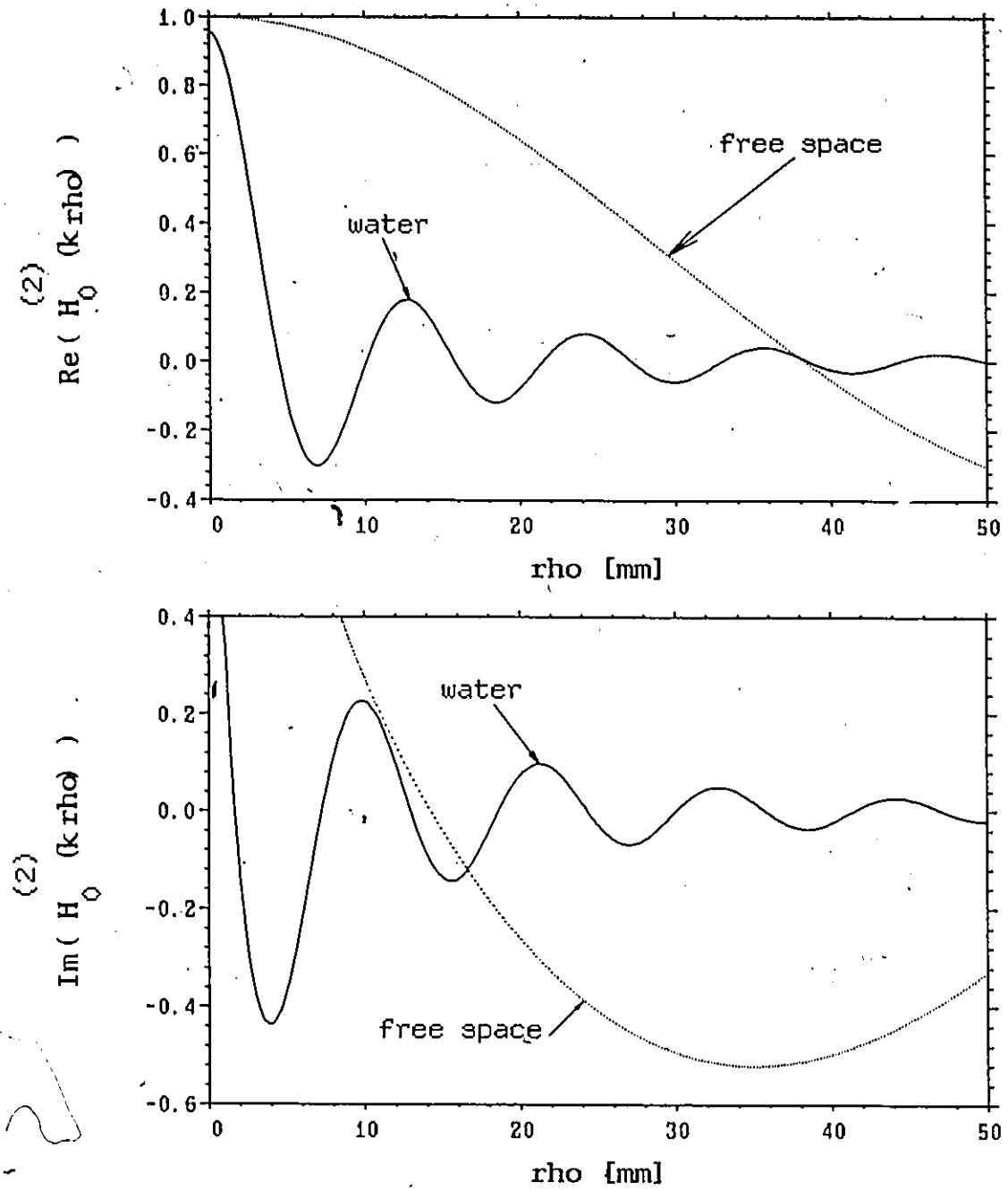


Figure 4: $H_0^{(2)}(k_0\rho)$ and $H_0^{(2)}(k_A\rho)$
Hankel function (kernel) in free space (k_0) and in distilled water at the nominal operating condition ($\kappa_A = 76.5 - j11.3$) as a function of the distance from the origin of the impulse ρ (rho).

Similarly, the two-dimensional *inverse Fourier transform* is defined by

$$f(x,y) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{f}(u,v)e^{j2\pi(ux+vy)}\,du\,dv \tag{38}$$

A set of sufficient conditions for eq. (37) to exist is:

1. $\int_{-\infty}^{\infty} |f(x,y)|\,dx\,dy < \infty$, i.e. the function must absolutely integrable.
2. $f(x,y)$ must have only a finite number of discontinuities.
3. $f(x,y)$ must have no infinite discontinuities.

Jump discontinuities are handled in the inverse transform by averaging about the discontinuity. For the one-dimensional case, we write (see Bracewell [1978, p. 7])

$$f(x_0) \leftarrow \frac{1}{2}(f(x_0+0) + f(x_0-0))$$

BANDLIMITED. A function $f(x,y)$ is *bandlimited* if its Fourier Transform $\tilde{f}(u,v)$ vanishes outside $u^2+v^2 \leq k^2$ where k is a real constant.

Transforming the scalar Helmholtz equation eq. (25)

$$\nabla^2 E_s(\underline{x};\theta) + k_A^2 E_s(\underline{x};\theta) = J_{eff}(\underline{x};\theta) \tag{39}$$

into Fourier space, we have (see Davies [1984, p.178+])

$$\widetilde{J_{eff}}(\underline{u};\theta) = (|\underline{u}|^2 + k_A^2)\widetilde{E_s}(\underline{u};\theta) \tag{40}$$

where $\underline{u} = (u,v)$ are the coordinates in Fourier space.

The solution to the inverse problem from eq. (40) is straightforward, and requires the Fourier Transform of the scattered field $\widetilde{E_s}(\underline{u};\theta)$ to be known everywhere in $u-$space. However, because the object and hence $J_{eff}(\underline{x};\theta)$ is of finite support, $\widetilde{J_{eff}}(\underline{u};\theta)$ and hence $\widetilde{E_s}(\underline{u};\theta)$ must necessarily be of infinite support, if they are not identically zero (see Landau and Pollack [1961, p.70]).

This does not preclude using $f(x,y)$ as an approximate band-limited function (see Landau and Pollack [1962]). The difficulty with this approach is that the scattered field $\widetilde{E_s}(\underline{u};\theta)$ be measured everywhere in $\underline{x}-$space, including inside the object. Clearly, this is an impractical approach.

2.2.4 Fourier Domain - Band-limited Data

In this section, we consider the use of Wolf's formulation [1969] which is based on spatially bandlimited data. This assumption is useful because the inversion relation becomes a Fourier Transform which can easily be computed with a Fast Fourier Transform algorithm (see, for example, Gonzalez and Wintz [1987, p.87]). This is the principal advantage of this formulation. However, Wolf presented his results for "weakly-scattering objects". Due to the wide range of dielectric constants we consider (refer to Table 1), this assumption is not applicable to our problem.

Consider a three-dimensional source distribution $F(\underline{x})$ imbedded in a *lossless* ambient medium and irradiated by a monochromatic plane wave with a wavevector $\underline{k} \triangleq k_A \underline{s} \in R$, i.e. $u_i(\underline{x}) = e^{j\underline{k}\cdot\underline{x}}$, where $\underline{s}$ is the direction of the incident wave. For *weakly-scattering objects*, i.e.

$$|E_s| \ll |E_i| \tag{41}$$

a good approximation to the scattered field is given by the first Born approximation

$$E_s(\underline{x}) = -\frac{j}{4\pi}\int_D F(\underline{x}) e^{j\underline{k}\cdot\underline{x}} G(|\underline{x}-\underline{x}'|)\, d\underline{x}' \tag{42}$$

where $G(|\underline{x}-\underline{x}'|)$ is the appropriate Green's function. For this problem, the scattering potential $F(\underline{x})$ (eq. (29)) is bandlimited to the *visible part* of the spectrum, i.e. u, v such that

$$u^2 + v^2 \leq k^2 \tag{43}$$

and its Fourier Transform can be determined from knowledge of the two-dimensional Fourier components of the scattered field in two planes outside the object. This approximation is useful because it permits reconstruction of an object "layer by layer" since the effect of the scattered field on the other layers is neglected.

Mittra [1973, p.361] considered the problem of determining the aperture field f of an antenna by probing the radiated field g in the Fresnel region (see Figure 5), i.e. the region in which the fields are related by the Rayleigh diffraction formula

$$g(x,y) = \int_{-b}^{b}\int_{-a}^{a} f(x',y')h(x',y',x,y)\, dx'\, dy' \tag{44}$$

where

$$h(x',y',x,y) = -\frac{1}{2\pi}\frac{\partial}{\partial z}\left(\frac{e^{-jkr}}{r}\right) \tag{45}$$

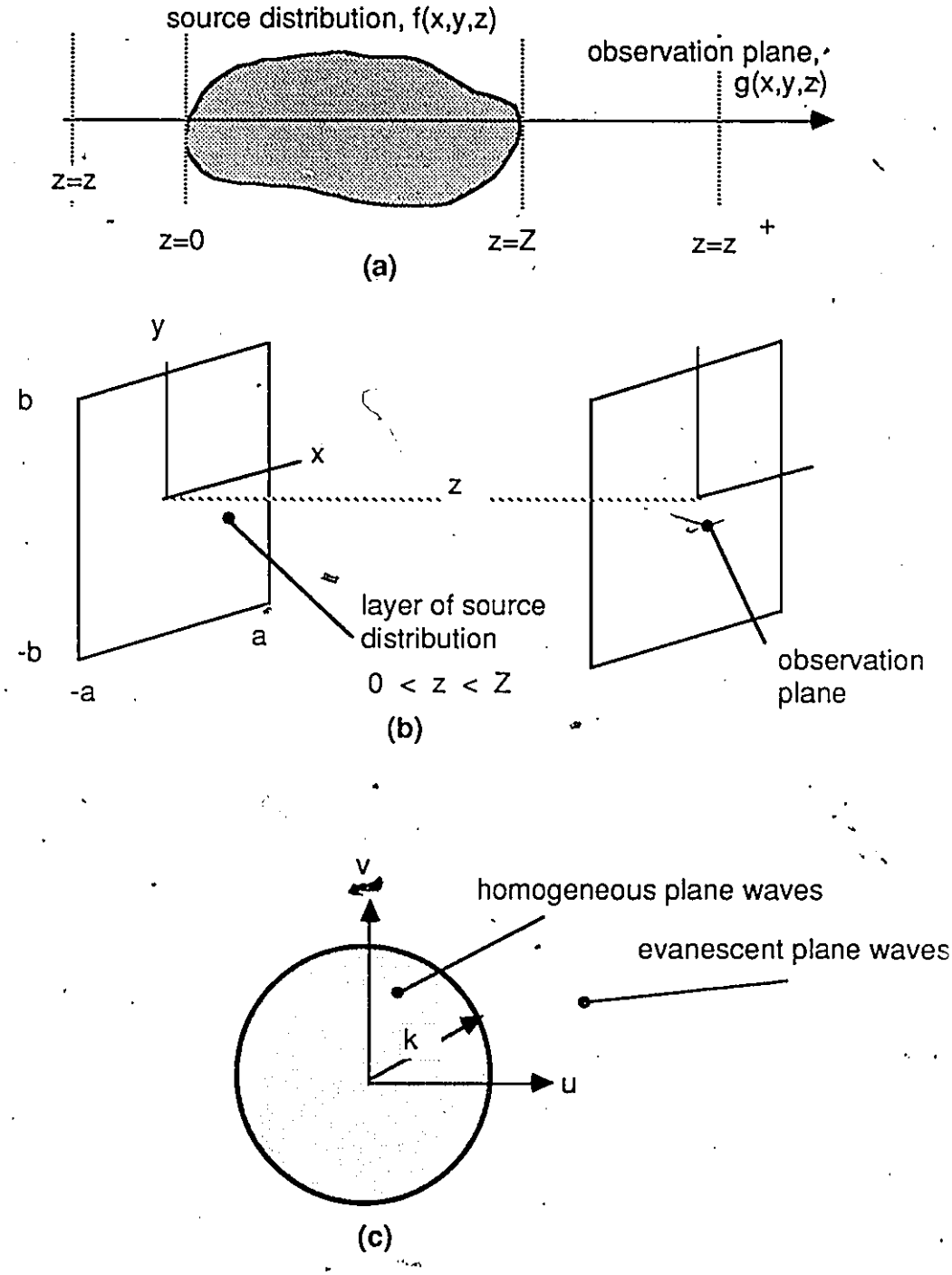


Figure 5: Reconstruction Model using the Layer-by-Layer Approach
This approach is appropriate if the data is band-limited, as is assumed with the Born approximation. (Adapted from Mittra [1973].)

This can be expressed as a convolution

$$g(x,y) = h(x,y) * * f(x,y) \tag{46}$$

An approximation which inverts only the visible part of the spectrum is

$$f(x,y) = \overline{h(x',y',x,y)} * * g(x,y) \tag{47}$$

This inversion formula is exact if the distribution is bandlimited inside the visible part of the spectrum. (If we are using "layer by layer" reconstruction, then his results are applicable).

The assumption of band-limited data in a lossless ambient medium was used by the group of Baribaud et al. [1985] in their work on microwave imaging. Their approach was similar to that of Mittra.

Another approach was that of Pichot et al. [1985]. Reconsider the $E_s - J_{eff}$ relation (eq.(33)) for a lossless ambient medium $k = Re(k_A)$, and reconstruct instead the quantity

$$K(\underline{x};\theta) = J_{eff}(\underline{x};\theta)/E_i(\underline{x};\theta) \tag{48}$$

which we note is a function of the angle of incidence of the incident wave θ (see Figure 6). The Green's function for problem can be expressed as an integral (see Tyras [1969, p.114])

$$\frac{j}{4}H_0^{(2)}(k\sqrt{(x-x')^2+(y-y')^2}) = \frac{j}{2}\int_{-\infty}^{\infty}\frac{e^{-j\sqrt{k^2-4\pi^2\nu^2}|y-y'|}}{\sqrt{k^2-4\pi^2/\nu^2}}e^{-j2\pi\nu(x-x')}d\nu \tag{49}$$

with $Im(\sqrt{k^2-4\pi^2\nu^2}) \leq 0$

Substituting eq.(49) into eq.(33), after some manipulation the following inversion formula is obtained

$$\widetilde{K}(u,v;\theta) = j2\sqrt{k^2-4\pi^2\nu^2}e^{j2\pi\Omega(\nu)Y_0}\hat{\Psi}(\nu,Y_0;\theta) \tag{50}$$

where

$$\Omega(\nu) = Re(\frac{-k}{2\pi}+\frac{1}{2\pi}\sqrt{k^2-4\pi^2/\nu^2}) \in R$$

$$\alpha = \nu\cos\theta - \Omega(\nu)\sin\theta \in R$$

$$\beta = \nu\sin\theta + \Omega(\nu)\cos\theta \in R$$

and

$$\tilde{\Psi}(\nu,Y_0;\theta) = \int_X \frac{E_s(\underline{x};\theta)}{E_i(\underline{x};\theta)}e^{j2\pi\nu X}dX$$

is the Fourier Transform of the normalized scattered field on the line $Y = Y_0$. The Fourier Transform of K is obtained by irradiating the object at various angles θ and measuring the

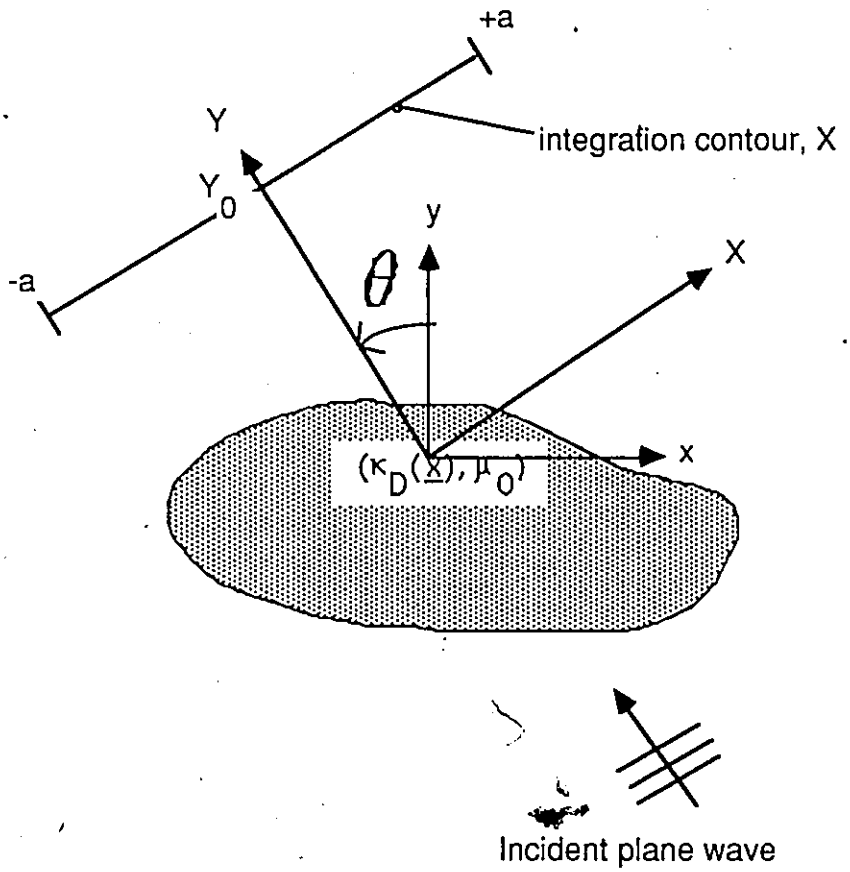


Figure 6: Reconstruction Model used by Pichot et al. [1985]
The reconstruction is performed by measuring the scattered field along a segment on $Y = Y_0$ for plane waves incident upon the object at angles θ.

scattered field $E_s(\underline{x};\theta)$ (called *projections*) , obtaining $\widetilde{K}(u,v;\theta)$, averaging the transform over the values at each projection, and finally Fourier Transform inverting the result.

This approach has two disadvantages for the problem we consider. First, it assumes that the ambient medium is lossless. As water is used in their experimental work, this assumption is not valid. Second, the quantity they reconstruct is a function of the incident wave, thus each projection actually gives information about different objects: $K(\underline{x};\theta_1), K(\underline{x};\theta_2), \ldots$.

Chapter 3

SOLUTIONS OF THE OPERATOR EQUATION $Kf = g$

In this chapter, *solutions* to the operator equation

$$Kf = g \tag{51}$$

are discussed, where $f \in X$ and $g \in Y$, and K is a linear, bounded non selfadjoint compact operator. The operator K_H defined in eq. (33) is compact if $X = Y = L_2$. (Mathematical preliminaries for this chapter are reviewed in Appendix E.) This problem is discussed because the term "solution" is not intuitively clear when the relation (51) is not satisfied exactly. The classical definition of "solution", i.e. f is a solution if $Kf = g$ is exactly satisfied and the element $f \in X$ in unique in X, is not appropriate. Due to uncertainties in measuring g or in finite-precision computation, it may be impossible to satisfy the relation eq. (51) exactly. In this case it is appropriate to look for an element f which minimizes some norm, e.g.

$$\|Kf - g\|_2 \tag{52}$$

generated by the inner product

$$(u, v) = \int_\Omega u\bar{v}\, d\Omega \tag{53}$$

(This norm is useful because it is a continuously differentiable function of f.) Hence the relation eq. (51) may have (1) no solution, (2) a unique solution, or (3) multiple (perhaps, an infinite number of) solutions. (This is analogous to solutions of a set of n linear equations in n unknowns, $Ax = y$, where the number of solutions depends on the determinant $\det(K)$

and the consistency of the equations). By searching for a solution to eq. (51) we are in fact *inverting* the equation.

Consider a function f_0 in the nullspace of K, i.e. $f_0 \in \ker(K)$ or $Kf_0 = 0$. Given any solution of eq. (51), $f + f_0$ is also a solution. Hence, an operator with a non-empty null space will not have a unique solution, unless constraints are imposed. For the inverse scattering problem operator considered K_H (eq. 33), the nullspace consists of the set of so-called "non-radiating sources" which produce no fields outside the domain of the object. Although such non-radiating sources can exist for *inverse source* problems (where a set of arbitrary sources produce the external fields), they are non-physical for inverse scattering problems because of the extra constraints added by the constitutive relations. The uniqueness of the solution to the inverse scattering problem has been proven by Stone [1981].

Working in the function space $X = L_2$, the elements of L_2 are "energy"-limited (see Appendix E). The subspace $X_A \subset X$ is the set of elements of X which satisfy some physical admissibility constraints, e.g.

$$l_R \leq Re(h(f)) \leq u_R, \quad l_I \leq Im(h(f)) \leq u_I \tag{54}$$

where $h(f)$ is some function (to which physical constraints can be applied) and $l_R, u_R, l_I, u_I \in R$ are constants. Such elements are called *admissible solutions.* (In our case, $h(f)$ will represent the function which gives the complex relative permittivity from the effective current, eq. (36), and the above inequalities constrain the real and imaginary parts of the complex relative permittivity.) The subspace $X_{A\epsilon} \subset X_A$ is the set of elements of X_A which also satisfy

$$\|Kf - g\|_2 \leq \varepsilon \tag{55}$$

i.e. which satisfy eq. (51) to a fixed error ε. Elements of the set $X_{A\epsilon}$ are called *feasible solutions.*

Further, only least squares solutions are considered because they are easier to deal with.

> LEAST-SQUARES SOLUTION. An element $f_{LS} \in X$ is said to be a least-squares solution of eq. (51), if for a given $g \in \mathrm{dom}(K^\dagger)$, i.e. in the domain of the pseudo-inverse of K,
>
> $$\|Kf_{LS} - g\| = \inf\{\|Kf - g\|_2 : f \in X\} \tag{56}$$

A least-squares solution satisfies the set of *normal equations*

$$K_{LS} \triangleq K^*Kf = K^*g, \quad \text{or} \quad Kf = g + g_1, \text{ where } K^*g_1 = 0$$

(see Kammerer [1972] or Tikhonov [1977]) which is found by setting the first variation of the norm $\|Kf - g\|_2$ to zero (see Nashed [1981, p. 222]). This equation is equivalent to the $Kf = Pg$ where P is the projection operator onto ran(K) (see Bertero [1986, p. 60]). We denote the set of all least-squares solutions of eq. (51) by

$$S_{LS} = \{f : \|Kf - g\|_2 = \inf_{f \in X} \|Kf - g\|_2\} \tag{57}$$

We shall see that our operator K_H is unstable and that its solution represents an *ill-posed* problem, and the linear system obtained by discretization is *ill-conditioned*. In this chapter, this instability will be examined and quantified. Also, methods of solving the unstable linear system $[K][f] = [g]$ (where [ ] denotes a vector or matrix) obtained by discretization of eq. (51) will be explored.

3.1 Stability and Ill-posed Problems

Two concepts are now examined: an ill-posed problem and a stable problem.

> WELL-POSED. A problem $K : X \mapsto Y$ is well-posed relative to spaces (X, Y) in the sense of Hadamard if the solution f_* possesses the properties of
>
> - uniqueness : the solution f_* is unique in X for every $g \in Y$
> - existence : a solution f_* exists for every $g \in Y$
> - continuity : the solution $f \in X$ depends continuously on the data, i.e. on g or on any other variable on which f depends (for example, on the position vector $\underline{x}$.)

A problem which does not satisfy all of these equations is called *ill-posed.* If after certain additional constraints are imposed on the solution, the solution is stable (insensitive) relative to the data perturbations, then we say the problem is *conditionally well-posed.* The problem may be further classified by the *degree of ill-posedness* , i.e. by the rate of decay of the

singular values decay to zero (see Kristensson and Vogel [1986, p. 467]), but this notion is rather nebulous.

Phillips [1962] showed that the Fredholm integral equation of the first kind (eq. (51)) is ill-posed in the sense of Hadamard by using the Riemann-Lebesgue lemma

$$g(\omega, x) = \int_0^1 K(x, x') e^{j\omega x'} dx' \quad \to 0 \quad \text{as} \quad \omega \to \infty \tag{58}$$

In other words, g is not sensitive to the addition of arbitrarily high frequency components. This also shows that a small change in the data δg (a variation of g) may correspond to a large change in the solution δf. Let $f_1(x)$ be a solution of eq. (51) for $g = g_1(x)$. Then the solution $f_2(x) = f_1(x) + N\sin(\omega x)$ is also a solution of eq. (51) with

$$g_2(x) = g_1(x) + N \int_a^b K(x, x') \sin(\omega x) dx'$$

The L_2 distance between g_1 and g_2 is

$$d_2(g_1, g_2) = \|g_1 - g_2\|_2 = |N| \sqrt{\int_a^b \left(\int_a^b K(x, x') \sin(\omega x') dx' \right)^2 dx}$$

and if the value of ω is sufficiently great, the distance $d_2(g_1, g_2)$ can be made arbitrarily small (see Tikhonov [1977, p. 3+]). Thus the problem eq. (51) is ill-posed: the solution does do not depend continuously on the data in the L_2 setting. Further, one expects that $g(\omega, x) \to 0$ as $\omega \to \infty$ more quickly for a smooth kernel than for a sharply peaked kernel because $e^{j\omega x'}$ would oscillate quickly between positive and negative values, thus "cancelling" out the integral (Phillips [1962]).

For a well-posed (in some sense) problem, we can use a *condition number* to quantify the relative error propagation from g to the solution f.

CONDITION NUMBER. Let $Kf = g$ be an operator equation, and let δf be a variation in f and δg a variation in g. Then the condition number $\kappa(K)$ is given by

$$\frac{\|\delta f\|_X}{\|f\|_X} \leq \kappa(K) \frac{\|\delta g\|_Y}{\|g\|_Y}$$

where $\kappa(K) = \|K\|_X \|K^{/1}\|_X$ and the symbol $K^{/1}$ is used to denote the desired inverse.

The condition number indicates the relative distance of K from the nearest non-invertible operator. (For the matrix case, see Conte and de Boor [1980, p. 177].) This concept is important because computed solutions are performed in finite precision arithmetic. Consider a machine with the floating point format

$$d = \pm(0.d_1 d_2 \ldots d_t) \times \beta^e \qquad 0 \leq d_i \leq \beta, d_1 \neq 0, L \leq e \leq U \tag{59}$$

which can be described by the 4-tuple (β, t, L, U) where β is the radix of the computation, t the precision, L the underflow limit, U the overflow limit, and e an integer called the exponent (see Conte and deBoor [1980, p. 7]). Let x be a floating point number and $fl(x)$ its floating-point representation

$$fl(x) = x(1 + \delta)$$

Then the maximum possible value for $|\delta|$ is called the *unit roundoff* and is denoted by u. (This number is related to *roundoff error* which is the difference between the exact value of a number resulting from a floating-point operation, x, and its machine representation, $fl(x)$.) A problem is *well-conditioned with respect to the precision used* if it possesses a solution which is stable relative to some small changes in the data, i.e. $u\kappa(K) \ll 1$, where we say that "$\kappa_2(K)$ is low". A problem is *ill-conditioned with respect to the precision used* if $u\kappa(K) \ll 1$ does not hold, so, in effect, the unit roundoff u controls the accuracy of the computed solutions. In this case, the computed solution may have no resemblance to the (exact) solution to the problem.

The condition number is also useful because it limits the convergence rate of some projection methods of solution of equations, e.g. the method of successive approximations, steepest descent, and conjugate gradients. These methods all lead to least-squares solutions of the operator eq. (51) (see Householder [1975, p. 123]), and generally exhibit slow convergence, typically as $O(1/n)$. The method of conjugate gradients is further explored in Chapter 4.

3.2 Inverses and Solutions

A number of inverses have been studied in the past. Some of these will be examined here. In particular, we discuss the classical inverse K^{-1} with a solution f_{-1}, the generalized inverse $K^{\dagger}$ with a solution $f_{\dagger}$, and the regularized inverses K^{α} with the family of solutions $\{f_\alpha\}$. The stability properties of these inverses are now examined using the eigenfunction expansion

of the linear operator K. In this section, we first give results for operators (possibly infinite dimensional), and then for matrix (finite dimensional) operators. Further details may be found in the cited references.

Consider again eq. (51). The *eigenvalues* of this equation are the set of values $\lambda(K) = \{\lambda_i\} \in C$ for which

$$Kf_i = \lambda_i f_i \tag{60}$$

where $f_i \neq 0$. Also define $\mu_i \triangleq 1/\lambda_i$ when $\lambda_i \neq 0$. The set of eigenvalues $\lambda(K)$ is called the *point spectrum* of K, and the f_i are called the *eigenvectors* of K. Equation (51) will have solutions only if $f \perp \ker(K^*)$, but this condition is not generally sufficient because $\text{ran}(K)$ may not be closed (Stakgold [1979, p. 357]).

Now, if K is compact and one-to-one, then the classical inverse of K defined on its range can be represented by (see Naylor and Sell [1982, p. 482] or Smithies [1958, p. 145])

$$f_{-1} = K^{-1}g = \sum_{i=1}^{\infty} \frac{(g, u_i)}{\mu_i} f_i \tag{61}$$

where $\{u_i\}$ is an orthonormal set of eigenvectors of KK^*, and $\mu_i \to 0$ as $i \to \infty$. The domain of K^{-1} is all g for which the series converges. For ill-posed problems, we cannot say that the solution $f = K^{-1}g$ is the solution for two reasons : (1) g may not belong to KX in which case there is no solution, or (2) f may not be stable because K^{-1} is not continuous. (If K is not one-to-one, then eq. (51) may not have a unique solution.) In the latter case, it is reasonable to look for solutions such that Kf is as close to g as possible. This can be done by projecting g onto KX and finding the least-squares distance, eq. (52).

Least-squares solutions to eq. (51) may not be unique. However by restricting ourselves to the least-squares solution of minimum norm (LSMN) , i.e. the minimum $\|f\|_2^2$ produced by the linear *pseudo-inverse* operator, $K^\dagger$ we have a unique solution if and only if $\ker(K) = \{0\}$. (For a discussion on the pseudo-inverse and its properties, see Nashed [1976]).

> LSMN. An element $f_\dagger \in X$ is said to be a *least-squares minimum norm* solution of (51) for a given $g \in \text{dom}(K^\dagger)$ if $f_\dagger \in S_{LS}$ and $f_\dagger$ has the minimum norm of all least squares solutions
>
> $$f_\dagger = K^\dagger g = \{f : f \in S_{LS} \quad \text{and} \quad \|f\| \leq \min_{x \in S_{LS}} \|x\|\} \tag{62}$$

To further explore these solutions, the concept of a *singular system* is needed.

SINGULAR SYSTEM. Consider the eigenvalue problem for the two operators

$$K_{LS} = K^*K, \quad \text{and} \quad KK^* \tag{63}$$

which are both compact, symmetric, non-negative definite and self-adjoint if K is compact. The point spectrum $\lambda(K)$ for both operators is identical, and the eigenvalues are real and non-negative. Let u_n denote the set of eigenvectors for KK^*

$$KK^*u_n = \lambda_n u_n \tag{64}$$

and v_n the set of eigenvectors for K^*K

$$K^*Kv_n = \lambda_n v_n \tag{65}$$

where the eigenvalues are ordered $\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$ (see Naylor and Sell [1982, p. 477]). Let $\sigma_n \triangleq \sqrt{\lambda_n}$ for $\lambda_n > 0$. The positive *singular values* $\{\sigma_n\}$ together with the singular functions $\{u_n, v_n\}$ form the *singular system* for the compact operator K : $\{u_n, v_n; \sigma_n\}$, $n = 1, 2, ..., p$. It is always possible to choose $\{u_n\}, \{v_n\}$ so that the equations

$$Kv_n = \sqrt{\lambda_n}u_n, \quad K^*u_n = \sqrt{\lambda_n}v_n \quad \text{for} \quad n = 1, 2, \ldots, p, \quad \lambda_n > 0 \tag{66}$$

are satisfied, and $\{u_n\}$ and $\{v_n\}$ are orthogonal.

In general, an operator in an infinite dimensional space may not have eigenvalues, but since the operator K is compact, the following is known.

THEOREM. Let K be a compact operator on a Hilbert Space. Then the spectrum of $K, \sigma(K)$, is finite, or at most countably infinite and $\lambda = 0$ is the only possible point of accumulation. (From Naylor and Sell [1982, p. 452]).

Since K_H is compact, then the operators $K_H^*K_H$ and $K_HK_H^*$ have a finite or countable set of eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots$ with $\lambda_n \to 0$ in the latter case (see Nashed [1976, p. 193+]). Now recall that a necessary and sufficient condition for the existence of a least square solution is

$$\sum_i \frac{|(g, v_i)_Y|^2}{\lambda_i^2} < \infty \tag{67}$$

Note that if $\lambda_i \to 0$ as $i \to \infty$, then $|(g, v_i)_Y|^2$ must go to zero as $i \to \infty$ (see Bertero [1986, p. 62]). If a LSMN solution exists, it is given by

$$f_\dagger = K^\dagger g = \sum_i \frac{(g, v_i)_Y}{\sqrt{\lambda_i}} u_i \tag{68}$$

The last type of inverse operator we consider is the regularized inverse K^α where $f_\alpha = K^\alpha g$ and f_α is the element that minimizes the *additive functional*

$$M_\alpha(f) = \|Kf - g\|_2^2 + \alpha\Omega(f), \quad \alpha > 0 \tag{69}$$

where the *regularizing functional* $\Omega(f) \geq 0$ for all $f \in X$. This method transforms the operator equation (51) into an operator which is continuous and which has a unique solution. (In the case of the integral operator K_H, it transforms from a Fredholm integral equation of the first kind to a more stable equation of the second kind.) For simplicity, consider the less general case $\Omega(f) = \|Lf\|_2^2$ where L is a closed linear differential with dense domain and closed range. Such a choice comes from the observation that ill-posed problems have oscillating solutions whose fluctuations are reduced by minimizing derivatives of f, e.g. $\Omega(f) = \|f\|^2$ (see Phillips [1962]). Such a choice also makes the problem more tractable. The normal equations corresponding to this problem can be found from the Euler-Lagrange equation to be

$$(K^*K + \alpha L^*L)f = K^*g, \quad \alpha > 0 \tag{70}$$

where the operator $(K^*K + \alpha L^*L)$, $\alpha > 0$ is one-to-one and its inverse is continuous (see Nashed [1974, p. 305]). Formally, the solution is given by

$$f_\alpha = K^\alpha g = (K^*K + \alpha L^*L)^{-1} K^*g \to K^\dagger g \quad \text{for} \quad g \in \text{dom}(K^\dagger) \quad \text{as} \quad \alpha \to 0 \tag{71}$$

$$= \sum_{n, \lambda_n \neq 0} \frac{(g, Ku_n)}{\alpha + (1-\alpha)\lambda_n} u_n \tag{72}$$

(see Nashed [1981, p. 228]).

If the measured g is contaminated by an amount δg such that $\|\delta g\| \leq \epsilon$ then

$$f_\alpha^\epsilon = \sum_n \frac{(g, Ku_n)}{\alpha + (1-\alpha)\lambda_n} u_n + \sum_n \frac{(\delta g, Ku_n)}{\alpha + (1-\alpha)\lambda_n} u_n \tag{73}$$

The second term is the error in the solution due to uncertainty in g. It can be made small by making α large, but this leads to larger difference in satisfaction of $\|Kf - g\|_2$. If we take $\alpha \to 0$, then f_α^ϵ is close to the LSMN solution $K^\dagger g$. However, as $\alpha \to 0$, the system tends to be poorly conditioned, and hence numerically unstable. Nashed [1981, p. 226] suggests $\alpha = \epsilon^2$ as a choice that "works".

3.3 Discretization of the continuous operator $Kf = g$

Now consider the methods of transforming the (continuous) operator equation $Kf = g$ into the matrix equation $[A][x] = [y]$. Two symbols are used for the operator (K and $[A]$) for two reasons: (1) to permit discrimination between the operator and matrix formulations, and (2) to emphasize that they may represent (slightly) different problems.

One method which has found many applications is based on the following.

MINIMAL FUNCTIONAL THEOREM. Consider the operator equation $Kf = g$ where, $f, g \in C$. If the operator K is positive definite (and hence, by definition, self-adjoint), then the solution to the equation is the function which minimizes the functional

$$J(f) = (Kf, f) - (f, g) - (g, f) \tag{74}$$

This theorem is the foundation of the *Ritz method*, and can be extended to non self-adjoint problems (see Chen and Lien [1980].) However, this method applied to a non-selfadjoint problem does not guarantee better approximations to the unknown function f as the degree of approximation is increased (see Sarkar [1984, p. 1162].)

Alternatively, the *method of projections* can be applied, in a manner which is more straightforward (see Reinhardt [1985, chapter 2]). In the discussion that follows, we stress that the operator K need not be positive-definite, it need not even be symmetric, nor necessarily linear. We assume that the inner product is defined on a Hilbert space, but a more general setting is possible (op. cit.).

Consider a separable Hilbert space H with basis $\{\phi_i\}_{i=1}^{n_*}$, $1 \leq n_* \leq \infty$ and a set $M \subset H$ which is dense in H.

THEOREM. If $u \in H$ is orthogonal to all elements of a set M with basis functions $\{\psi_i\}_{i=1}^{m_*}$, $1 \leq m_* \leq \infty$, which is dense in H then u is the null element of H, i.e.

$$(u, \psi_i) = 0, \quad i = 1, 2, 3, \ldots, m_* \tag{75}$$

implies $u = 0$ (see Rektorys [1980, p. 79]).

For an element $f_* \in X$ which satisfies

$$(Kf_* - g, \psi_i) = 0 \quad \text{for every} \quad i = 1, 2, \ldots, m_* \tag{76}$$

then, according to the theorem, we have

$$Kf_* - g = 0$$

In other words, f_* is the solution of $Kf = g$ in X.

Now, let $\{\phi_i\}, \phi_i \in \mathrm{dom}(K)$ be a basis for $X \subseteq H$ (called *expansion functions*), and $\{\psi_i\}, \psi_i \in \mathrm{ran}(K)$ be a basis for $Y \subseteq H$ (called *testing functions*). For simplicity, the expansion functions are assumed to satisfy the boundary conditions of the operator, K. The sets of expansion functions and testing functions are, in fact, both *basis functions* for the spaces X and Y respectively. We now look for an approximate solution f_n

$$f_n = \sum_{i=1}^{n \le n_*} \alpha_i \phi_i \tag{77}$$

where $\alpha_i \in C$. We define the two subspaces spanned by the basis functions $X_n = \{\phi_i\}_{i=1}^{n} \subseteq X$ and $Y_n = \{\psi_i\}_{i=1}^{m} \subseteq Y$. Substituting eq. (77) into eq. (76),

$$(Kf_n - g, \psi_i) = 0, \quad i = 1, 2, \ldots, m \tag{78}$$

which represents $m \le m_*$ equations in $n \le n_*$ unknowns. (This equation can be written in operator form as

$$P_n K f = P_n g \tag{79}$$

where P_n is a *projection operator*. Such operators have a range which is closed (Bertero [1986, p. 58]), and if K if infinite dimensional, they project the problem to a finite dimensional space. In order for them to converge to the exact solution, then $P_n g \to g$ as $n \to n_*$.) In this case, the discretization procedure will give us the "best" solution from the family of solutions prescribed by the given sets of basis functions, $\{\phi_i\}, \{\psi_i\}$.

For simplicity, now assume that the operator K is linear and thus that eq. (78) reduces to the linear system of m equations n unknowns.

$$\begin{aligned}
(K\phi_1, \psi_1)\alpha_1 + (K\phi_2, \psi_1)\alpha_2 + \ldots + (K\phi_n, \psi_1)\alpha_n &= (g, \psi_1) \\
&\ldots \\
(K\phi_1, \psi_m)\alpha_1 + (K\phi_2, \psi_m)\alpha_2 + \ldots + (K\phi_n, \psi_m)\alpha_n &= (g, \psi_m)
\end{aligned} \tag{80}$$

If we denote

$$[A]^{(n)} = [a_{ij}], \quad a_{ij} = (K\phi_i, \psi_j), \quad [x] = \{(\alpha_i)\}, \quad \text{and} \quad [y] = \{(g, \psi_i)\}$$

we have the system

$$[A]^{(n)}[x] = [y]$$

and have thus reduced the problem of solving an operator equation eq. (51) to the deceptively simple problem of solving a system of m linear equations in n unknowns, eq. (80). This set of equations is entirely formal, without any necessary physical interpretation. (This is in contrast to the Ritz Method which may thought of as minimizing an *energy* functional).

It must be stressed that solving such a linear system on a digital computer, we are not solving the original *continuous* problem, but another *discretized* problem whose discrete solution approximates in some sense the continuous solution of the continuous problem.

Special cases of the method of projections for linear operators include the *Petrov-Galerkin method* (in a Hilbert Space setting), the *Bubnov-Galerkin method* (if $\psi_i = \phi_i$), and the *method of moments* (if $\psi_i = M\phi_i$ where M is a one-to-one mapping from X to Y) (see Harrington [1967]). The method of moments itself includes the *method of least squares* if K is one-to-one (see Reinhardt [1985,p. 39]).

Equation (51) is uniquely solvable for every $g \in Y$ for $n \geq n_0$ when $P_n K$ is a one-to-one mapping of X_n to Y_n, i.e.

$$\ker(P_n K) = \{0\} \tag{81}$$

Convergence and related topics are discussed in Reinhardt [1985, chapter 2] and Sarkar [1984].

An instance of the method of projections which is particularly easy to compute because it avoids difficult integrations is the *collocation* or *point matching* method (see Lebedev et al. [1979, p. 404-413]). Here, $\psi_i = \delta(\underline{x} - \underline{x}_i)$ is the impulse distribution and the equation $Kf = g(\underline{x}_i)$ is solved at m discrete points $\{\underline{x}_i\}_{i=1}^m$ called the collocation points. One disadvantage with this method is that the approximate solution may vary considerably with the position of the collocation points. Lebedev et al. [1979, p. 408] suggest that this effect may me minimized by taking "a sufficient number of points and distribute them over the domain and boundary".

Following Greenstadt [1959], a popular choice as a basis for the unknown function $f(\underline{x})$ are *pulse functions* $P_i(\underline{x})$ which partition the domain of f into contiguous subdomains (called *cells*), with the assumption that f is constant in each of the subdomains.

$$P_i(\underline{x}) \triangleq \begin{cases} 1 & \underline{x} \in \text{ cell } i \\ 0 & \underline{x} \notin \text{ cell } i \end{cases} \tag{82}$$

3.4 The Matrix Equation $[A][x] = [y]$

Consider computed solutions to the system of m linear equations in n unknowns:

$$[A][x] = [y] \tag{83}$$

where $[A] \in C^{m \times n}$ is a matrix, and $[x] \in C^n$ and $[y] \in C^m$ are vectors. Let $[x]$ be the unique exact solution to the system and let $[\hat{x}]$ be the computed solution. Recall that the solutions of a set of linear equations are continuous functions of the coefficients and the right-hand sides, but this modulus of continuity (whose least upper bound is given by the condition number, κ_2) may be very large (see Szidarovsky and Yakowitz [1978, p. 212-220]).

We now discuss some results for finite-dimensional (matrix) operators which are analogous to those discussed in section 3.2 for general operators.

If the matrix $[A]$ is not singular, i.e. $\det([A]) \neq 0$, then $[A]$ has an inverse $[A]^{-1}$ and the solution is given by

$$[x]_{-1} = [A]^{-1}[y] \tag{84}$$

However, a general $m \times n$ matrix $[A]$ can be decomposed into a form called the *singular value decomposition* (SVD) of $[A]$

$$[A] = [U][\Sigma][V]^H \tag{85}$$

where $[U]$ and $[V]$ are unitary (i.e. $[U][U]^H = [I], [V][V]^H = [I]$), and $[\Sigma]$ is a rectangular diagonal matrix of the same size as $[A]$ with non-negative real diagonal elements called the *singular values* of $[A]$: $\sigma([A]) = \{\sigma_i\}_{i=1}^p$ where $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_p$ and $p = \text{rank}([A])$ (see Golub and Kahan [1965] and the LINPACK Users' Guide [1979]). The singular values are the positive square roots of the non-negative eigenvalues of $\{\lambda_i\}$ of $[A]$. (This method is the same as that in eq. (66) with $[A]^H$ replacing $[A]^*$.) SVD is useful because the κ_2 condition number of the matrix $[A]$ is given by

$$\kappa_2([A]) = \|[A]\|_2 \|[A]^\dagger\|_2 = \frac{\sigma_1([A])}{\sigma_p([A])} \tag{86}$$

where the $p-$norm of a matrix $[A]$ $\|[A]\|_p \triangleq \sup_{x \neq 0} \frac{\|[A][x]\|_p}{\|[x]\|_p}$. If $[x] = (x_1, x_2, \ldots, x_n)$, then for $p = 2, \|[x]\|_2 = \sqrt{\sum_i |x_i|^2}$. and for $p = \infty, \|[x]\|_\infty = \max_i |x_i|$. The singular values can be accurately computed to the unit roundoff u, i.e. to $O(1 + \beta^{-t})$, by the algorithm in the LINPACK subroutine ZSVDC (Cline et al. [1982, p. 74]).

The condition number is useful in giving an upper bound on the sensitivity of the problem $[A][x] \doteq [y]$. If we consider only real values for a non-singular matrix $[A] \in R^{m \times n}, [x] \in R^m$ and $[y] \in R^n$, Golub and van Loan [1984, p. 24-26] show that the system under perturbation can be expressed as

$$([A] + \epsilon[F])[x](\epsilon) = [y] + \epsilon[f], \quad x(0) = [x] \tag{87}$$

where $[F] \in R^{n \times n}$ and $[f] \in R^n$ are "small" perturbations, and the relative error in the solution x is

$$\frac{||[x(\epsilon)] - [x]||}{||[x]||} \leq \kappa([A])\epsilon[\frac{||[F]||}{||[A]||} + \frac{||[f]||}{||[y]||}] + O(\epsilon^2) \tag{88}$$

Chapter 4

INVERSE SCATTERING - Object Reconstruction Methods

In this chapter, the mathematical apparatus reviewed in the previous chapter is used to compute solutions to the operator K_H, eq. (33). Specifically, consider computing the complex relative permittivity of an object $\kappa_D(\underline{x}), \underline{x} \in D$ and the scattered field $E_s(\underline{x};\theta), \underline{x} \in O$ when the object is irradiated by a plane wave propagating at an angle θ with respect to the $+y$-axis, $E_i(\underline{x};\theta)$ (see Figure 1 for the geometry). This chapter retains the assumptions of Chapter 2, and hence *it is not necessary to assume that the object is "weakly-scattering", nor that the ambient medium is lossless.*

In practice, the scattered field E_s is measured in a two-step process; (1) without the object, the incident field $E_i(\underline{x};\theta)$ is measured, and (2) with the object inserted, the total field $E(\underline{x};\theta)$ is measured. The difference between these two measurements is the scattered field

$$E_s(\underline{x};\theta) = E(\underline{x};\theta) - E_i(\underline{x};\theta) \tag{89}$$

A system to perform such measurements in described in Appendix B.

Based on the available mathematical apparatus and physical reasonableness, we restrict ourselves to function which are in L^2, i.e.

$$\int_{\Omega=D,O} |u(\underline{x})|^2 d\underline{x} < \infty, \quad u = E_i(\underline{x};\theta), E(\underline{x};\theta), \text{ or } E_s(\underline{x};\theta) \tag{90}$$

which is also a Hilbert space, H.

In this chapter, two topics are discussed: (1) a careful model of the integral operator K_H is formed so that it can by discretized by the method of projections, and (2) the standard

methods of solving a system of linear equations are reviewed. Due to the limitations of the standard methods of solving a highly ill-conditioned system of linear equations, a new method of handling such problems is proposed. It is based on the method of conjugate gradients.

4.1 Applying the Method of Projections to the Operator K_H

Chapter 3 described how the method of projections could be applied to operators such as K_H. Traditionally, this is a low frequency method (i.e. cell size small compared to λ_A) which transforms an integral equation to a system on linear equations. Historically, it has been used to about 600 MHz for direct scattering problems, beyond which its computational cost becomes "prohibitive" (Kastner [1983].) For higher frequencies, geometrical optics (the $\lambda \rightarrow 0$ approximation) is normally used.

In order to apply the Method of Projections correctly, we must find an adequate set of basis functions for the object J_{eff} as well as for the scattered field E_s. Hence, care must be taken in modelling

- the object, J_{eff}
- the operator, K_H (in most cases it must be approximated), and
- the scattered field E_s.

Further, the choice of method of solving the resulting system of linear equations will also influence the solution due to their varying sensitivities to ill-conditioned problems.

4.1.1 The Object Model

Historically, Richmond [1965] first applied the method of projections (specifically, the point matching method with pulse expansion functions) to the direct problem, DP(κ_D). He divided the object into a grid of square cells with edge size d [m]. Actually, he considered instead circular cells with radius $a_c = d/\sqrt{\pi}$ because such cells provide an analytically tractable integral. Richmond found that "for accurate results"

$$d \leq d_R \triangleq 0.2\frac{\lambda_0}{\sqrt{\kappa'_D}} \tag{91}$$

where "numerical calculations has shown that little error is incurred in approximating square cells with circular cells of the same cross-section area." The concept involved here is to keep J_{eff} and therefore κ_D and E (the $E-$ field inside the cell) "nearly uniform over each cell". Of course, such a model ignores the boundary condition $[\hat{n} \times \mathbf{E}] = \mathbf{0}$, i.e. the tangential components of E are continuous. (With such a model, a charge density ρ_s is required at the surface on each cell in order to satisfy the boundary condition $[\hat{n} \cdot \mathbf{D}] = \rho_s$, but is usually ignored [Hagmann, 1987].) Under these assumptions, the system of linear equations obtained has been observed by many authors to be quite stable, i.e. the condition number of the matrix κ_2 is small, typically less than 10^4 (see also Hagmann [1987, p. 21]). This is typical of Fredholm integral equations of the second kind (with the total field inside the object $E(\underline{x})$ being the unknown), of which the equation for this direct problem is an example. (Recall that the inverse problem leads to a first kind equation.)

Later, Jones [1974, p. 580] pointed out that the kernel of the integral equation $\frac{j}{4}H_0^{(2)}(k_A\rho)$ is oscillatory, with period $T = 2\pi/Re(k_A)$ (at least in the far-field) and "to prevent miscalculation" choose $a_c \leq \frac{\lambda_A}{5} = \frac{\lambda_0}{5\sqrt{\kappa'_A}}$, or

$$d \leq d_J \triangleq \frac{\sqrt{\pi}\lambda_0}{5\sqrt{\kappa'_A}} \tag{92}$$

Hagmann et al. [1977] later derived another cell-size estimate by examining the properties of the Helmholtz equation which they considered local to each cell, i.e. $\nabla^2 E + k_D^2 E = 0$, where $k_D^2 = \omega^2\mu_0\epsilon_0\kappa_D$, and κ_D is the complex relative permittivity of the cell. Each cell can be considered a homogeneous, linear and isotropic space. They found in the two-dimensional case that if

$$d \leq d_H \triangleq \frac{2}{|k_D|} = \frac{\lambda_0}{\pi\sqrt{\kappa'_A}} \tag{93}$$

then the field variations are not "large" if pulse expansion functions are used.

For the operating condition defined in Chapter 1 (7), we have $d_R = 2.29$ mm, $d_J = 4.05$ mm, $d_H = 3.64$ mm, or

$$d \leq d_{max} = \min\{d_R, d_J, d_H\} = 2.29 \text{ mm} \tag{94}$$

4.1.2 The Operator K_H

The operator $K_H : J_{eff} \rightarrow E_s$ is a Fredholm integral equation of the first kind with kernel $\psi(\underline{x}, \underline{x}') = \frac{j}{4}H_0^{(2)}(k_A|\underline{x} - \underline{x}'|)$ (eq. (32).) Its inherent instability was discussed in Chapter

3. For such an integral equation to be self-adjoint, the kernel must be symmetric, i.e. $\psi(\underline{x},\underline{x}') = \overline{\psi(\underline{x}',\underline{x})}$ (see Friedman [1982, p. 238]). This is not the case here, hence K_H *is not self-adjoint.* Since $K(\underline{x},\underline{x}'), \underline{x}' \in D, \underline{x} \in R^2 \setminus D$ is continuous, it is bounded, and hence *compact* (see Friedman [1982, p. 237+] or Ramm [1986]). Compact operators have a range which is non-closed (Bertero [1986, p.61]).

However, operators in an infinite dimensional space, e.g. non self-adjoint compact operators, need not have eigenvalues (Ramm [1986, p. 6]). Instead, if we consider the operator corresponding to the least-squares problem

$$K_{H,LS} \triangleq K_H^* K_H \tag{95}$$

which is compact (see Chapter 3), then eigenvalues do exist and the condition number of the problem is given by

$$\kappa_2(K_{H,LS}) = \frac{\sigma_{max}}{\sigma_{min}} \to \infty \tag{96}$$

(see Bertero [1986] for operator form, and Householder [1975, p. 81-82] for the matrix form) since $\lambda_{min} \triangleq \sigma_{min}^2 \to 0$ by the Theorem in Chapter 3. Because any computable approximation to K_H is necessarily finite dimensional (say an $M \times N, M \geq N$ matrix $[K_H]^{(N)}$), it must have a basis , say $\{b_i\}_{i=1}^N$, with spectrum

$$\sigma([K_H]^{(N)}) = \{\lambda \in C : \det(\lambda \mathrm{I} - [K_H]^{(N)}) = 0\} \tag{97}$$

which is non-empty and contains at most N complex numbers (Naylor and Sell [1982, p. 415]).

Hence if $[K_H]$ is a matrix approximation to the non-selfadjoint operator K_H, the condition number $\kappa_2([K_H])$ will approach infinity as the $[K_H]$ better approximates K_H

$$\kappa_2([K_H]^{(N)}) \overset{N\to\infty}{\longrightarrow} \kappa_2(K_H) \to \infty \tag{98}$$

because $N \to \infty$ with better approximations and hence, $[K_H]^{(N)}$ has more eigenvalues. So, the problem is unstable.

From this point, when we write $K_H J_{eff} = E_s$ we will be solving the least-squares problem

$$K_H^* K_H J_{eff} = K_H^* E_s \tag{99}$$

4.1.3 The Scattered Field Model

As a basis for the right-hand side of eq. (51) a set of impulses at M observation points

$$w_m(\underline{x}) = \delta(\underline{x} - \underline{x}_m), \quad \underline{x}_m \in O \quad m = 1, 2, \ldots, M \tag{100}$$

provides an easy evaluation of the inner products on the right-hand side of eq. (81). Thus, this choice leads to a great reduction in the computational time of the solution. These testing functions are chosen so that it is simpler to investigate the difficult problem of overcoming the instability inherent in solving the ill-conditioned matrix problem.

4.1.4 Discretization of K_H

Because the operator K_H is not self-adjoint the application of the Minimal Function Theorem is not straightforward. Hence, the Method of Projections shall be used. We now apply this method.

The object is modelled by N square cells of edge size d on a grid with cell centers at $D = \{\underline{x}_n = (x_n, y_n)\}$ and pulse functions, $P_n(\underline{x})$ (eq. (82)). Following Richmond [1965], we use circular cells of radius $a_e = d^2/\sqrt{\pi}$. The contribution of a single cell n to the scattered field at observation point m is

$$h_{mn} \triangleq \int_0^{2\pi} \int_0^{a_e} H_0^{(2)}(k_A \rho) \rho' \, d\rho' \, d\phi = \begin{cases} \frac{2}{k_A^2}(\pi k_A a_e H_1^{(2)}(k_A \rho) - 2j) & \text{if } \underline{x}_m \text{ is in cell } n \\ \frac{2\pi a_e}{k_A} J_1(k_A a_e) H_0^{(2)}(k_A \rho) & \text{if } \underline{x}_m \text{ is not in cell } n \end{cases} \tag{101}$$

which is analytically tractable. Here, the distance from cell to the observation point is $\rho = |\underline{x}_m - \underline{x}'_n|$, $\underline{x}_m \in O$, $\underline{x}'_n \in D$. The first integral applies if the observation point, $\underline{x}_m$, is at the center of the cell, $\underline{x}'_n$. The scattered field is modelled by M impulses at locations $O = \{x'_m = (x'_m, y'_m)\}$. The set O is called the observation contour.

Hence, the linear system of equations we obtain from eq. (33) is

$$s[K_{mn}][J_{eff}()] = [E_s()] : m = 1, 2, ..., M; \quad n = 1, 2, \ldots, N \tag{102}$$

where

$$s \triangleq \frac{j\pi a_e}{2k_A} J_1(k_A a_e) \in C$$

$$[K_{mn}] \triangleq [H_0^{(2)}(k_A |\underline{x}_m - \underline{x}'_n|) \in C]$$

$$[J_{eff}()] \triangleq [J_{eff}(\underline{x}'_n) \in C, \underline{x}'_n \in D]$$

$$[E_s()] \triangleq [E_s(\underline{x}_m) \in C, \underline{x}_m \in O]$$

From equation (36), the complex relative permittivity can be obtained

$$\begin{aligned} \kappa_D(n) &= \kappa_A + \frac{J_{eff}(n)}{k_0^2 E(n)} \\ &= \kappa_A + \frac{J_{eff}(n)}{k_0^2 (E_i(n) + \sum_{i=1}^{N} K_{in} J_{eff}(i))} \end{aligned} \tag{103}$$

where the integral for the internal scattered field $E_s(\underline{x}'), \underline{x}' \in D$ has been replaced by a summation over the N cells of the object.

4.2 Solving the Linear System of the Discretized Problem

The problem of solving the system of linear equations (102) is now considered in order to compute *physically feasible* solutions. When dealing with measured data, we must assume that the input data may contain noise. The proper model for our problem was described at the beginning of Chapter 3, and formally represents the optimization problem of minimizing a quadratic functional (eq. (55)) subject to non-linear constraints (eq. (54)). Algorithms developed for such problems have been limited to several variables due to the computational load. Thus, more standard methods of solving the problem are explored. Specifically, we use the standard methods of solving a system linear equations as a base: Gaussian Elimination, Least-Squares Minimum Norm, Tikhonov Regularization, and Conjugate Gradients.

A number of methods have been proposed to solve ill-conditioned linear problems by "filtering out" the highly oscillatory components of the computed solution. For example, it is possible to truncate the "smallest" singular values after finding the singular value decomposition (SVD) of the $M \times N$ matrix $[K_H]^{(N)}$, or to drop the "smallest" (in norm) columns of $[K_H]^{(N)}$ in the LSMN method. In either case, the range of the truncated operator $\text{ran}([K_H]^{(N)})$ is no longer the range Y of the original operator, and the null space $\ker([K_H]^{(N)})$ does not contain only $\{0\}$, hence, the solution is not unique. In both these cases, if $\kappa_2([K_H]^{(N)})$ is "large" with respect to machine precision, then the computed solutions can be rather inaccurate (see Graves and Prenter [1978] for SVD truncation, and Bjorck [1967] for choosing the number of columns in the LSMN method.)

4.2.1 Gaussian Elimination (GE) Solutions

Gaussian Elimination is appropriate when the matrix $[K_H]$ is square, i.e. $M = N$, and thus prevents overdetermined ($M > N$) systems of linear equations from consideration. As will be seen later, this is a slight handicap because the κ_2 condition number of the linear system often improves when $M > N$.

The method of Gaussian Elimination has the important property that no general system of linear equations can be solved with fewer arithmetic operations than are required by this method, if and only if the elements of the system are handled only by rows or only by columns (cited in Sarkar et al. [1981, p. 847]).

Gaussian Elimination is an example of a *direct method*. Such methods provide algorithms guaranteed to lead to the exact solution (neglecting roundoff and other errors) after a finite number of mathematical operations on scalars (Householder [1975, p. 91]). However, roundoff error becomes pronounced not only for highly ill-conditioned problems, but also for problems which require a large number of arithmetic operations which are dependent on each other (Sarkar et al. [1981]).

It can be shown that Gaussian Elimination produces a solution that has about $t \log_{10} \beta - \log_{10}(\kappa_\infty(A))$ correct decimal digits in the solution (Golub and van Loan [1984, p. 72].) This method is also normally less sensitive to roundoff errors than the conjugate gradient method (Hestenes [1980, p. 142].)

Gaussian Elimination solutions are computed with LINPACK subroutines ZGEFA and ZGESL (see Dongarra et al. [1979].)

4.2.2 Least-Squares Minimum Norm (LSMN) Solutions

Least-squares minimum norm (LSMN) solutions using the pseudo-inverse for solution of integral equations of the kind we consider (eq. (33)) was proposed by Golub and Kahan [1965]. The purpose of using this technique is to dampen spurious oscillations in the solutions. The LSMN solution to such a problem is unique (see Chapter 3). It can be computed with a number of algorithms, one of which a projection method based on Gram-Schmidt orthogonalization (see Dorny [1980, pp. 382-385]). This algorithm was implemented in FORTRAN (subroutine PSEUDO) by Ney [1983, Appendix B.1] in his Ph. D. thesis, using the numerically more stable Modified Gram-Schmidt (MGS) method. Essentially, at each step, the vector with the largest norm is chosen (the *pivot vector*), and all other vectors

are made orthogonal to it. Hence, vectors of small norm (which are likely to be inaccurate results of a previous subtraction) are avoided until the end of the procedure, thus tending to minimize error propagation (see Dorny [1980, p. 379].) As the procedure is well-explained in the above references, it is not repeat here.

As will be seen in the next chapter, problems which are highly ill-conditioned or when the input data is noisy give LSMN solutions which have elements outside the known bounds (eq. (11)). Because of its nature, it is not possible to incorporate constraints into the LSMN algorithm. For this reason, other algorithms are sought.

In this work, LSMN solutions are computed using subroutine PSEUDO. Columns with norms less than ten times the unit roundoff u (approximately 10^{-16}) the norm of the largest column are truncated.

4.2.3 Tikhonov Regularized (TReg) Solutions

The method of Tikhonov regularization (TReg) (see Tikhonov [1963a, 1963b] or Franklin [1974]) is based on eq. (70) which in matrix form is

$$([K]^H[K] + \alpha[L]^H[L])[f] = [K]^H[g] \tag{104}$$

where $[L]$ is a smoothing operator which is used to constrain the space of possible solutions. It has been proposed in numerous papers as method of solving the ill-conditioned systems associated with ill-posed problems. By suitable choice of α, the resulting linear system can be made well-conditioned and is thus solvable by standard methods such as Gaussian Elimination.

This procedure is not easy to implement unless $[L]$ is a derivative operator $\frac{\partial^k()}{\partial x_i^k}$ of some order, k. In this case, however, the smoothing operator does not reflect the physical situation in object reconstruction problems where discontinuities in the complex relative permittivity of the object may occur.

In this work, subroutine TIKHREGD (shown in Appendix D) is used with the regularizing functional $\Omega(f) = \|f^{(2)}\|_2^2$ (following Phillips [1962]), and the derivatives approximated by finite differences. The solution to the problem is chosen by finding the family of solutions for $\alpha_i = 10^{-i}, i = 1, 2, \ldots$, and choosing the smallest i for which the residual $\|[K][f] - [g]\|_2^2$ is less than the known error in $[g]$, $\|[g] - [g_\epsilon]\|_2 \leq \varepsilon_g$. If $[g]$ is known exactly, then this corresponds to the least-squares minimum norm solution.

It is interesting to note that for $[L] = [I]$ the identity operator, the choice of $\alpha = \sigma_N^2/\sigma_S^2$, where σ_N^2 and σ_S^2 are the "noise" and "signal" powers respectively leads to the *optimum inverse* of Foster [1961]. Such an inverse leads to the pseudo-inverse of section 3.2 in a limiting case of error free data.

4.2.4 Conjugate Gradient (CG) Solutions

The method of conjugate directions is a general method for finding least-squares solutions of operator equations $[K][f] = [g]$ in N variables (see Hestenes [1980, chapter II]). Essentially, it solves the system by N linear minimizations in N suitably chosen directions, called *conjugate directions*. Geometrically, the method is equivalent to finding the center of an N-dimensional ellipsoid given an initial guess on the surface of the ellipsoid. Indeed, the conjugate gradient method (where the direction vectors are mutually conjugate) and the Gaussian Elimination method (where the direction vectors are the N unit vectors $\{e_i\}_{i=1}^N$) are special cases.

The importance of this method is not so much that it converges to a least-squares solution (because it does so rather slowly for ill-conditioned problems as is seen later), but because of its iterative nature there is the possibility of adding constraints at each iteration step. Another advantage is that the algorithm can be started with a given initial guess, and converge to the nearest least-squares solution to the initial guess. For exact input data, the method is guaranteed to converge. Practically, the computed solutions are within "engineering accuracy" of better than 0.1% of the exact solution (see Sarkar [1986]).

We now review the algorithm (10:4) from the classic paper by Hestenes and Stiefel [1952], and then describe a new procedure based on this method. The brackets ([ ]) surrounding matrix/vector quantities will be omitted for clarity.

Consider the finite-dimensional (matrix) linear system

$$Ax = b \tag{105}$$

where $A \in C^{M\times N}, x \in C^N, b \in C^M$, whose exact solution is x. Let $\hat{x}$ be the computed solution, with initial guess $\hat{x}_0$. The conjugate gradient (CG) algorithm finds a sequence of *direction vectors* $\{p_i\}_{i=1}^{\cdots}$ which are *mutually conjugate*, i.e.

$$p_i^H A p_j = 0, \quad i \neq j, \quad i = 1, 2, \ldots \tag{106}$$

and a set of magnitude scalars $\{\alpha_i\}_{i=1}^{\cdots}$ such that

$$\hat{x}_i = \hat{x}_0 + \sum_i \alpha_i p_i \tag{107}$$

The process of finding the next $\hat{x}_{i+1}$ is called an *iteration (step)*, with

$$r_i \triangleq b - A\hat{x}_i \tag{108}$$

being *residual* at the $i-$th iteration.

If A is Hermitian, i.e. $A = A^H$, then this method can be interpreted as minimizing the functional

$$F(x) = \frac{1}{2}x^H Ax - x^H b$$

If A is not Hermitian, as is the case here, then the functional minimized is

$$F(x) = \|x - \hat{x}_i\|_2^2$$

which is the error between the exact solution x and the i-th estimate of $\hat{x}$. The direction vectors are derived from $F'(x)$ at each iteration step in such a way as to minimize the error between the approximate solution x_{i+1} and the projection of x onto the space spanned by the direction vectors $\{p_k\}_{k=1}^{i}$

Algorithm CG

1. Initialization.
   Initial guess $\hat{x}_0$, and termination parameter EPS given
   $r_0 = b - A\hat{x}_0$
   $p_0 = A^H r_0$

2. Iteration. (Restart at Step 1 every R iterations).
   2.1: $\alpha_i = \frac{|A^H r_i|^2}{|Ap_i|^2}$ (see step 3)
   2.2: $\hat{x}_{i+1} = \hat{x}_i + \alpha_i p_i$
   2.3: $r_{i+1} = r_i - \alpha_i A p_i$ (see step 3)
   2.4: $\beta_i = \frac{|A^H r_{i+1}|^2}{|A^H r_i|^2}$
   2.5: $p_{i+1} = A^H r_{i+1} + \beta_i p_i$
   2.6: Terminate if $|\alpha_i p_i| \leq$ EPS.

3. Termination. If $r_{i+1} = 0$, then $A\hat{x}_i = b$ and $\hat{x}_i$ is the solution of the system.
   If $(Ap_{i+1}, p_{i+1}) = 0$, then $Ap_{i+1} = 0$, and it follows that $Ar_{i+1} = 0$. The point $\hat{x}_i$ is thus the point at which $|b - Ax|^2$ attains its minimum, and is thus a least-squares solution.

The residuals $\{r_i\}_{i=1}^{n}$ normally oscillate, and may not decrease until after many iterations (see Hestenes and Steifel [1952, p. 410], Sarkar [1986, p. 9]), but at each iteration are orthogonal to each other.

As mentioned in the 'Iteration' step, it is possible to restart at the 'Initialization' step every R iterations in order to reduce roundoff error which can otherwise accumulate in the inner product computations, specifically from inaccuracies in the computation of the vectors p_{i+1} and r_{i+1} (Hestenes [1980,p. 165], Hestenes and Steifel [1952, p. 420].) In other words, the roundoff error prevents the sequence of vectors $\{p_i\}$ and $\{r_i\}$ from being orthogonal. Even with this correction, increasing the number of iterations indefinitely when computing with finite-precision arithmetic will eventually lead to roundoff error. A characteristic curve for CG performance is shown in Figure 7. Near the minimum of this curve, the computed solution does not change much from iteration to iteration. It is at this point that the algorithm is terminated for our problem (if it is ill-conditioned.) This method is implemented in subroutine CONJGRAD, shown in Appendix D.

The conjugate gradient algorithm finds the nearest least-squares solution x_{NLS} to the initial guess $\hat{x}_0$ (Nashed [1976, p. 381]). Let $P_\eta x$ be the orthogonal projector of x onto $\ker(A)$, and P_S be the closest projector of x onto the set of all least-squares solutions S_{LS}

$$P_S\hat{x}_0 = A^\dagger b + P_\eta\hat{x}_0 \tag{109}$$

Thus, if

$$\|x - P_S x\| = \inf\{\|x - w\| : w \in S_{LS}\} \tag{110}$$

then the computed solution is

$$x_{NLS} = P_S x_0 = A^\dagger b + P_\eta\hat{x}_0 \tag{111}$$

The convergence rate of this method if $\mathrm{ran}(A)$ is closed is at least geometric

$$\|x_n - P_S\hat{x}_0\| \leq c_2\left(\frac{\kappa_2^2 - 1}{\kappa_2^2 + 1}\right)^N \tag{112}$$

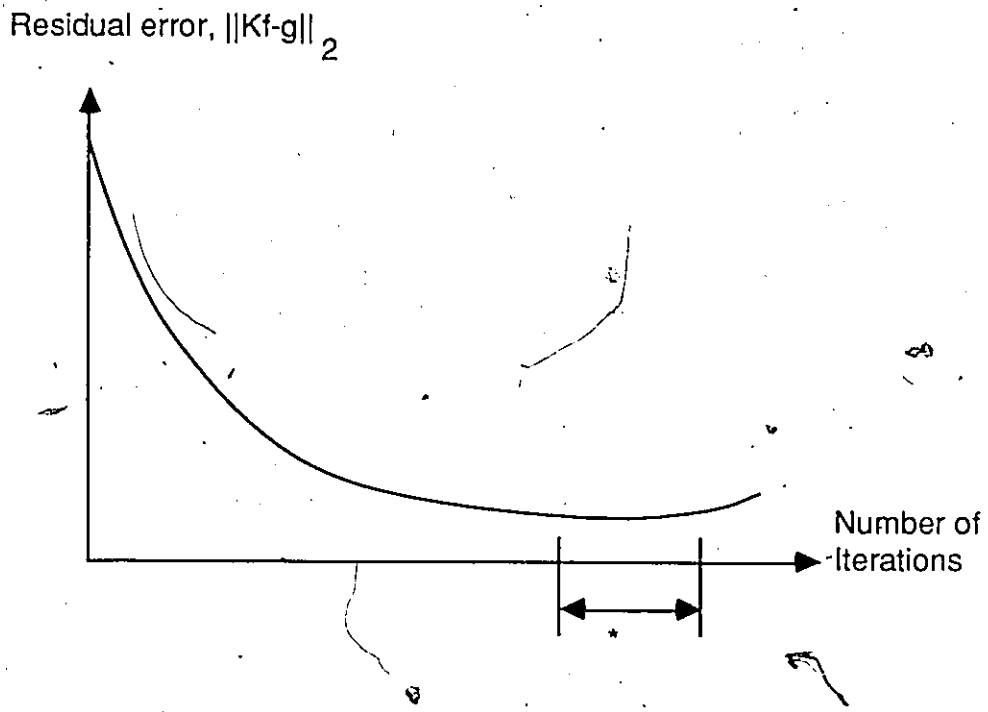


Figure 7: Typical Residual Error Characteristic in the CG Method.
Residual error $\|Kf - g\|_2$ vs. the number of iterations in the Conjugate Gradient Method. The error is the difference between the measured g and the computed estimate $\hat{g} = Kf$. The computed "solution" is taken in the region before the residual error begins to increase (marked by a '*').

where c_2 is a constant (see McCormick and Rodrigue [1975, p. 281].) Hence, if the singular values are not clustered together, as is the case in an ill-conditioned problem, then the convergence rate is slow. Theoretically, this method leads to the exact solution in finite number of steps $\leq N(\sigma)$, where $N(\sigma)$ is the number of independent singular values of A^*A in the finite dimensional space in which the problem is being solved, assuming no truncation or roundoff error (see Sarkar [1981, p. 851].) Whenever we deal with a finite dimensional operator, we are dealing with a closed operator (see Naylor and Sell [1982, pp. 267-268]), hence this geometric convergence rate applies. However, the range of the continuous operator K_{II} from which the discretized problem is derived is not closed because $\sigma_i = 0$ is an accumulation point of the spectrum of $K_{II,LS}$ (see Bertero [1986, p. 58]). Thus, the pseudo-inverse $K_{II}{}^{\dagger}$ is no longer continuous (McCormick and Rodrigue [1975]). The transformation from $\mathrm{ran}(A)$ onto $\ker(A)^{\perp}$ is no longer one-to-one and the convergence rate is no longer geometric but is $O(1/n)$ (see Nashed [1976, p. 381]).

4.2.5 Adding Constraints

Here, we discuss some ways to use the conjugate gradient method for solving ill-conditioned problems.

It is not easy to add constraints to the standard conjugate gradient algorithm because the function we seek $\kappa_D(\underline{x})$ is not the solution of the linear system, $J_{eff}(\underline{x})$, but is a non-linear function of $J_{eff}(\underline{x})$ (eq. (103)). In principle, it is possible to apply such constraints at each iteration step by choosing the "best" direction vector p_i and magnitude α_i. To that end, a function which indicates how far from an admissible solution we are can be used. For this we introduce a *penalty function* $\mathrm{pf}(h)$, where $h_n = \kappa_D(n; J_{eff})$, and $\kappa_D(n; J_{eff})$ is the complex relative permittivity for cell n. Given the physical bounds imposed by eq. (11), the following function penalizes values outside the meaningful physical bounds quadractically (see Figure 8)

$$\mathrm{pf}(h(x)) = \sum_{n=1}^{N} p_{nR} + \sum_{n=1}^{N} p_{nI} \tag{113}$$

where

$$\begin{aligned} p_{nR} &= |Re(h_n(x)) - l_R|^2 \quad \text{if } Re(h_n(x)) < l_R \qquad \text{(infeasible)} \\ &= 0 \qquad\qquad\quad \text{if } l_R \leq Re(h_n(x)) < u_R \quad \text{(feasible)} \\ &= |Re(h_n(x)) - u_R|^2 \quad \text{if } Re(h_n(x)) > u_R \qquad \text{(infeasible)} \end{aligned}$$

and p_{nI} is similarly defined for the imaginary part with $Im()$ replacing $Re()$ and l_I and u_I

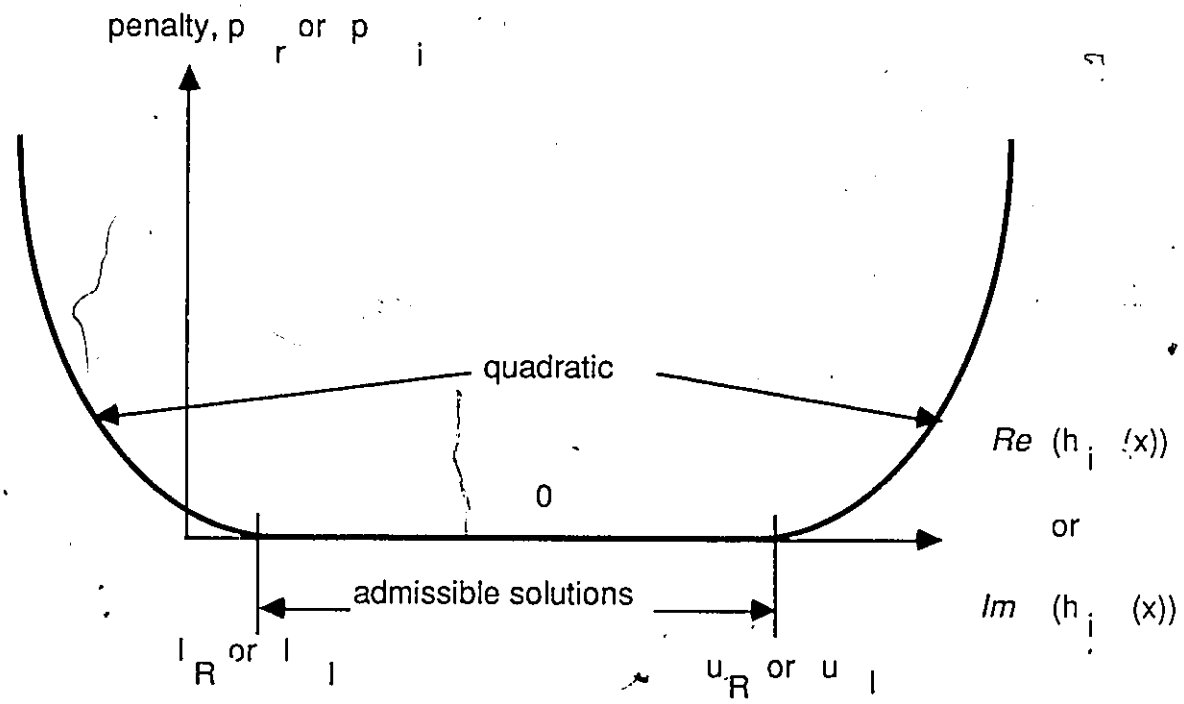


Figure 8: Penalizing "Inadmissible" Solutions using Penalty Functions

replacing the bounds l_R and u_R, respectively, as given in eq. (11). As a simplified procedure we can use the direction p_i computed by the conjugate gradient algorithm, and choose the step size α_i so as to minimize the penalty function, pf(h). Thus, we can choose the most feasible of possible solutions at each iteration. However, the residuals in the conjugate gradient method do not necessarily decrease at the same rate from iteration to iteration. In fact, it often happens that the residual drops several orders or magnitude in the last iterations step(s). (This phenomenon also occurs as columns are added to the solution in the least-squares minimum norm (LSMN) method previously described.) Hence, using the penalty function to guide the update vector $\alpha_i p_i$ is not intuitively reasonable.

After the algorithm converges, it is possible that they are still elements of the complex relative permittivity vector κ_D which are not admissible. They can be *clipped* (by assigning the value of the nearest bound: l_R or u_R for the real parts, and l_I or u_I for the imaginary parts) and the process restarted. We call the procedure of clipping the κ_D and executing the conjugate gradient method until convergence to a given EPS a *pass*. Using this concept, we introduce the following *new method* for computing solutions to the inverse problem $IP(\kappa_D)$.

Algorithm CG with clipping

1. Choose an initial guess $J_{eff(0),0}$.
   Select how well the equations are to be satisfied, ε_{Es}.
   $p = 0$

2 DO UNTIL (convergence occurs)
   $p = p + 1$
   With initial guess $J_{eff(p-1),0}$, execute Algorithm CG until the residual is sufficiently small, $\|K J_{eff} - E_s\|_2 \leq \varepsilon_{Es}$.
   Clip all $\kappa_{D(p)}(n)$ with are outside the bounds in eq. (11) $\rightarrow \kappa_{D(p+1)}(n)$, $n = 1, 2, \ldots N$.
   Solve the direct problem DP(κ_D) : $\kappa_{D(p+1)} \rightarrow J_{eff(p+1)}$ to find the new initial guess.
   ENDDO

At each iteration, one can also impose constraints directly on J_{eff}, the constraints being linear. Here, this implies constraining the effective current

$$J_{eff}(\underline{x}; \theta) = k_0^2(\kappa_A - \kappa_D)E(\underline{x}; \theta) \tag{114}$$

By initially assuming $E(\underline{x};\theta) \approx E_i(\underline{x};\theta)$, i.e. a Born approximation, we can write

$$J_{eff}(\underline{x};\theta) = k_0^2(\kappa_A - \kappa_D)E_i(\underline{x};\theta) \qquad (115)$$

Using this approximation and the bounds on κ_D (eq. (11)), constraints on both the real and imaginary part of J_{eff} can be imposed. Upon solving the linear system for $J_{eff}(\underline{x};\theta)$ and then computing $\kappa_D(\underline{x})$, one can solve the direct problem to find the total field inside the object D, and thus refine (hopefully) the approximation for $E(\underline{x};\theta)$. Computer simulations have shown that this approach does not work well because constraining $J_{eff}(\underline{x})$ does not adequately constrain κ_D, thus leading to solution which are inadmissible.

Chapter 5

INVERSE SCATTERING - Simulated Object Reconstructions

Consider the inverse problem IP(κ_D) which can be described in operator form $K_H J_{eff} = E_s$. The method of projections with pulse basis functions for the unknown (J_{eff} in this case) and impulse basis functions for the measured (E_s) is a standard engineering approach to solving the such a problem, i.e. the *point-matching method.* Although other basis functions may be more appropriate, these bases are pursued because it is important to understand how they work and what their limitations are. As previously explained, such inverse problems are inherently ill-posed, and lead to linear systems which are ill-conditioned. Because the parameters (i.e. number of cells, N, number of observation points, M, etc.) at which the ill-conditioning affects the solution depends on the precision used in computing, the details of the computer/compiler used in this work are described in Appendix C. (Roughly, about sixteen (16) decimal digits of precision are available for computing using VAX-FORTRAN/77 in DOUBLE COMPLEX.)

This chapter has three goals. First, to explore the solutions obtained by discretization of the continuous operator with the point-matching method and solving the resulting linear system by a number of methods. In particular, to see how these methods work with data (i.e. E_s) measured to a various precisions, and how the observation points influence the solution. Second, to further explore the LSMN solution method originally suggested for this inverse problem by Ney [1983]. Third, to propose a new method (conjugate gradient with clipping) which can be applied (with some success) to problems which are highly ill-conditioned.

The model for the digital computer simulations presented in this chapter now follows.

The object D lies in the x-y plane and is discretized into N square cells ($=N_0 \times N_0$ for two-dimensional objects) of cell edge size d. A plane wave $E_i(\underline{x};\theta)$ propagating in the +y-direction (see Figures 1 and 9) irradiates the object at the nominal operating condition f = 3.156 GHz and T=20C (7). The object is immersed in an ambient medium (distilled water) with complex relative permittivity

$$\kappa_A = 76.5 - j11.3 \tag{116}$$

Because the solution depends on the accuracy of the input data, the *input data* is specified by the set

$$I = \{E_i(O), E(O), E_i(D), \text{precision(F)}, \text{precision(O)}\} \tag{117}$$

i.e. the incident fields inside the object D, $E_i(D)$, and on the observation contour O, $E_i(O)$, and the total field on the observation contour, $E(O)$. F signifies the E-field data (E_i or E). Since a network analyzer is used for electrical measurements, F is measured in magnitude [dB] and phase [degrees]. O signifies the position data (x-y coordinates) measured in meters [m]. These input data I are assumed measured to some finite precision, expressed in decimal places [dp] or "full" in the case of full computer/compiler precision. Further, the positions are rounded to precision(O) decimal places in order to simulate a $E-$field probe positioning system of finite precision. (Typically, precision(O) = 6dp or about 1 μm is used in this work.)

The software structure used in the simulations presented here consists of three programs: GEOMEPS, ISI1 and PLOTEPS. Program GEOMEPS solves the direct problem of finding the scattered fields $E_s(\underline{x};\theta)$ on some observation contour O, e.g. a line or a circle, using Richmond's method [1965] (subroutine SCAT). These values are then rounded to precision(F) decimal places *in order to simulate a loss of accuracy*, and then passed to the object reconstruction program, ISI1. ISI1 passes the results of the reconstruction process to program PLOTEPS which produces graphs and images.

Assume that the quantities in I are measured on some sort of ratio meter, and hence the data are normalized with respect to some quantity and are expressed in [dB] for magnitude and [deg] (degrees) for phase. For example, if a network analyzer is used, then the S-parameter S_{21} is measured. In this case, J_{eff}/c_1 is computed where $c_1 \in C$ is a constant, rather than J_{eff}. However, when the complex relative permittivity κ_D is computed from J_{eff}, this scale factor c_1 vanishes. This is discussed in section 5.1. Appendix B describes an experimental $E-$field probing system operating at about 3.156 GHz which was used in this

work. Unfortunately, sufficient accuracy for use in the reconstruction algorithms developed here could not be attained.

In order to compare the original and the reconstructed objects, it is convenient to do so with a scalar reconstruction fidelity metric. For this, the average root mean-square error per cell e_{RMS} is proposed in section 5.2.

The remainder of the chapter discusses reconstructions simulated on a digital computer. Section 5.3 considers one-dimensional problems (i.e. a single line of cells) which are useful in demonstrating the computational problems involved in finding solutions to this ill-posed problem. Section 5.4 considers two-dimensional objects which are considerably more difficult to solve because of the rapidly growing condition number (κ_2) of the problem.

When modelling objects with a square grid of square cells it is not uncommon for some of the cells to have no object in them, e.g. the space between two rods. Such cells (called *ambient cells*) are not, in general, known a priori. In these ambient cells, $J_{eff}(\underline{x};\theta) = 0$. The other cells are called *material cells* and belong to the object. Spacings between the N cells of the object and the $M \geq N$ observation points are given between cell centers. The *width* (x-dimension) is measured as cell edge size d × number of cells in that dimension (N_0), as is the *thickness* (y-dimension). Cells are numbered from lowest x to highest x (left to right) and from lowest y to highest y (bottom to top).

In this chapter, we identify the operator K with K_H, and the functions f with the effective current J_{eff}, and g with the scattered field E_s for notational simplicity. The former set of symbols are often used in this chapter.

5.1 Measurements with a Ratiometer

If the fields F are measured with a network analyzer, a reference signal is needed. If this is done, then all measurements are normalized. Consider a system where all measurements are normalized to a single point (x_{cal}, y_{cal}), where the $E-$field is $E_{cal} \triangleq E_i(x_{cal}, y_{cal})$. Hence, the governing equation is given by dividing eq. (33) by E_{cal}

$$\frac{E_s(\underline{x};\theta)}{E_{cal}} = \int_D \frac{j}{4} H_0^{(2)}(k_A|\underline{x} - \underline{x}'|)\frac{J_{eff}(\underline{x}';\theta)}{E_{cal}} d\underline{x}' \tag{118}$$

where

$$\frac{E_s(\underline{x};\theta)}{E_{cal}} = \frac{(E(\underline{x};\theta) - E_i(\underline{x};\theta))}{E_{cal}} \tag{119}$$

is the difference between two measured quantities. In other words, when eq. (33) is solved $J_{eff}(\underline{x};\theta)$ is scaled by the complex constant E_{cal}.

To see that this scaling does not affect the complex relative permittivity $\kappa_D(\underline{x})$ we divide both the numerator and denominator of the second term of eq. (36) by E_{cal}

$$\kappa_D(\underline{x}) = \kappa_A - \frac{J_{eff}(\underline{x};\theta)/E_{cal}}{k_0^2 E(\underline{x};\theta)/E_{cal}} \tag{120}$$

without affecting the relation between J_{eff} and κ_D.

Thus, normalizing the input data with respect to a fixed point will cause $J_{eff}(\underline{x};\theta)$ to be computed to within a complex constant, but does not affect the computed result for the sought complex relative permittivity $\kappa_D(\underline{x})$.

5.2 Reconstruction Fidelity Metrics

As we shall later see, in most cases we are unable to reconstruct the objects exactly except for certain simulated data, usually with input data I of high accuracy. Hence, it is convenient to find a scalar metric for the fidelity of reconstruction which indicates how close the reconstruction is to the original.

A number of such metrics have been used in the field of digital image processing. Many of them use statistical data about the images, e.g. gray-level histograms. As objects as small as 25 cells are often difficult to reconstruct, sufficient data is not available to supply reliable statistics. Hence, these metrics are excluded here.

The most common metric is the *root-mean-square error* per "pixel" or cell $e_{RMS}(f)$. If $f(i,j)$ is the original $M \times N$ function, and $\widehat{f}(i,j)$ is the reconstructed function, then

$$e_{RMS}(f) \triangleq \frac{\sum_{i,j} |f(i,j) - \widehat{f}(i,j)|^2}{M \times N} \tag{121}$$

where the summation is over all cells. (When the argument (f) is omitted, we assume the function is the complex relative permittivity, κ_D.) Although this metric is known to be deficient in many cases, e.g. a constant offset in the data would lead to a large error, it is still universally used. For errors in the effective current J_{eff}, it is more appropriate to use the *normalized root-mean-square error*

$$e_{NMS}(f) \triangleq \frac{\sum_{i,j} |f(i,j) - \widehat{f}(i,j)|^2}{\sum_{i,j} |f(i,j)|^2} \tag{122}$$

because the J_{eff} tend to large in magnitude (typically, about 10^5) and this metric gives a relative error. Slaney and Kak [1984, p. 870] have used e_{NMS} for object reconstruction of the complex refractive index $n = \sqrt{\kappa_D}$, but since the metric depends on the complex relative permittivity of the ambient medium, it is not suitable in this work.

e_{RMS} is used because of its simplicity, but it is well known that this metric may not be correlated with the ability to extract relevant information from the reconstruction (called *image intelligibility*). In the problems considered here, it has been empirically observed for simple objects, i.e. those with a few material cells, the structure can often be deduced with error e_{RMS} of the order one. For more complex objects, errors e_{RMS} of the order of 0.1 or more given reconstructions which bear little relation to the original objects, or have dielectric constants which are clearly unreasonable, e.g. $\kappa'_D < 0$.

5.3 One-dimensional Objects

In this section, we explore the problems involved in reconstruction by considering objects which can be modelled by a single line of cells. We refer to such objects as *one-dimensional* even though they have a width of one-cell, d. Computed fields in I are normalized to E_i at (x_{cal}, y_{cal}) = (0,-10) mm, and expressed in [dB] and [degrees].

For simplicity, we consider two types of objects because they represent extremes. Objects which are of high contrast with respect to the ambient medium (acrylic, $\kappa_D = 2.6 - j0.0$) generally lead to more difficult numerical problems, because of the larger magnitude of effective current J_{eff}. (See the definition of J_{eff}, eq. (35).) A low-contrast object (sea water, κ_D = 75.5 - j18.1) is used to demonstrate that object of low contrast generally have slightly better solutions, because of the smaller effective current J_{eff}.

5.3.1 Number of cells and the condition number

Consider a line of N cells of edge size d packed into a 40 mm segment along the x-axis as shown in Figure 9 (Problem P-1A). The scattered field E_s is measured at $M = N$ observation points equally spaced along a 70 mm segment at y = 10 mm.

Objects of N = 8, 16, 32, and 64 cells, are packed into the same width of 40 mm. Two of the cells are made of acrylic and the others are ambient cells as shown in Table 2. This table also shows the cell edge size d, as well as the condition number κ_2 of the resulting linear system, and the error e_{RMS} in reconstructing with the LSMN algorithm. (These

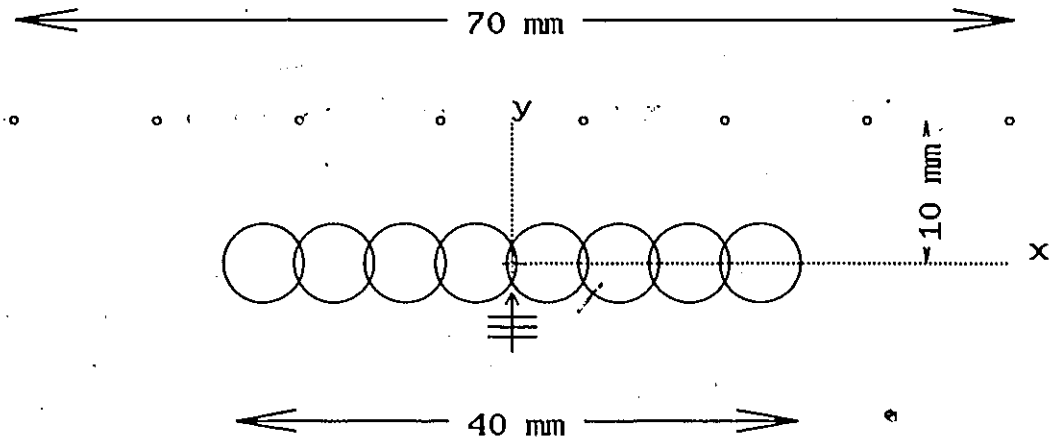


Figure 9: Geometry for PROBLEM P-1A
The object cells are on a 40 mm segment, and the observation contour is a 70 mm segment on y = 10 mm. The number of cells N is varied. In this figure, $N = 8$ cells. Fields are normalized to (0,-10 mm).

No. of cells N	Material cell numbers	Cell edge size d [mm]	κ_2 (ZSVDC)	e_{RMS} (LSMN)
8	4,6	5	2.7×10^{1}	4.1×10^{-11}
16	8,10	2.5	8.4×10^{6}	1.5×10^{-3}
24	12,14	1.667	6.5×10^{10}	5.1
32	16,18	1.25	3.6×10^{13}	0.39
64	32,34	0.625	6.1×10^{14}	0.44

Table 2: Effect of Number of Cells on the Condition Number
$M = N$ observation points with the middle cells being acrylic. (Simulation, LSMN, precision(O)=6dp, precision(F) = full.)

errors are examples of what may occur and are not necessarily the best attainable.) Note that after the condition number κ_2 exceeds about 10^8, the error e_{RMS} increases rapidly and then decreases. This phenomenon has not been investigated, but may be due to the influence of the final column in the LSMN solution when columns are truncated (see section 4.2.2.)

As the cell edge size decreases, hence the number of cells in a given width increases, the condition number is seen to increase. Recall from Chapter 4, that as the object is better approximated (i.e. the cells are more densely packed) the condition number is expected to increase. (Actually, we have not shown this exactly as we have neglected the thickness d of the object in order to avoid having more than one row of cells. This problem is considered later.) *A very large condition number with respect to machine precision is not in itself an insurmountable problem, because the algorithm may be simply executed on a machine with more precision. However, increasing the density of the cells will increase the condition number* κ_2.

5.3.2 Number of impulse basis functions for E_s

In this section, it is demonstrated that the set of basis functions for the scattered field must be adequate to represent it. By choosing basis functions improperly, a problem other than the original may be solved, thus leading to solution which is incorrect in the sense that it does not correspond to the original physical problem. Now, we fix the contour of the observation points to a line segment, and vary the number of observation points M along this segment.

Consider again Problem P-1A (Figure 9) with N cells and M equally spaced observation points along a 70 mm long segment. Figure 10 shows the change in both the κ_2 condition number and the error e_{RMS} when the object is reconstructed using the LSMN method. The horizontal (x) axis M/N gives an indication of how overdetermined the system of linear equations is. In the upper figure, note that the condition number κ_2 saturates at about $M = 1.5N$. (Recall that increasing the number of cells N leads to an increased condition number. κ_2 could not be computed with the LINPACK subroutine ZSVDC for $N = 64$ which is a very ill-conditioned problem.) In the lower figure, the LSMN error e_{RMS} tends to saturate at about $M = 2N$. This can be explained theoretically by noting that some sets of impulses do not provide a sufficiently complete set of basis functions to model the scattered field E_s. Reconstructions with the LSMN algorithm for $N = 8$ cells are shown in Table 3, with input data of precision(F)=full. Figure 11 regraphs the data of Figure 10 to show that the trend in increased κ_2 condition number is to worsen the solution (in terms of e_{RMS}) for the LSMN method.

Therefore, increasing the number of observation points from $M = N$ to about $M = 1.5N$ can improve (decrease) the condition number, κ_2. Further, at least for LSMN solutions, this tends to decrease the error e_{RMS}.

5.3.3 Methods of solution and Input data precision

Figure 12 shows the e_{RMS} errors in the solutions computed by the Gaussian Elimination (GE) (LINPACK subroutine ZGESL), LSMN (Ney's [1983] subroutine PSEUDO), Tikhonov regularization (with $\Omega(f) = \|f\|_2^2$) (subroutine TIKHREGD) and the conjugate gradient (CG) method (subroutine CONJGRAD) with 10000 iterations. The e_{RMS} error is shown against the number of digits of precision in the input data for the total and incident fields, F. PROBLEM P-1A with $M = N$ observation points is used so that Gaussian Elimination

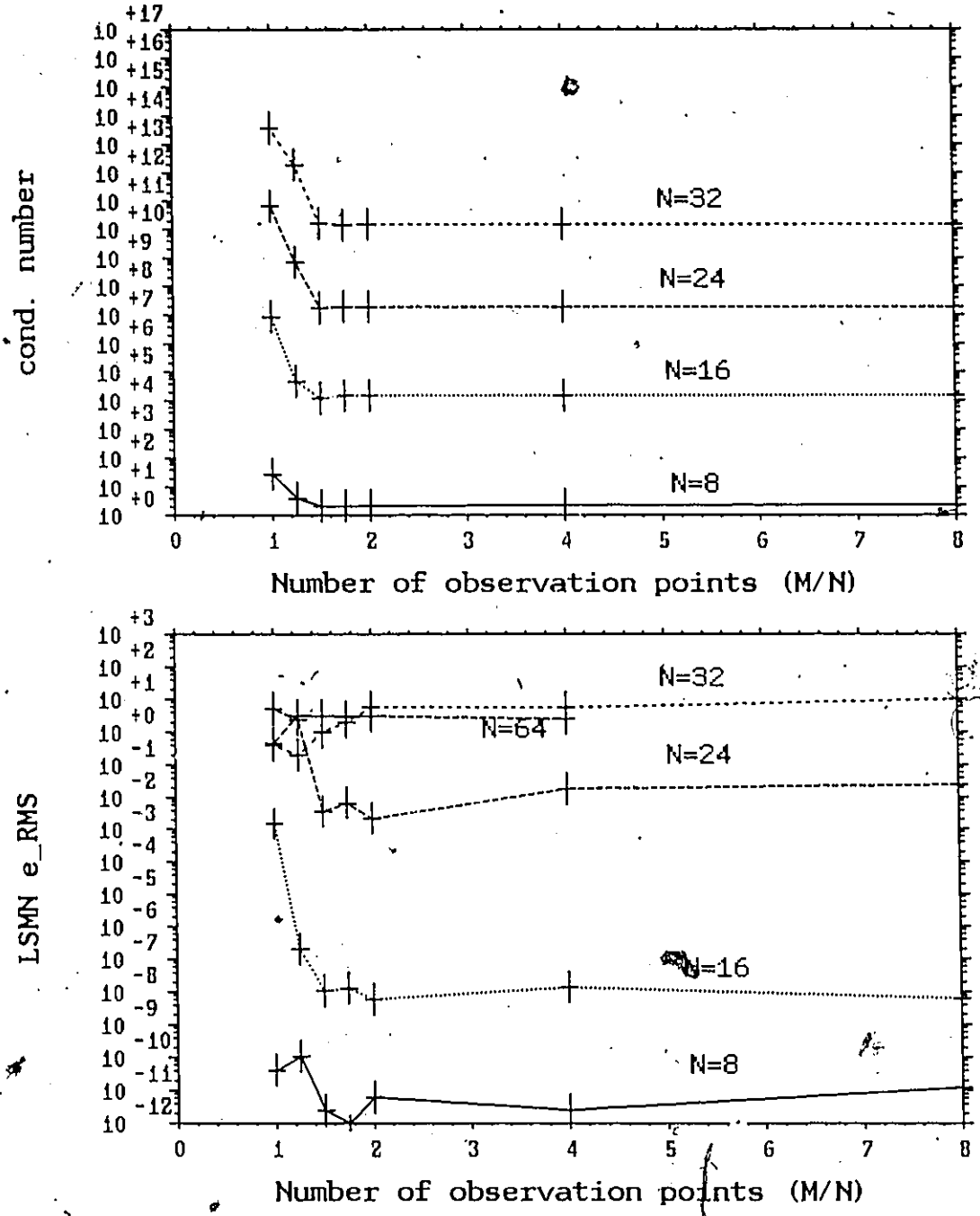


Figure 10: Effect of the number of observation points, M, on the computed solutions An example using PROBLEM P-1A with M observation points. In the upper figure, the condition number κ_2 is shown against the number of observation points M normalized with respect to the number of cells, N. This indicates how "overdetermined" the resulting system of $M \times N$ linear equations is. In the lower figure, the object consists of two high contrast (acrylic) cells in the middle of the "object", separated by a single cell. (Simulation, LSMN, precision(O)=6dp, precision(F)=full)

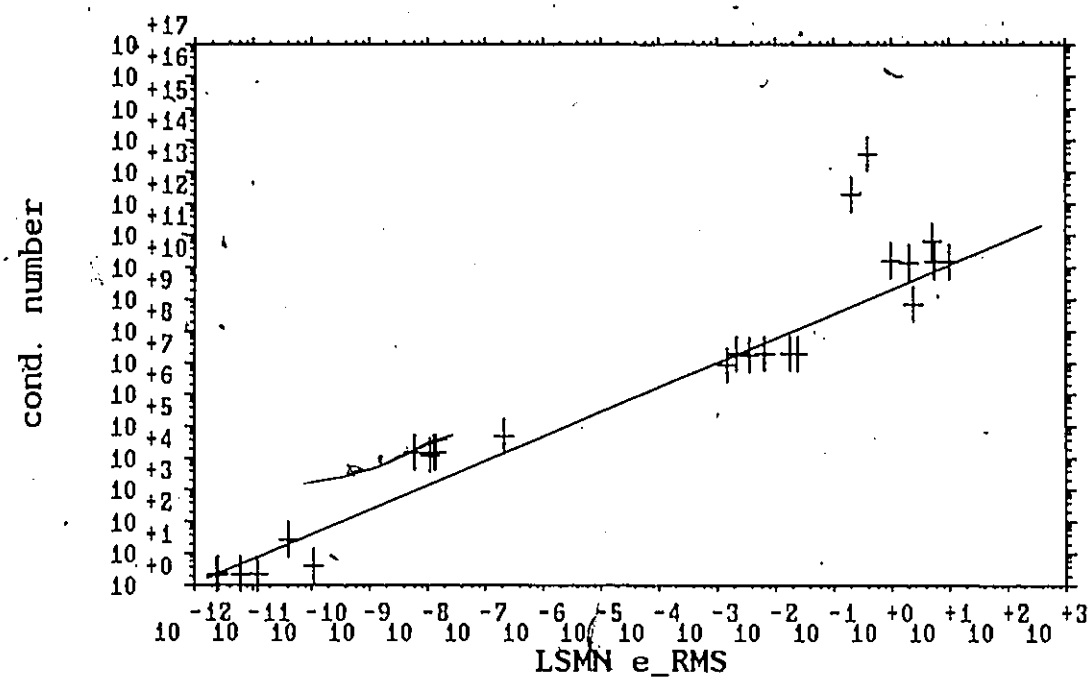


Figure 11: Condition number κ_2 and the computed LSMN error, e_{RMS}
This figure shows the data in Figure 10 redrawn to shown the influence of the condition number on the solution computed by the LSMN method. (Simulation, LSMN, precision(O)=6dp)

No. of observation points, M	κ_2	Computed Solutions $\kappa'_D = 2.60$	$\kappa''_D = 0.00$	$\kappa'_A = 76.50$	$\kappa''_A = -11.30$	e_{RMS} (κ_D)
8	27	2.60	0.91×10^{-10}	76.50	-11.30	0.41×10^{-11}
16	2.2	2.60	-0.38×10^{-10}	76.50	-11.30	0.61×10^{-11}
32	2.3	2.60	-0.12×10^{-3}	76.50	-11.30	0.25×10^{-11}
64	2.3	2.60	0.71×10^{-10}	76.50	-11.30	0.12×10^{-10}
128	2.3	2.60	-0.47×10^{-10}	76.50	-11.30	0.77×10^{-11}
256	2.3	2.60	0.12×10^{-11}	76.50	-11.30	0.11×10^{-11}

Table 3: Effect of varying the number of impulses, M, within a segment
Tabulated are the κ_2 condition number and the computed values for both the acrylic cells (D) and the ambient cells (A) for example PROBLEM P-1A with $N = 8$ cells.
(Simulation, LSMN, precision(O)=6dp, precision(F)=full.)

can be used in the comparison. (All the other methods can work with $M \neq N$.) The three curves in each graph represent a low condition number problem ($\kappa_2 = 27$, $N = 8$ cells), a mid condition number problem ($\kappa_2 = 8.4 \times 10^6$, $N = 16$ cells), and a high condition number problem ($\kappa_2 = 3.6 \times 10^{13}$, $N = 32$ cells). *If the input data is not perfectly accurate (which can be simulated by data of finite precisions), then there may be a number of solutions satisfying* $\|Kf - g\| \leq \varepsilon_g$ *where the error in* g, $\varepsilon_g \triangleq \|g - g_e\|$ *is the norm of the error in the input* g, as noted by Phillips [1962].

For the low κ_2 condition number (with respect to machine precision) problem, it is clear that any of the methods give excellent results, with the input field data precision as low as two (2) decimal places.

For the mid condition number problem, Gaussian Elimination gives the best results, but the LSMN results are excellent as well. These methods can be used with the input field data precision as low as six (6) decimal places. The conjugate gradient (CG) method has convergence problems, probably due to the computing precision available.

For the high condition number problem, all methods give bad results due to the lack of computing precision. For both mid and high condition problems, the Tikhonov regularization methods give errors e_{RMS} which are less dependent on the input field data precision.

Figure 13 shows similar graphs, except that the material cells are made of the low

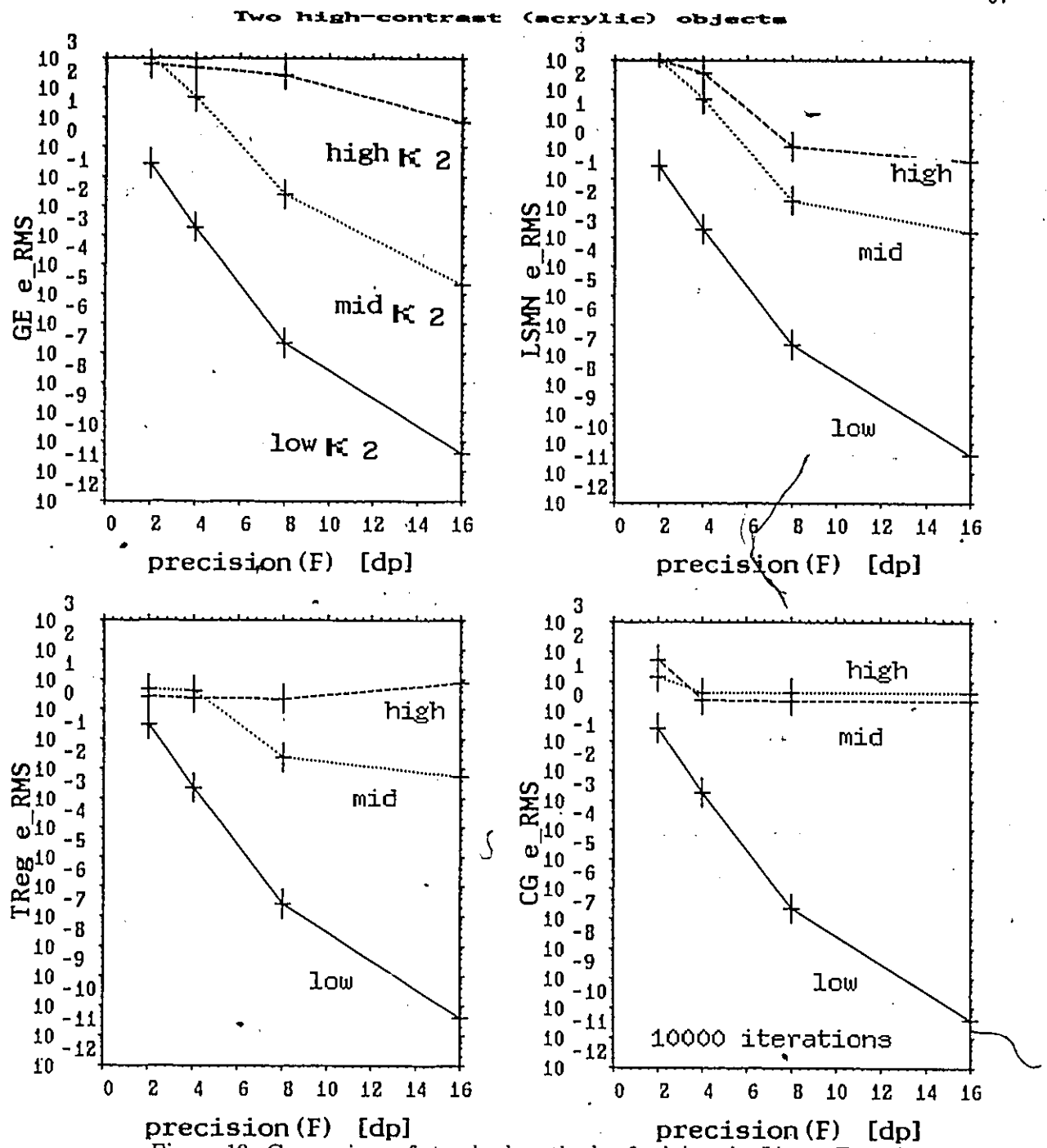


Figure 12: Comparison of standard methods of solving the Linear Equations
An example using PROBLEM P-1A with $M = N$ observation points and two high contrast (acrylic) cells in the middle of the "object" (see Table 2.) The three cases represent a *low* condition number $\kappa_2 = 27$, a *mid* condition number $\kappa_2 = 8.4 \times 10^6$, and a *high* condition number $\kappa_2 = 3.6 \times 10^{13}$. The set F is the set of field measurements E and E_i represented in magnitude [dB] and phase [degrees]. (Simulation, precision(O)=6dp)

contrast sea water. The curves are essentially the same as for Figure 12, except that e_{RMS} errors are slightly lower.

For a 'feel' of the meaning of the magnitude of the error e_{RMS} refer to Figure 14. Shown are least-squares minimum norm (LSMN) solutions for input fields precisions(F) = 8 dp, 4 dp, 2 dp and 1 dp. The precision(F) = 8 dp solutions gives the same graph as for the original. Note that the 1 dp solution has error $e_{RMS} = 0.63$ and is an unacceptable reconstruction, while the 2 dp solution with error $e_{RMS} = 0.27$ is acceptable. *Thus, we require e_{RMS} to be less than about 0.3.*

Figure 15 compares the convergence rate of the conjugate gradient (CG) method with various M. Note that $M = N$ observation points seems to give the best convergence, i.e. minimum residual $\|Kf - g\|_2$.

Figure 16 compares the convergence properties of the conjugate gradient method (CG) and the Algorithm CG with with clipping method using PROBLEM P-1A as an example. For the case $N = 8$ (condition number $\kappa_2 = 27$), the residual using either method is very small, being of $O(10^{-17})$. Due to the overhead involved in computing the direct problem in order to apply the constraints in the conjugate method with clipping method, the clipping process increases the CPU time involved. However, computation time is small in both cases (about 10 CPU seconds). For the mid condition problem $N = 16$ (condition number $\kappa_2 = 8.4 \times 10^6$), the standard conjugate gradient iterations converge very slowly. Using about 1000 CPU s, the clipping process gives a smaller residual. The same is true for the high condition problem $N = 16$ (condition number $\kappa_2 = 3.6 \times 10^{13}$). Hence, we conclude that clipping used with the standard conjugate gradient iterations accelerates convergence for this problem.

Therefore, for low condition number with respect to machine precision problems, all methods of solution give excellent results, even with precisions as low as two (2) decimal places in magnitude [dB] *and phase* [degrees]. *However, none of the methods perform well with high condition number problems with respect to machine precision.*

5.3.4 Resolution

Consider the case of a line of $N = 32$ cells of edge size 1 mm along the x-axis as shown in Figure 17 (Problem P-1B). Here, the $M = N = 32$ observation points are along the line y = 10 mm. The condition number for this problem $\kappa_2 = 9.3 \times 10^{11}$ is very high with respect to machine precision. (Recall that κ_2 gives an upper bound to the sensitivity of the solution

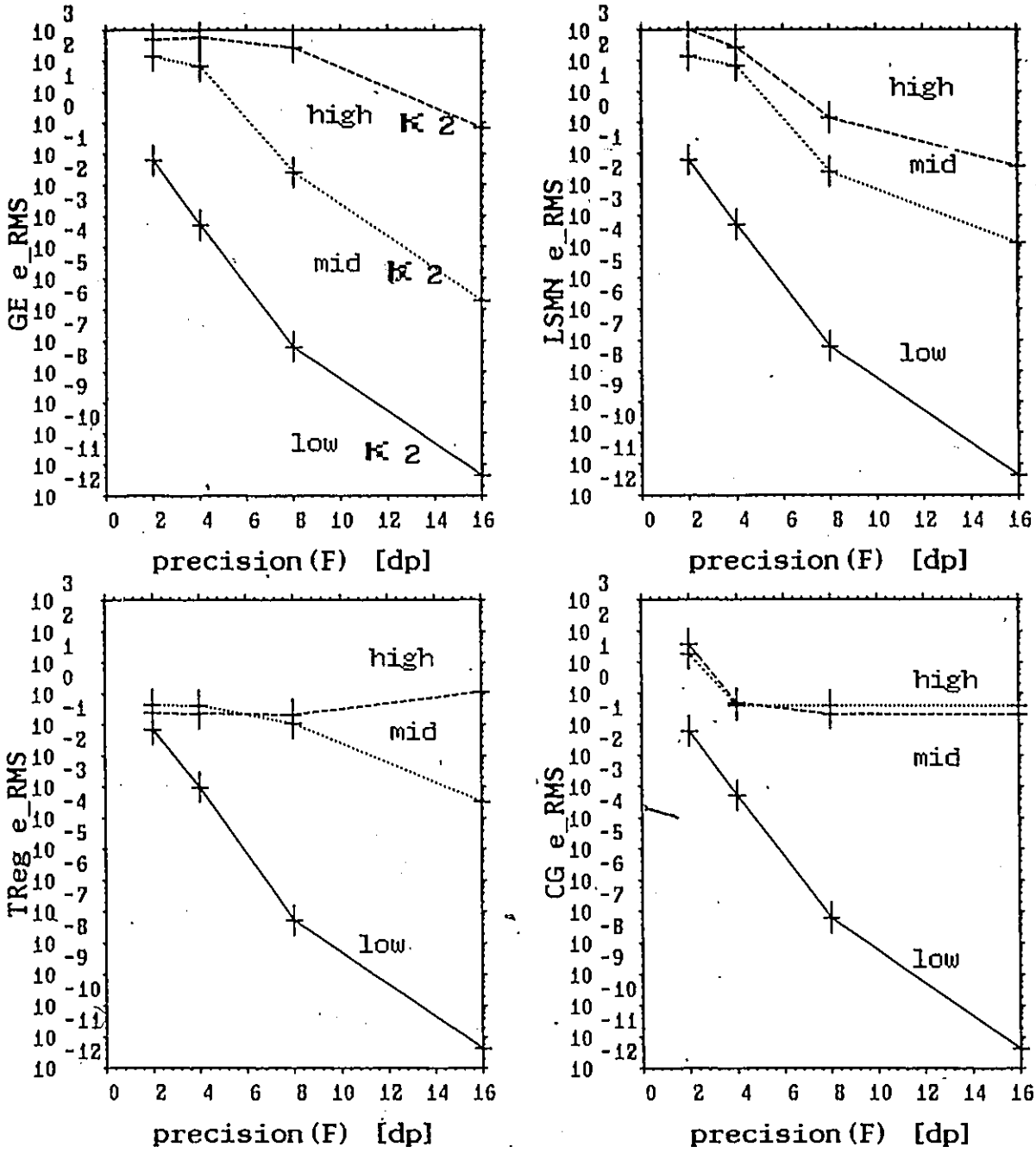


Figure 13: Comparison of standard methods of solving the Linear Equations. An example using PROBLEM P-1A with $M = N$ observation points and two low contrast (sea water) cells in the middle of the "object" (see Table 2.) The three cases represent a *low* condition number $\kappa_2 = 27$, a *mid* condition number $\kappa_2 = 8.4 \times 10^6$, and a *high* condition number $\kappa_2 = 3.6 \times 10^{13}$. (Simulation, precision(O)=6dp)

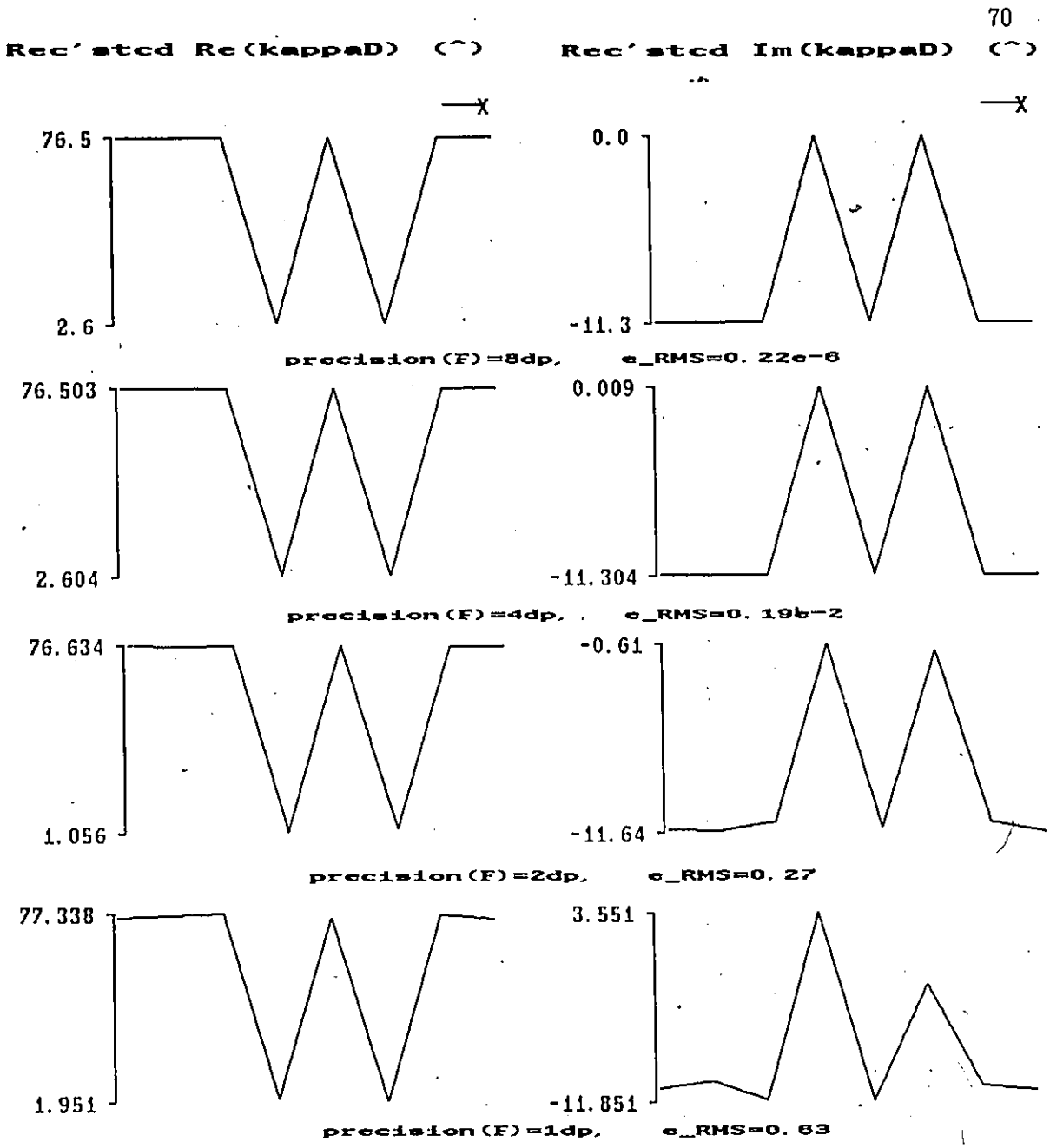


Figure 14: LSMN Solutions of increasing error e_{RMS}

PROBLEM P-1A with $M = N = 8$ observation points for input E-field precisions 8 dp, 4 dp, 2 dp and 1 dp in [dB] and [degrees]. From top to bottom, are shown solutions of *increasing* e_{RMS}. Note that for e_{RMS} less than approximately about 0.27, there is little difference between the graphs of the computed solution and the exact solution. (Simulation, LSMN, precision(O) = 6dp.)

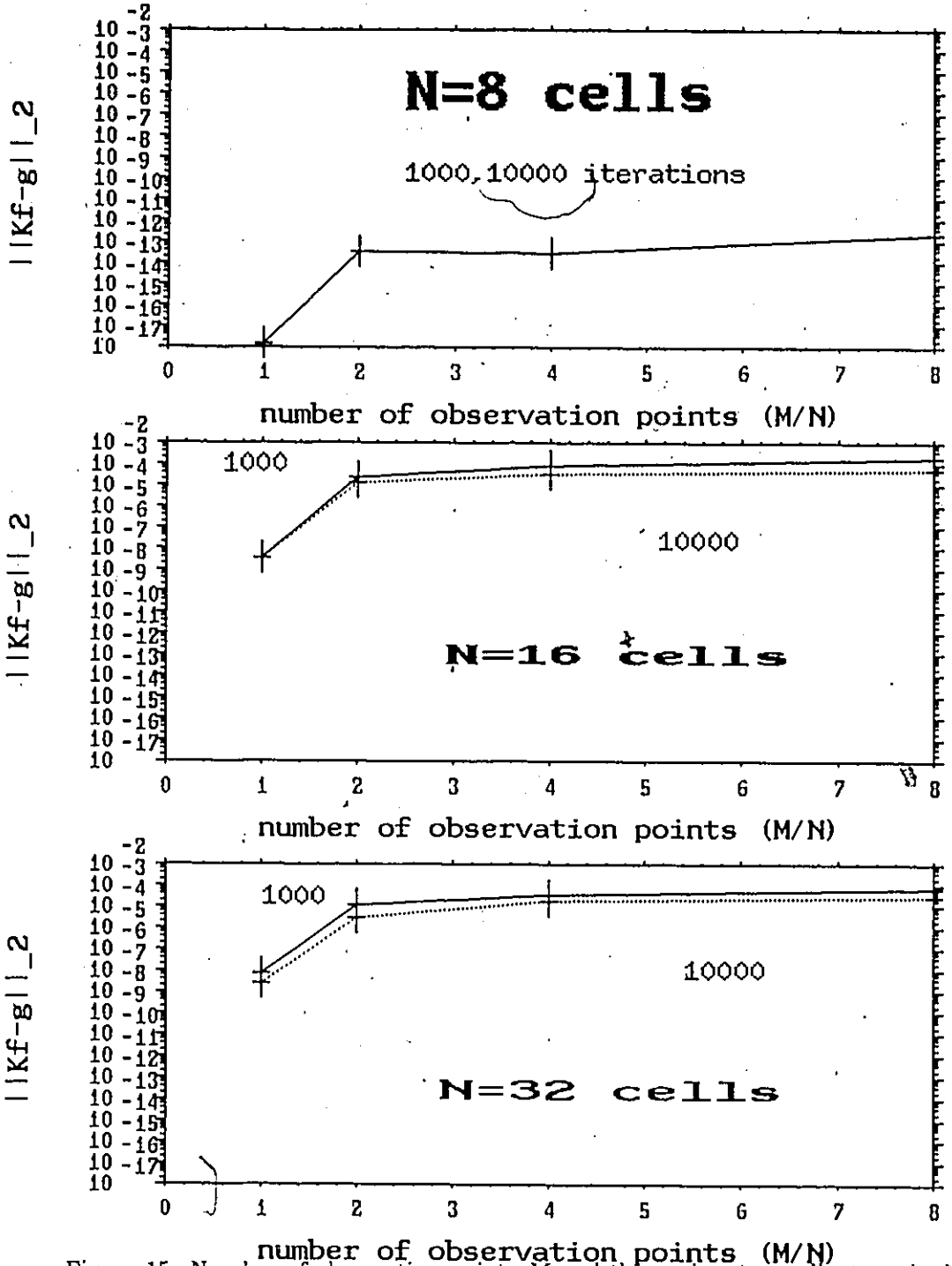


Figure 15: Number of observation points M and the conjugate gradient method
A maximum of 10000 iterations were used in the comparison with PROBLEM P-1A. Note the increasing residual $\|Kf-g\|_2$ as the number of observation points M is increased. With the conjugate gradient method, there is no benefit in considering more observation points than the number of cells N. (Simulation, CG, precision(O)=6dp)

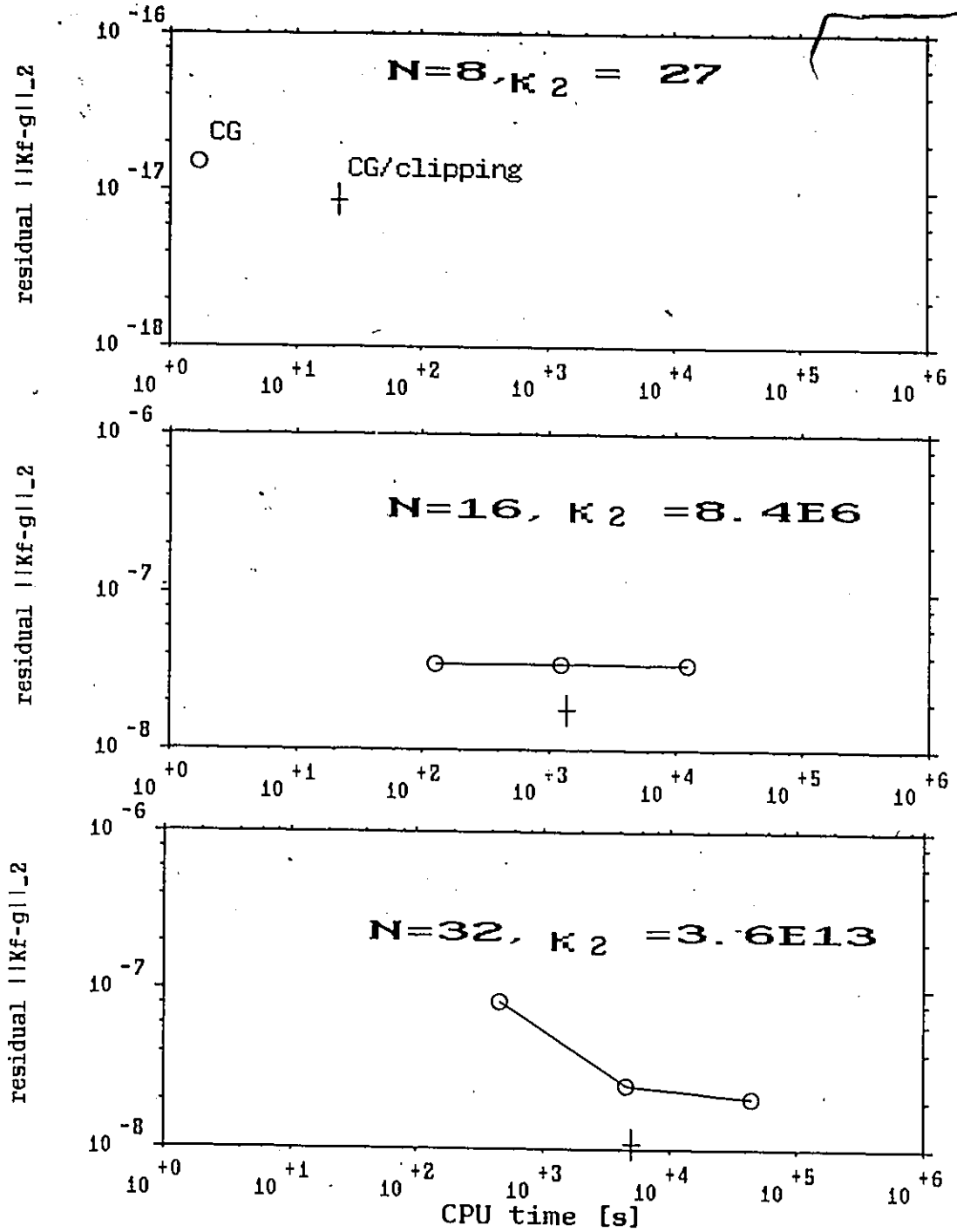


Figure 16: Comparison of convergence in two conjugate gradient based methods Residuals of the solutions to PROBLEM P-1A computed by the standard conjugate gradient method (o), and by the conjugate gradient method with clipping (+), $\varepsilon_g = 10^{-12}$ as CPU time is increased. The top, middle and bottom figures represent problems of increasing condition number, κ_2. Note that with clipping (+); the attained residual $\|Kf - g\|_2$ is smaller. (Simulation, precision(O)=6dp, precision(F)=full.)

Solution method No. of iterations	e_{RMS}	Residual $\|Kf-g\|_2$	CPU time [sec]
CG method, EPS=1×10^{-16}			
1000	4.2	3.50×10^{-8}	127
10000	4.2	3.50×10^{-8}	1235
100000	4.2	3.50×10^{-8}	12340
CG with clipping method, EPS=1×10^{-12}			
10 passes/MAXITER=100	2.146	1.80×10^{-8}	195
10 passes/MAXITER=1000	2.146	1.80×10^{-8}	1365

Table 4: Comparison of solutions for two conjugate gradient based methods for PROBLEM P-1A
N = 16 cells, and the condition number κ_2 is 8.4×10^6. As explained in Appendix D, EPS and MAXITER are used as termination criteria. Note that the error e_{RMS} in the conjugate gradient method does not decrease from 4.2 despite the use of much more CPU time, but, with clipping the error decreases from its initial value of 4.2 to 2.146. (Simulation, CG and CG/clipping , precision(O) = 6dp, precision(F) = full.)

to the input data.) The two acrylic cells are separated by a distance of s mm in order to study the resolution for this configuration.

Figure 18 show the scattered fields E_s along the observation contour for the configurations s = 31, 7, 5, 3, and 1 mm. Reconstruction was performed with **Algorithm CG with clipping** and conjugate gradient termination parameter EPS = 10^{-12}. (Using the conjugate gradient method alone gives larger residuals and larger errors e_{RMS} regardless of the number of iterations applied as shown in Table 4.) Table 5 summarizes the reconstruction process. The 'Resolved in ? passes' is only an approximate measure because of the termination criterion used (see section 4.2). Figure 19 shows the convergence patterns

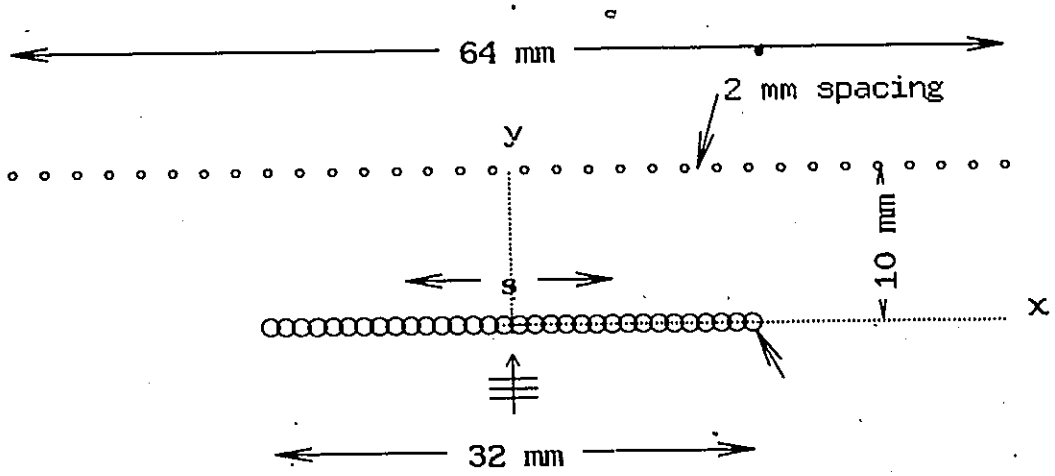


Figure 17: Geometry for PROBLEM P-1B

Two acrylic rods ($\kappa_D = 2.6 - j0.0$) separated by a distance s. This problem has a very high condition number ($\kappa_2 = 9.3 \times 10^{11}$) with respect to machine precision ($u\kappa_2 \approx 2.6 \times 10^{-5}$.) Dimensions are between the outside edges of the structures. Fields are normalized to (0,-10 mm).

Separation s [mm]	Resolved in ? passes	e_{RMS} (CG/clip)	Residual $\|Kf - g\|_2$	Computation Time (CPU seconds)
31	20	0.012	9.4×10^{-10}	811
7	1000	0.746	5.0×10^{-9}	104094
5	1000	0.458	2.0×10^{-9}	114808
3	2000	0.245	4.2×10^{-9}	348501
1	500	0.172	1.6×10^{-9}	41382

Table 5: Resolution Test for Algorithm CG with clipping
Example using Problem P-1B. See Figure 20 for the solutions. (Simulation, Algorithm CG with clipping, precision(O) = 6dp, precision(F) = full.)

for the problem, and Figure 20 shows the object reconstructions. Note that not only are the objects well resolved, but the values are very nearly those of the original except for the imaginary part of the s=7 mm case.

Although a large amount of computing time is required, objects with a separation of only $s = 3$ mm $< \lambda_A/2 = 5.7$ mm can be reconstructed. This demonstrates that the object is not spatially bandlimited to $\lambda_A/2$ as assumed by other researchers, e.g. Pichot et al. [1985].

Although it requires a large amount of CPU time, the **Algorithm CG with clipping** *method can be used to resolve objects better than $\lambda_A/2$. Further, it gives better results than that obtained by any of the four standard methods.*

5.3.5 Summary - One-dimensional objects

No matter what method is used to solve the system of linear equations resulting from discretization, the error e_{RMS} in the solution depends on the accuracy of the input data for two reasons. First, because the continuous problem itself is inherently ill-posed, the solution will be unstable with respect to small errors in the input data. Second, the methods of solving

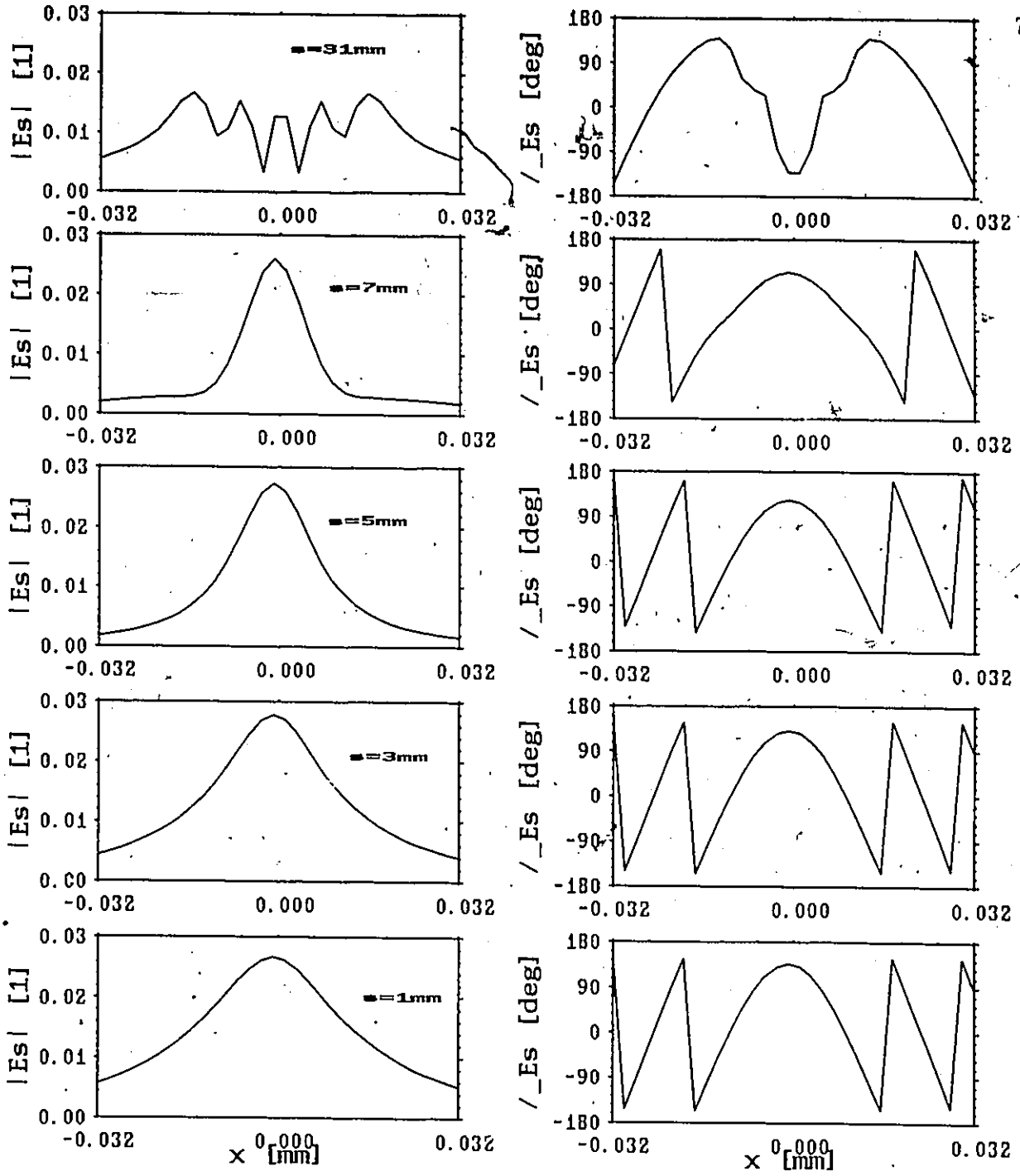


Figure 18: Simulated scattered fields $\dot{E}_s$ for PROBLEM P-1B
The magnitude [1] and phase [degrees] of the scattered E-fields are shown for two acrylic cells ($\kappa_D = 2.6 - j0.0$) separated by a distance s[mm]. (Simulation.)

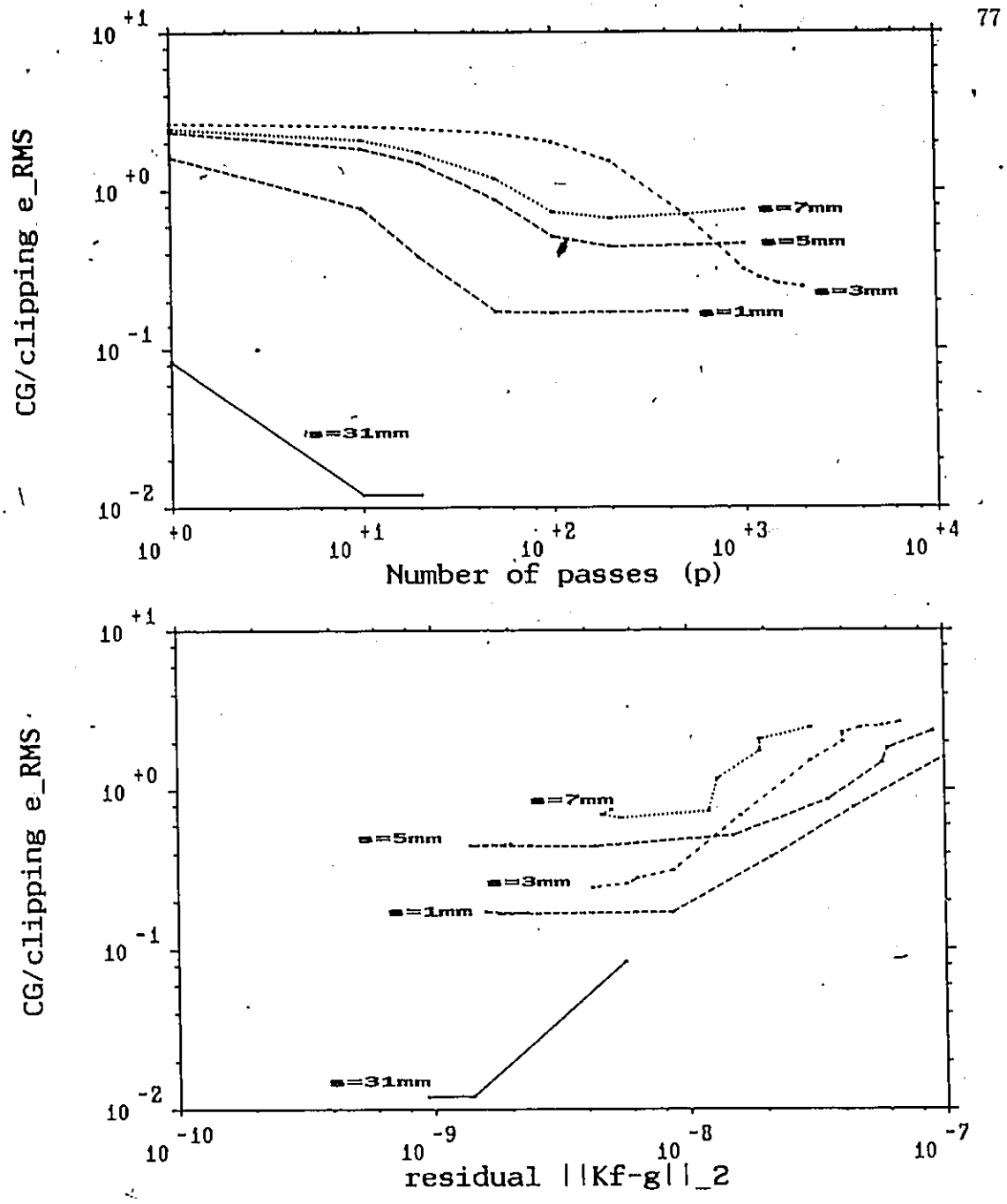


Figure 19: Convergence of a highly ill-conditioned problem using the conjugate gradient with clipping method

PROBLEM P-1B with two acrylic ($\kappa_D = 2.6 - j0.0$) cells separated by s mm in an ambient medium with $\lambda_A/2 = 5.7$ mm. The condition number $\kappa_2 = 9.3 \times 10^{11}$. The upper graphs shows that the method converges sufficiently to resolve the two cells for all values of s. Certainly, the case $s = 3$mm can be resolved. The bottom graph shows that the error in the solution (e_{RMS}) decreases with the residual $||Kf - g||_2$. (Simulation, precision(O)=6dp, precision(F)=full.)

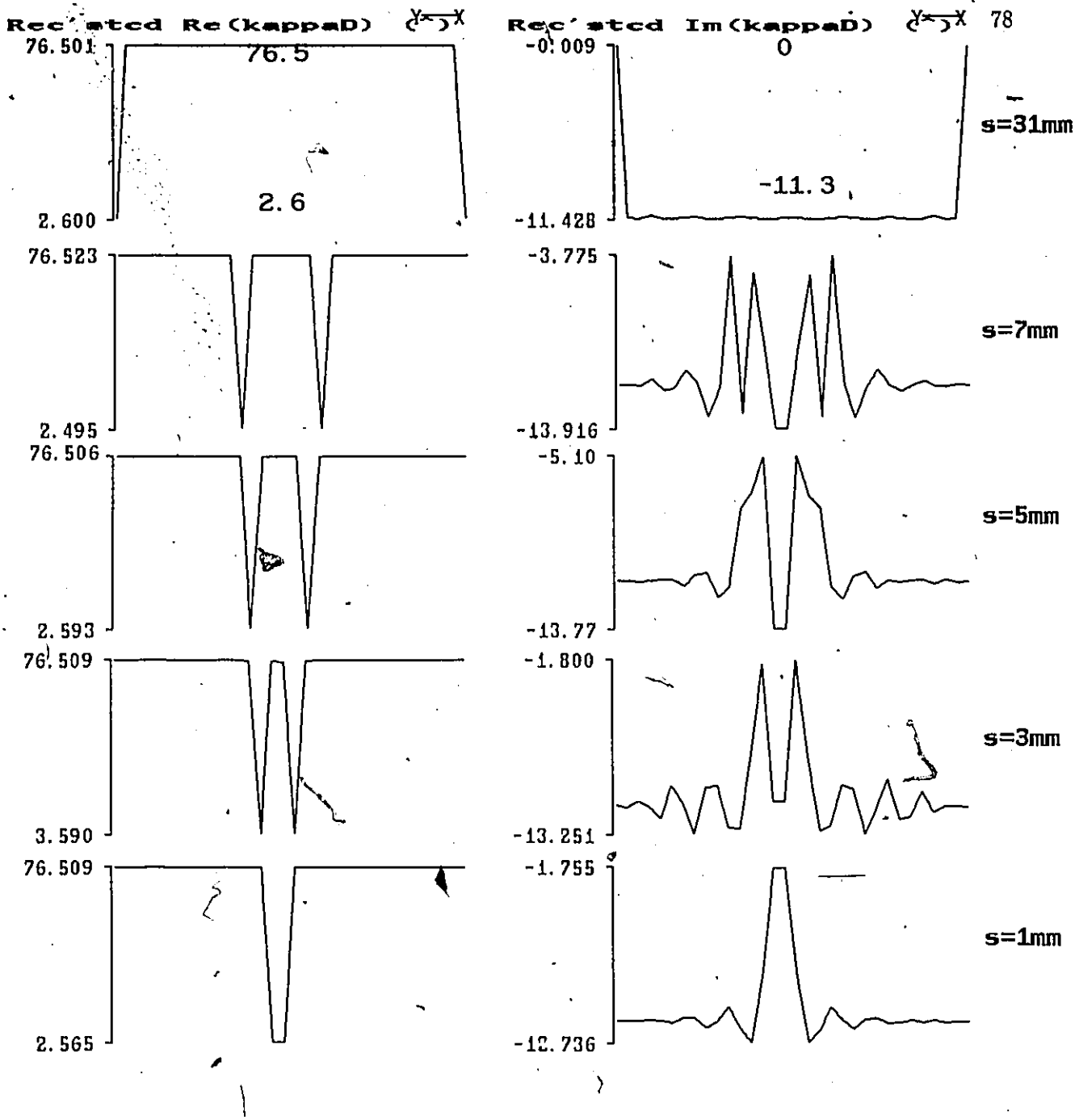


Figure 20: Reconstructions for objects separated by s [mm]
Object reconstruction using PROBLEM P-1B with the two acrylic ($\kappa_D = 2.6 - j0.0$) cells separated by s mm. The condition number $\kappa_2 = 9.3 \times 10^{11}$ in all cases. The ambient medium has $\lambda_A/2 = 5.7$ mm. Note that the real part of the solutions (on the left hand side) are better than the imaginary parts (on the right hand side.) (Simulation, Algorithm CG with clipping, precision(O)=6dp, precision(F)=full.)

the linear system will further amplify the instability through the condition number κ_2. The different methods have different susceptibilities to roundoff errors which are connected with the condition number relative to machine precision.

Problems with low condition number (with respect to machine precision) present no problems in reconstruction, even when the input $E-$field data (F) is not very precise. For a simple problem with condition number $\kappa_2 = 27$, excellent reconstructions were obtained with precision(F) as low as two (2) decimal places. Problems with a mid condition number ($\kappa_2 \approx 10^6$) can give accurate reconstructions with input field precision to about eight (8) decimal places. For problems with a high condition number, all standard methods have difficulties and it is necessary to incorporate constraints on the solution. For problems whose input field data is known to full precision, the conjugate gradient with clipping method has been shown capable of resolving objects less than one-half the ambient wavelength apart. The use of the conjugate gradient method with clipping and imprecise input data has not been investigated.

5.4 Two-dimensional Objects

In this section, the reconstruction of two-dimensional objects is considered. First, we discuss some more important points in using the method of projections as a discretization method for the IP(κ_D) problem. These were omitted from the previous section, because they are generally much more prominent in two-dimensional objects where the condition number of the linear system becomes large with respect to machine precision, $\kappa_2 >\approx 10^8$ on the DEC/VAX computer used (see Appendix C). Up to this point, the multiplication of two floating point numbers can be represented exactly using the approximately 16 decimal digits available on the computer.

In section 5.4.1, the question of how to choose the set of observation points is discussed. Section 5.4.2 considers the reconstruction of two-dimensional objects using the least-squares minimum norm (LSMN) method. This problem was discussed by Ney [1983].

In this section, fields F are computed normalized to E_i at (x_{cal}, y_{cal}) = (0,-35 mm) and are expressed in [dB] and [degrees].

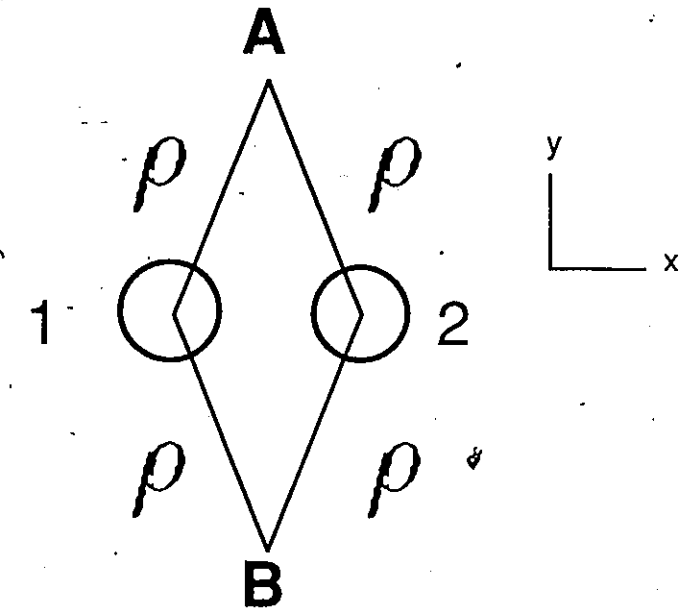


Figure 21: A Badly Conditioned Problem
Two equal size cells and two observation points arranged in such a way so that all relevant distances are equal.

5.4.1 Choice of Observation Points

Not all observation loci lead to equally stable systems of linear equations (eq. 102). The elements $K_{mn} = H_0^{(2)}(k_A \rho_{mn})$ are essentially determined by the set of observation points because the cells are generally fixed in number and location for other reasons. Therefore, the condition number κ_2 will be determined by the set of observation points, O.

Figure 21 shows a simple problem with two cells, and two observation points which are chosen so that they are at equal distances from both centers. Since $\rho_{1A} = \rho_{2A} = \rho_{1B} = \rho_{2B} = \rho$, we have two linearly dependent equations

$$K = \begin{bmatrix} H_0^{(2)}(k_A\rho) & H_0^{(2)}(k_A\rho) \\ H_0^{(2)}(k_A\rho) & H_0^{(2)}(k_A\rho) \end{bmatrix} \tag{123}$$

Therefore, the κ_2 condition number is infinite. This occurs because of the circular symmetry of the kernel $H_0^{(2)}(k_A\rho)$, i.e. the independence of this function from the angular coordinate ϕ. (It is also for this reason that a large number of cells are needed to accurately represent the angular variation of scattered field of an object when using these functions

with *no* angular variation.) Hence, the first point is that observation points must be chosen to 'break' the angular symmetry of $H_0^{(2)}(k_A\rho)$ as much as possible.

For two-dimensional problems, it is desirable to choose observation points as close to the interior of the object as possible. In this region, the kernel $H_0^{(2)}(k_A\rho)$ varies more rapidly than in its asymptotic region ($k_A\rho \to \infty$) (see Figure 4), thus reducing the condition number κ_2 of the problem. (For this reason, the linear system required to compute the total field inside the object for the direct problem leads to a system which has a low condition number.) Therefore, for objects whose support is assumed to be square, the choice of points which vary slightly about a circle is a good choice.

Observation point selection algorithm

One algorithm which proves successful in reducing κ_2 by choosing the observation points $O = (\rho_i, \theta_i)$ is the following. First, for simplicity, assume an equiangular separation of the M observation points,

$$\theta_m = (m-1)\frac{360^\circ}{M}, \quad m = 1, 2, \ldots, M \tag{124}$$

Next, search for a nominal radius, r_0, which minimizes κ_2 for $O = (r_0, \theta_m)$ by computing κ_2 for these matrices. This can be shown using PROBLEM P-2A (Figure 22) with the result in Table 6. Although $r_0 = 23$ mm has the lowest condition number in this range, the interval [22,24] mm has a large change in κ_2 and is thus not stable. Hence, choose $r_0 =$ 27 mm. Next, vary the radius r_m about the nominal value choosing values from a uniform distribution

$$r_m \in U(r_0, \delta r) \tag{125}$$

where $U(r_0, \delta r)$ is a uniform distribution in the interval $[(1-\delta r), (1+\delta r)r_0]$, and a δ is a user-chosen parameter. Table 7 shows the effect of applying eq.(125) for a number of δr. Note that for the case of the 3×3 cell object, the condition number κ_2 is reduced from 1.2×10^{17}, i.e. "infinity", to 8.5×10^2. *Thus, this heuristic does reduce* κ_2.

The problem of minimizing κ_2 can be expressed in mathematical terms but the solution of this problem is indeed difficult. Consider the rows of the matrix K_{mn} in eq. (102), where each row k, $k_k(\rho_k, \theta_k)$, corresponds to a different observation point (ρ_k, θ_k). To minimize κ_2 we desire maximizing row independence. This is achieved when all rows as orthogonal as possible from the other rows. (Finding the minimum orthogonal set is, in fact, the basic idea behind the LSMN method of section 4.2.) Thus, choose an initial point (ρ_1, θ_1) and use the Gram-Schmidt orthogonalization procedure to write the equations for the problem.

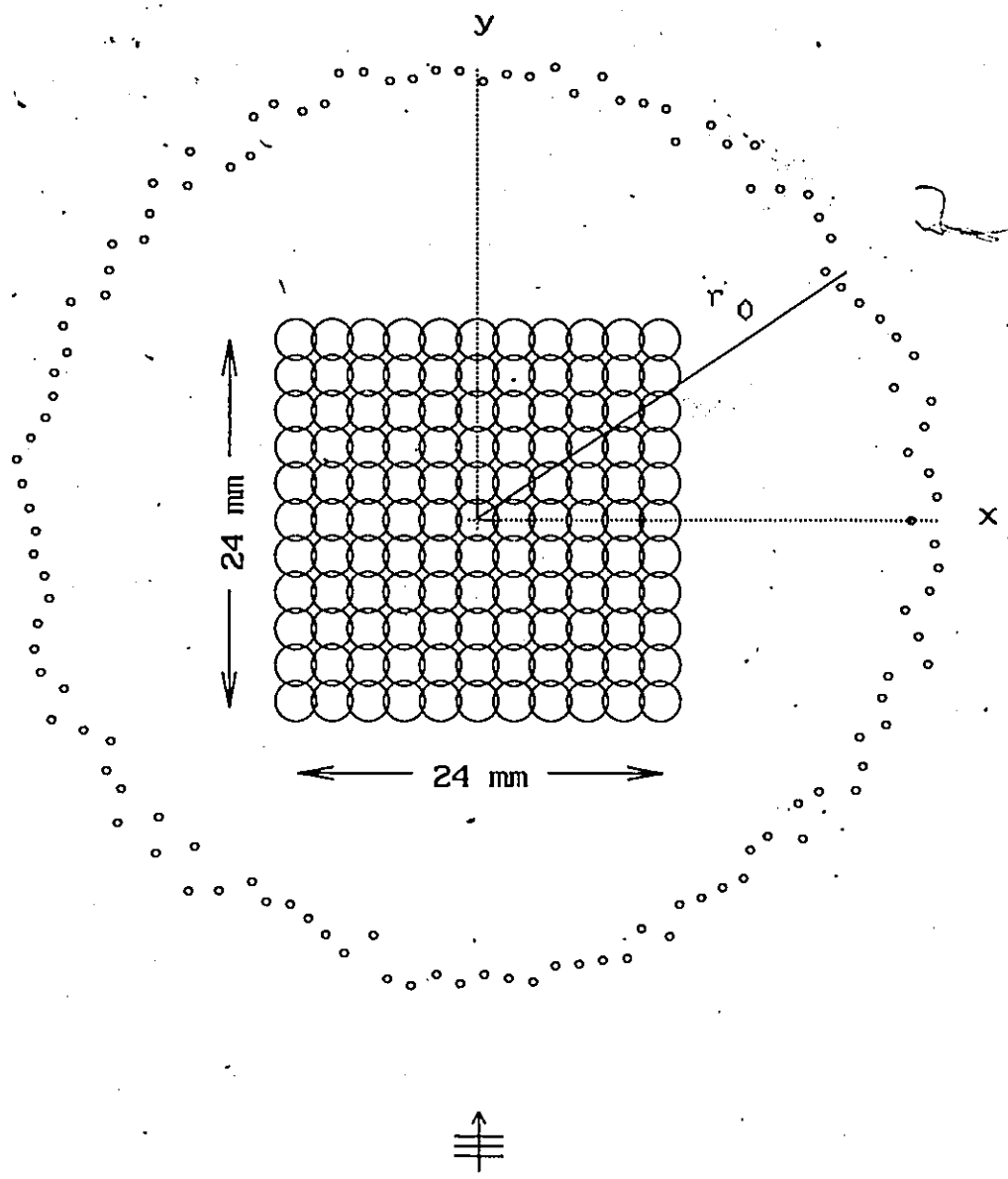


Figure 22: Geometry for PROBLEM P-2A
The object has dimensions 24 mm × 24 mm, and is discretized into a grid with $N = N_0 \times N_0$ cells. ($N_0 = 11$ is shown here with $U(27, 1.35)$ mm and fields normalized to (0,-35) mm.)

Radius of circle r_0 [mm]	Condition number κ_2
10	5.0×10^{15}
20	3.4×10^{14}
22	5.9×10^{13}
23	1.4×10^{14}
24	4.6×10^{13}
26	5.7×10^{13}
27	5.0×10^{13}
28	5.4×10^{13}
29	6.4×10^{13}
30	1.7×10^{15}

Table 6: Effect of r_0 (radius of O) on the condition number κ_2
An example using PROBLEM-P-2A with $N_0 = 11$ cells and $U(r_0, 0.25)$ mm.
(precision(O)=6dp, precision(F)=full)

Number of cells	Cell edge [mm]	κ_2 $r_m = r_0$	κ_2 $r_m = U(r_0, 0.25\text{ mm})$	κ_2, $r_m = U(r_0, 1.35\text{ mm})$	κ_2 $r_m = U(r_0, 2.00\text{ mm})$
3 × 3	8.000	1.2×10^{17}	6.3×10^{3}	1.2×10^{3}	8.5×10^{2}
5 × 5	4.800	1.7×10^{17}	1.3×10^{11}	2.6×10^{10}	1.9×10^{10}
7 × 7	3.429	7.0×10^{17}	2.1×10^{13}	4.7×10^{12}	7.5×10^{12}
9 × 9	2.667	4.6×10^{18}	2.6×10^{13}	3.4×10^{13}	2.3×10^{13}
11 × 11	2.182	1.5×10^{18}	5.0×10^{13}	1.2×10^{14}	1.2×10^{14}
13 × 13	1.846	n/c	n/c	n/c	n/c
15 × 15	1.600	n/c	n/c	n/c	n/c
17 × 17	1.412	n/c	n/c	n/c	n/c
19 × 19	1.263	n/c	n/c	n/c	n/c

Table 7: Effect of randomizing the nominal radius r_0=27 mm on κ_2
An example using PROBLEM P-2A with object dimension 24 × 24 mm, and $M = N$ observation points. Any number greater than $1/u \approx 10^{17}$ can be considered "infinite" in terms of machine precision. (Simulation, precision(O)=6dp)

We then minimize the difference between the chosen vectors k_k and those found by the orthogonalization procedure. If we assume that the $\{\theta_i\}_{i=1}^{M}$ are fixed, we can solve for $\rho_1, \rho_2, \ldots, \rho_M$ which must all be real numbers, and outside the domain of the object.

5.4.2 Using the LSMN method

One problem with the least-squares minimum norm (LSMN) method is how to choose the number of columns $k \leq N$ to use in order to obtain the optimum solution, i.e. minimum e_{RMS}. Using more columns to compute the solution, reduces the residual $\|K_H J_{eff} - E_s\|_2$, until the minimum residual (norm) is attained when using all $k = N$ columns. Thus, if the input field data F has full precision, then the solution should occur when the residual is minimum, i.e. when $k = N$. This does occur when the linear system is properly set up and the algorithm is computed with sufficient precision.

Consider Problem P-2B which is shown in Figure 23 . This is a rather complex object, with a large range of values of complex relative permittivity ranging from air $(1.0 - j0.0)$ to distilled water $(76.5 - j11.3)$. (See Table 1 for the other values.) This problem is mid-conditioned with respect to machine precision, $u\kappa_2 \approx 10^{-9}$ when $M = 2N = 50$ observation

points are used, and $r_m \in U(27, \pm 2)$ mm. Figure 24 shows both the original object and the reconstructed object when the input field data F has a precision of eight (8) decimal places. (For full precision, the reconstructed image is identical to that of the original.) Note that the reconstruction is excellent, despite the inaccurate input data. However, as the inaccuracy increases (precision(F) decreases) from eight (8) decimal places, the solution worsens rapidly.

As an example of an improper application of this method, consider $M = N'$ observation points lying on a circle of radius $r_0 = 27$ mm. Figure 25 shows the reconstruction with full input field F precision. Because of the "infinite" condition number κ_2, the reconstruction ($e_{RMS} = 0.6$) is bad.

5.4.3 Summary- Two-dimensional Objects

Problems with large condition numbers κ_2 lead to solution with large errors e_{RMS}. By choosing the observation points carefully, the condition number can be reduced, sometimes by many orders of magnitude. Apart from the increase in condition number, two-dimensional problems possess no special problems, and can be reconstructed under the same conditions as the one-dimensional problems of section 5.3.

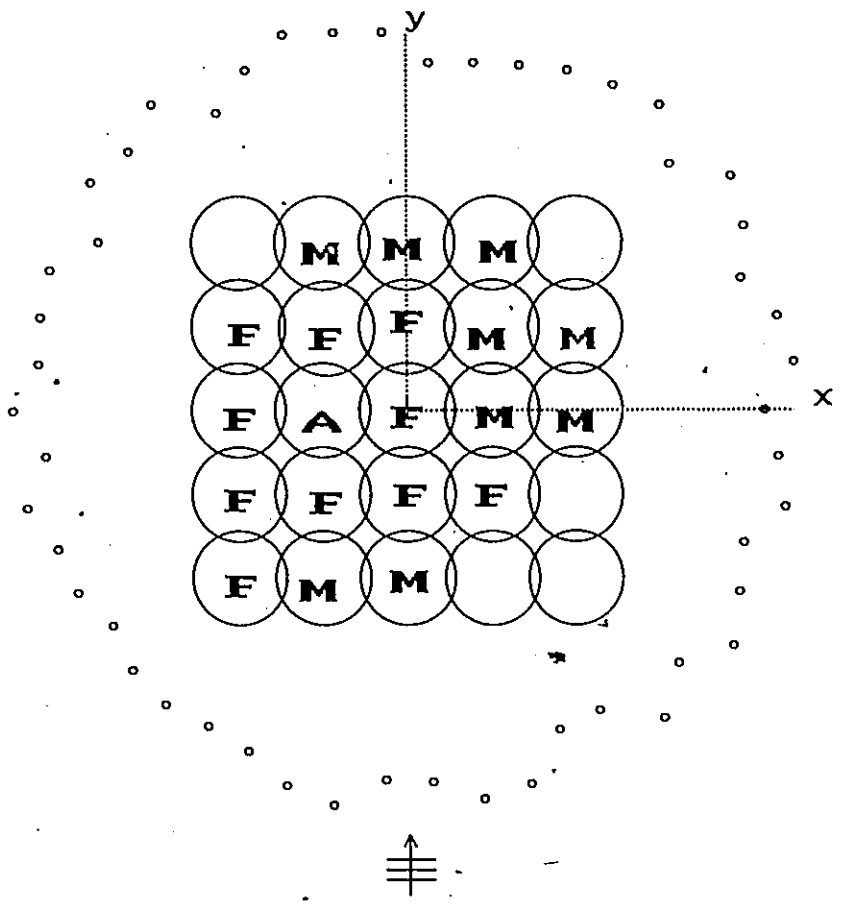
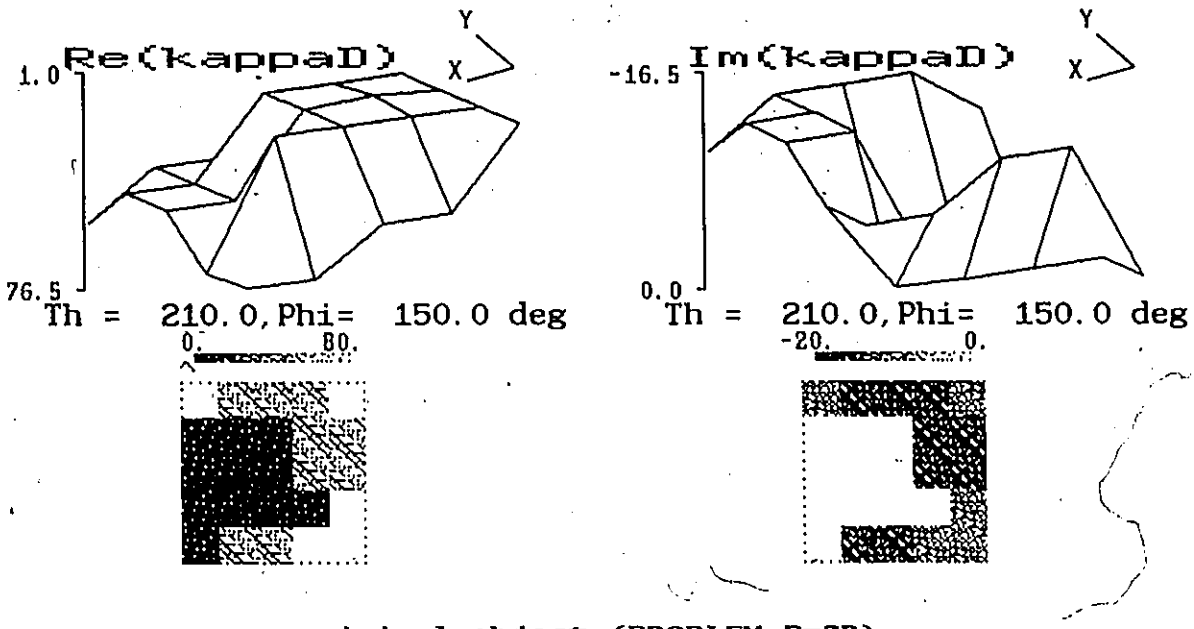


Figure 23: Geometry for PROBLEM P-2B
There are $N = 25$ cells : A for air, F for fat and M for muscle, and $M = 2N = 50$ observation points on an observation contour $r_m = U(27, \pm 2.0)$ mm. Fields are normalized to (0,-35) mm.



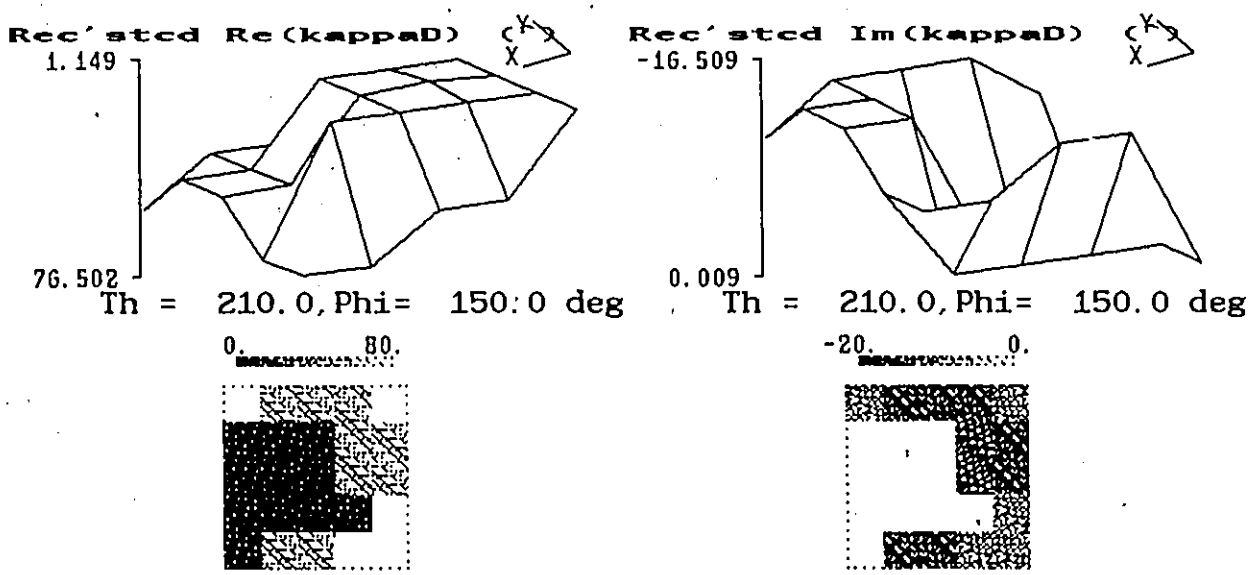


Figure 24: Reconstruction for PROBLEM P-2B with $M = 50$ observation points The error e_{RMS} is very low (0.051). For full input field precision, the reconstructed image is identical with the original. (Simulation, LSMN, precision(O)=6dp, precision(F)=8dp.)

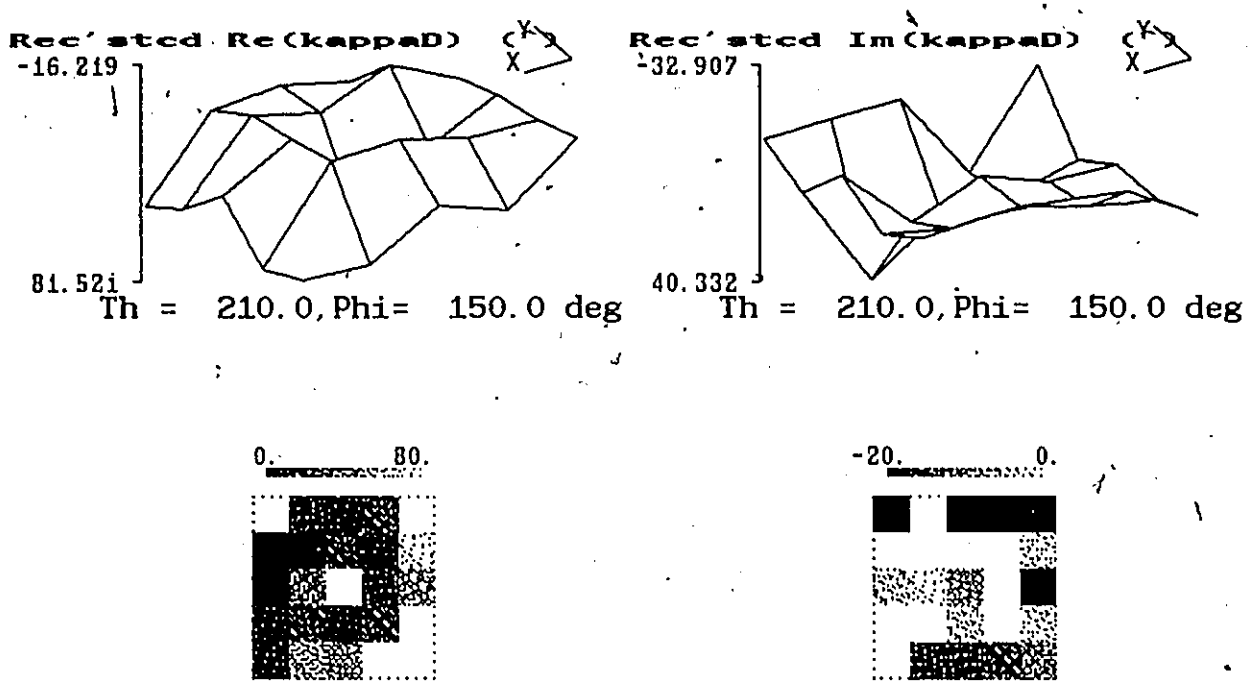


Figure 25: Reconstruction for PROBLEM P-2B with $M = 25$ observation points
By choosing the M observation points on a circle about the origin, the condition number κ_2 is very large (1.3×10^{17}), and the reconstruction is bad. An appropriate choice of point would lead to a very good reconstruction, even with imprecise input field data (precision(F) < full) as shown in Figure 24. (Simulation, LSMN, precision(O)=6dp, precision(F)=full.)

Chapter 6

CONCLUSIONS

In this thesis, the electromagnetic inverse problem IP(κ_D) of determining the complex relative permittivity of a low-loss dielectric object $\kappa_D(\underline{x})$ from its scattered electromagnetic field $E_s(\underline{x};\theta)$ at a set of points along an observation contour O was examined. From Maxwell's equations, the relationship between the effective current $J_{eff}(\underline{x};\theta) = (k_A^2 - k_D^2(\underline{x}))E(\underline{x};\theta)$ and the scattered field for a monochromatic ($e^{j\omega t}$) incident TM-field $E_i(\underline{x};\theta)$ was found to be a Fredholm integral equation of the first kind. This is a non-selfadjoint compact operator, the solution of which is an ill-posed problem. Upon discretization by the method of projections, the problem leads to a system of linear equations which are increasingly ill-conditioned as the discretization better approximates the original operator.

The original contributions of this thesis are

- object complex permittivity $\kappa_D(\underline{x})$ reconstruction in a *low loss* ambient medium taking into account the losses in the medium
- reduction of the instability of the set of linear equations resulting from discretization by a careful choice of the observation points on O by means of the condition number κ_2
- a comparison of the standard methods of solving a system of linear equations (Gaussian Elimination, Least Squares Minimum Norm, Tikhonov Regularization, and Conjugate Gradients) in solving systems of varying condition number κ_2 and input data precision
- **Algorithm CG with clipping** which permits solution of a highly ill-conditioned

system of linear equations using *physical* rather than *mathematical* constraints

- resolution of objects beyond the $\lambda_A/2$ limit imposed by the bandlimited "weakly-scattering" object approximation

This thesis extends the work of Ney [1983] by including the case of propagation in a low-loss ambient medium of complex relative permittivity κ_A, thus permitting reconstruction in a medium such as water. This method is exact in that there is no assumption about "weakly-scattering" objects. These methods generally employ the Born approximation, which implies a band-limited spectrum for both J_{eff} and E_s, as well as propagation in a lossless ambient, i.e. $Im(\kappa_A) = 0$. As in the work of Ney, the method of projections was used to compute solutions to the problem, with N pulse functions for the basis of J_{eff} and M impulses for the basis of the E_s.

A number of problems with the reconstruction process were isolated, and the knowledge was used to construct an algorithm with better, i.e. lower average mean-square error e_{RMS}, solutions. First, it was found that many problems associated with solving the discretized problem were due to the linear system having a large condition number κ_2 with respect to machine precision. (The condition number is an upper bound of the stability of the linear system.) **By carefully choosing the set of observation points, the condition number can be reduced.** In some cases, though, the condition number cannot be sufficiently reduced for the machine precision available. In this case, the only recourse is to compute on a machine with more precision.

Second, the number of observation points M must be chosen carefully with respect to the number of cells N, this depending also on the method used to solve the linear system resulting from discretization. For example, $M = 2N$ works best for least-squares minimum norm (LSMN) method and $M = N$ for conjugate gradient based methods. With "exact" input data and condition number low with respect to machine precision ($\kappa_2 \approx 10^8$ in this case), the solution generally improves, i.e. lower e_{RMS}, until a saturation point is reached. In this case, the LSMN method is adequate. However, as κ_2 increases the LSMN method is unable to incorporate constraints. **By modifying the method of conjugate gradients to "clip" the computed complex relative permittivities to their bounds known from physical constraints, it is possible to reduce the e_{RMS} error for highly ill-conditioned problems where the methods fail.** It is noted that this approach is merely a convenient way of implementing the non-linear optimization problem which represents the

desired formulation, but is both difficult to implement and very computationally intensive.

The spatial bandwidth limited approach used by other researchers limits the resolution of objects to $\lambda_A/2$, which is about 5.7 mm at the operating condition used (see (7)). By solving the equation exactly, we could resolve objects which were 1 mm in diameter and whose centers were 3 mm (about $\lambda_A/4$) apart.

However, the integral formulation used is limited in its applicability due to practical problems. First, due to the nature of the equation, it is difficult to consider objects with many cells because of the rapid increase in the condition number κ_2 of the problem. This can be combatted by using higher precision arithmetic, but it is difficult to find compilers which can support this arithmetic. Further, the computation time will also greatly increase. Second, the scattered fields which were measured using the MSCAN $E-$field probing system (Appendix B) which is based on the Hewlett-Packard (HP) 8410B network analyzer show that it is difficult to achieve measurements better than about one (1) decimal digit in either magnitude [dB] or phase [degrees]. The simulated results in Chapter 5 indicate that for inputs of such precision, object reconstruction cannot be achieved with sufficiently low error e_{RMS} (Figure 12.) Certainly, the accuracy of the measurement system can be improved. Further we have applied the method of projections with only the point-matching basis functions. Perhaps, testing functions other than impulses, e.g. the Bubnov-Galerkin method, may be able to compress information better and be able to compute more accurate solutions by creating systems of linear equations with smaller condition numbers. However, the limiting factor will still be the rapid increase in the condition number of the problem as the discretization of the operator becomes more refined.

The drawbacks imposed by the limitation on the number of cells are (1) the inability to reconstruct objects of arbitrarily fine detail, and (2) the difficulty of accurately modelling the perimeter of objects, thereby making it difficult for the object model to be used in computing an accurate estimate of its scattered field. In a practical case, this inaccurate field computation introduces more "noise" in the residual $\|K_H J_{eff} - E_s\|_2$ since $K_H J_{eff}$ is actually an estimate of the (reconstructed) scattered field.

From the work presented, it is clear that mathematical constraints alone (such as the minimum norm in the LSMN algorithm) are not sufficient to find solution to the object reconstruction problem, especially with "inexact" input data or finite precision computation. For this reason we incorporated physical constraints, i.e. bounds on the complex relative permittivity of the objects, into a method based on the elegant conjugate gradient

algorithm. This new method is more robust for our inverse problem, i.e. can solve problems with a higher condition number, than the standard methods of solving linear systems. However, the convergence properties of this method are unknown. This inverse problem is more accurately represented as a non-linear optimization problem in complex variables with non-linear constraints. Published algorithms for this problem are limited to several variables due to computational cost (time and memory.) For the inverse problem for *perfectly conducting* objects, Imbriale and Mittra also concluded that "... although there is no theoretical limitation to the size or detail of the body that can be reconstructed, there is a numerical limitation based on the accuracy of the data." We have found this to be true for material objects ($\sigma_D < \infty$) also. Further, there is still a great deal to be learned about use of the Method of Projections in solving inverse problems, especially those with inaccurate input data.

Appendix A

SCATTERING BY A HOMOGENEOUS CIRCULAR DIELECTRIC CYLINDER

The problem of $E-$ field scattering of a plane wave $E_i(0,0)e^{-jkx}$ by a homogeneous isotropic circular cylinder (κ_D, μ_0) of radius a in two dimensions was considered by Harrington [1961, p.233+,261] (Figure 26). Because an exact analytical expression for such a cylinder in a low-loss ambient medium (κ_A) exists, it provides a valuable reference, which we shall refer to as *the series solution for the circular material cylinder*

$$E_s(\rho,\phi) = E_i(0,0) \sum_{n=-\infty}^{n=\infty} j^{-n} a_n H_n^{(2)}(k_A\rho) e^{jn\phi} \tag{126}$$

where

$$a_n = \frac{-J_n(k_A a)}{H_n^{(2)}(k_A a)} \left[\frac{-\kappa_D J_n'(k_D a)/(\kappa_A k_D a J_n(k_D a)) - J_n'(k_A a)/(k_A a J_n(k_A a))}{\kappa_D J_n'(k_D a)/(\kappa_A k_D a J_n(k_D a)) - H_n^{(2)'}(k_A a)/(k_A a H_n^{(2)}(k_A a))} \right]$$

Consider the following setup which can be used to evaluate an experimental system used to measure scattered $E-$fields. PROBLEM CAL-1 consists of an acrylic rod of complex relative permittivity $\kappa_D = 2.6 - j0.0$ and radius 6.4 mm (nominal radius 1/4 inch) which is immersed in a distilled water ambient ($\kappa_A = 76.5 - j11.3$) at the operating condition f = 3.156 GHz and T = 20 C. Figure 27 compares the "exact" solution (given by eq. (126)) and the solution obtained by using the Method of Projections (MoP) with the object discretized

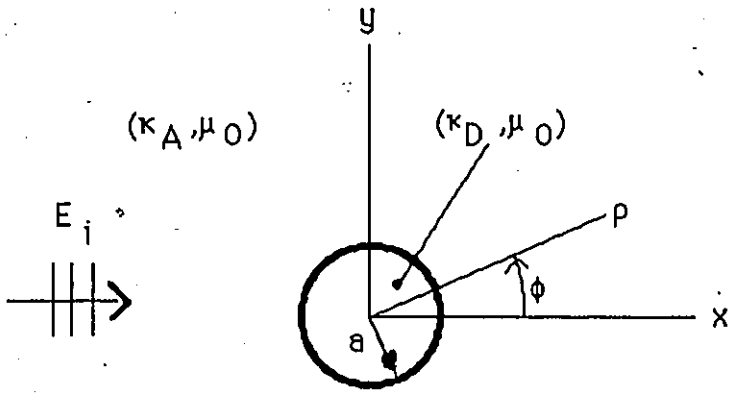


Figure 26: Scattering of a Plane Wave by a Homogeneous Circular Cylinder From Harrington [1961, p. 233].

into an $N_0 \times N_0$ grid of cells using the point-matching method. The Fortran 77 code is subroutine SCAT from Ney [1982, Appendix B.2], modified to allow field computation in a low loss ambient medium with complex relative permittivity κ_A. Note that for $N_0 = 17$ cells (cell edge size, $d = 0.753$ mm) the solution by the point-matching method is close to that given by eq. (126). Using, $N_0 = 13$ cells ($d = 0.923$ mm), however, introduces a larger error in the region around $\phi = 180°$, i.e. the back-scatter region.

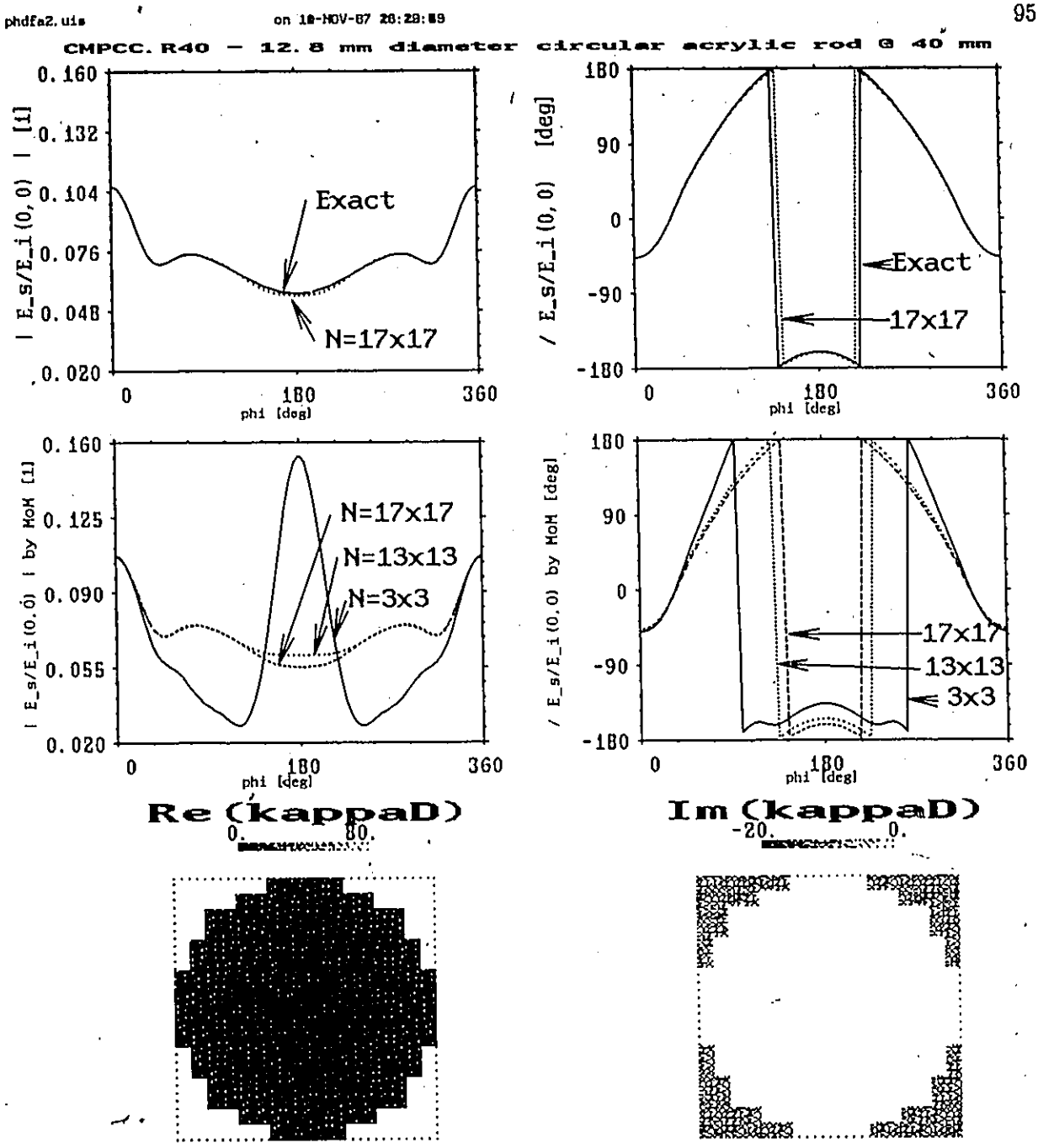


Figure 27: Scattered Field from a 12.8 mm diameter acrylic rod (PROBLEM CAL-1) The top two graphs show the series expansion ("exact") solution compared to MoP solution with $N = 17 \times 17$ cells. On the left side, are magnitudes normalized to $E_i(0,0)$ [1], and on the right are the phases [degrees]. The middle two graphs show the MoP solutions for $N = 3 \times 3, 13 \times 13$, and 17×17 cells. The bottom shows the object block model for the $N = 17 \times 17$ case.

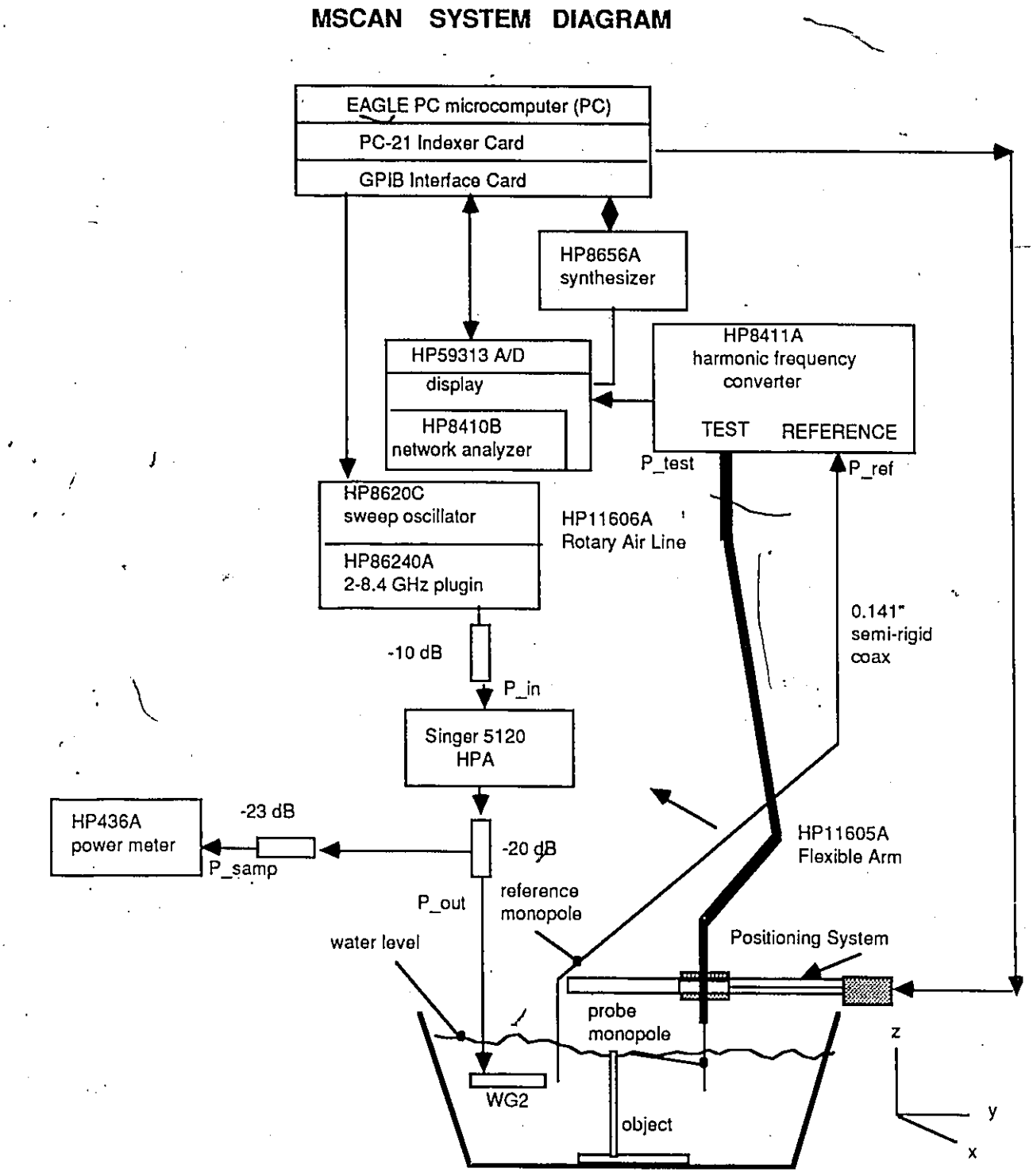
Appendix B

MSCAN - A MICROWAVE E–FIELD PROBING SYSTEM

MSCAN is an E–field probing system developed at the University of Ottawa to probe a vertically (z) polarized E–field in the x-y plane in a water (or other non-corrosive liquid) ambient medium. Its purpose is to measure the z-polarized E–field scattered by objects of size about 25 mm × 25 mm. The data collection hardware is shown in Figure 28. The system is controlled by an IBM-PC compatible (EAGLE-PC) computer (8088-8 MHz processor) with 640kB RAM and two 360k 5 1/4" floppy disk drives. Two VELEMEX B2521Q1J aluminum slides (with 1 mm pitch leadscrews) are driven by COMPUMOTOR M57-83S stepper motors (25000 steps/revolution) which in turn position a *probe monopole* in the x-y plane. The PC controls the stepper motors through a PC21 indexer card, and the Hewlett-Packard (HP) 8410B automated network analyzer (ANA) system through an IEEE/GPIB interface card.

An HP-8620C sweep oscillator with HP86240A 2.0-8.4 GHz unit provides the input power to an open ended waveguide, WG2 (see Figure 29), described in Assenheim et al. [1980]) through a SINGER 5120 high-power (10W) amplifier (HPA). The HPA was used to overcome the problems with the attenuation of water (measured at about 5 dB/cm). This system is phase-locked by a HP-8709A synchronizer and HP-8656A 0.1-990 MHz frequency synthesizer.

A *reference monopole* is clamped in a position which provides a convenient REFERENCE (0 dB) input to the HP8411A harmonic frequency converter. The probe monopole is connected to the TEST input of the HP8411A through an HP-11605A Flexible Arm and



University of Ottawa, Dept. of Electrical Engineering,Ottawa, CANADA.

Figure 28: MSCAN System Diagram

HP11606A Rotary Air Line. Both monopoles as well as the cabling are made from 0.141 inch semi-rigid coaxial cable.

The PC runs the (Microsoft) MS-DOS version 3.20 operating system. Software is written in MS/FORTRAN-77 version 3.31 except for some low-level I/O routines which are written in MS/Assembler MASM version 1.10.

Program MSCAN controls both the positioning system which moves the probe monopole in the x-y plane, and the network analyzer which collects the electrical measurements. The network analyzer subroutines of C. Sibbald [1985] were greatly modified to permit input from disk files and enhanced to permit measurements with signal averaging. This was necessary to improve the precision of the measured values by reducing the random errors. (Recall that a measurement process is *precise* if the dispersion of values is small, and is *accurate* if the values cluster closely together around the correct value, c (Cameron [1960, p. 137].) The measurement errors of the network analyzer can be separated into two classes - random errors and systematic errors. Random errors are due to variations due to the temperature drift of the components, noise and other physical changes in the measurement setup between the calibration and the measurements. This error is not repeatable. Systematic errors occur because of non-ideal components, e.g. leakage in system components. These errors are repeatable, and can be well compensated by a good system model.

B.1 Measurement Repeatability

A number of tests were performed to find the best operating environment for the system. Figure 30 shows the sample average (μ =mu) and sample standard deviation s for the transmission coefficient S_{21} in the HP8410B ANA with the RETURN port of the HP8743A Reflection/Transmission test set connected to its UNKNOWN port through the HP11605A flexible arm. It is seen that at least a two hour warmup period is required to allow the generators and oscillators of the network analyzer to stabilize their frequencies (HP8709A Synchronizer Manual, June 1969, p. 5-1). Figure 30 also shows μ and s for S_{21} measured after the HPA. Because of the instability of the HPA, the reference monopole was added as explained above.

Figure 31 shows the standard deviation s for measurement of S_{21} on our HP8410B in comparison with the measurement uncertainty in the HP specifications for the HP8410 and HP8510 as well as a 6-port ANA (Hoer [1979]). s was measured with one connection of

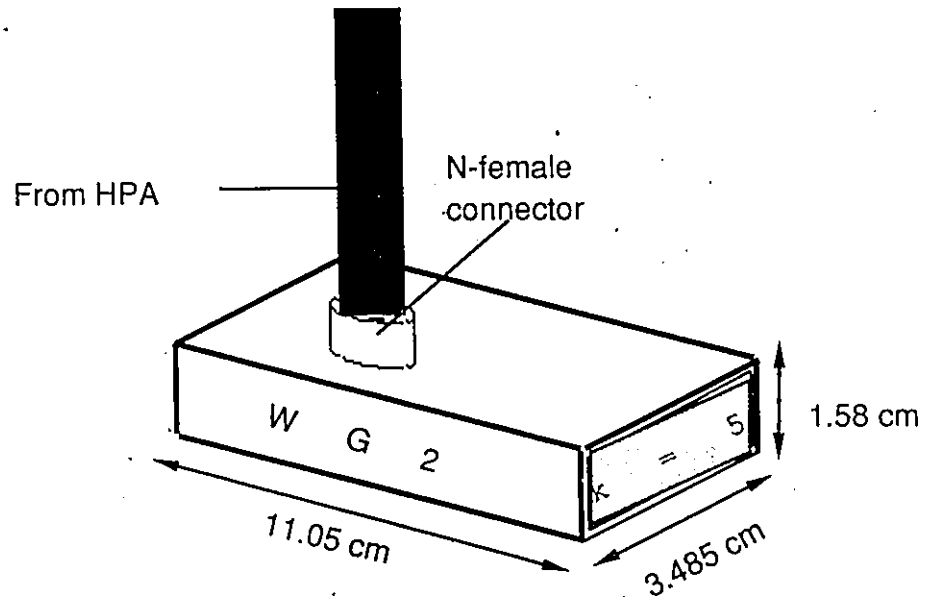


Figure 29: Open-ended Waveguide Radiator (WG2)
The waveguide is filled with a dielectric $\kappa = 5$. When the waveguide is immersed in water, the nearest minimum reflection coefficient to 3 GHz is at $f = 3.156$ GHz, where $|S_{11}| \approx -16$dB. Dimensions given are for the inside of the waveguide.

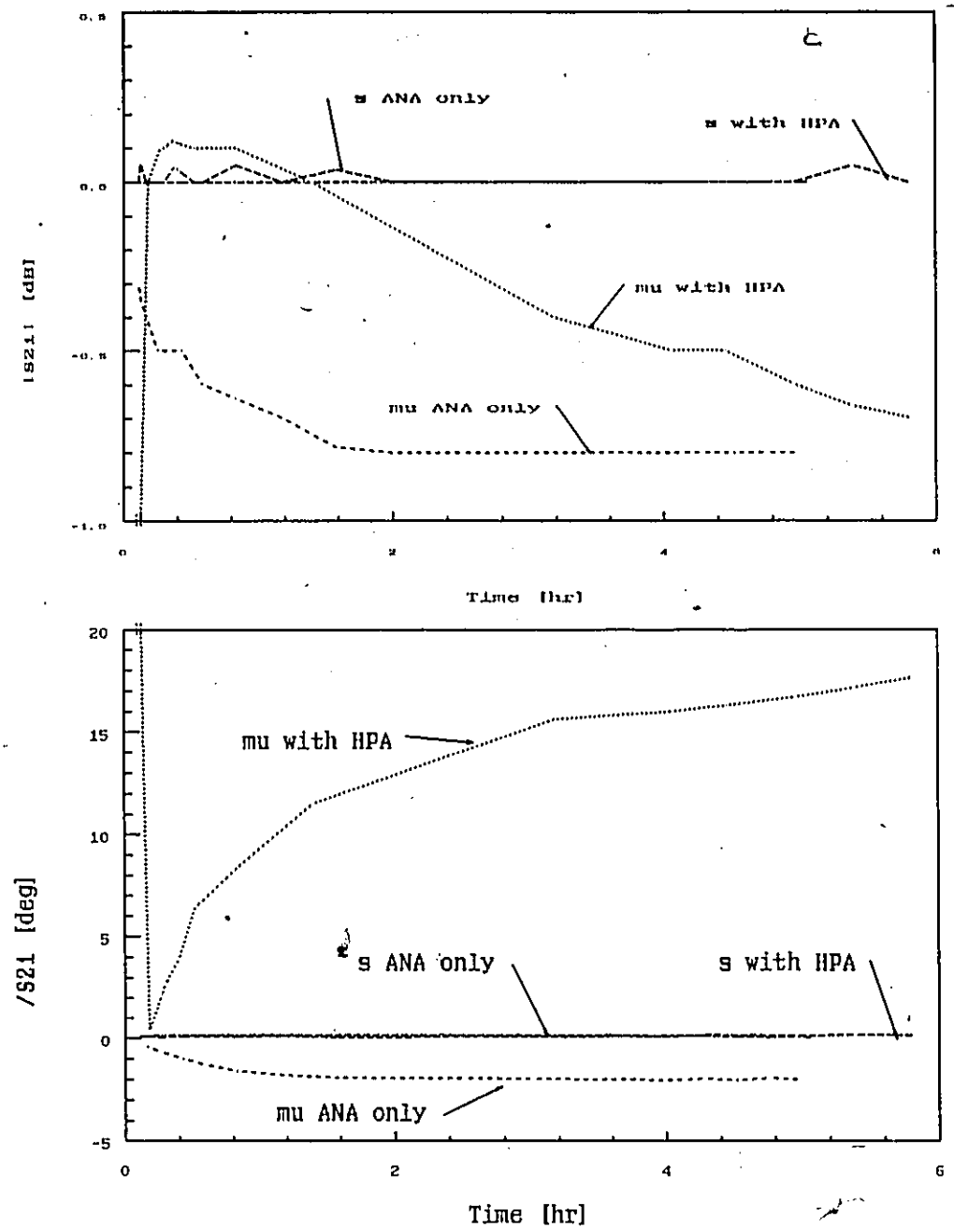


Figure 30: S_{21} measurement repeatability

Magnitude and phase repeatability for the network analyzer with the TEST port connected to UNKNOWN port (ANA) and through the high power amplifier (HPA). mu denotes the mean, and s the standard deviation.

S_{21} [dB]	Mag s [dB]	Phase s [deg.]
0	0.04	0.10
-10	0.10	0.10
-20	0.20	0.13
-30	0.20	0.31
-40	0.20	0.64
-50	0.46	1.46
-60	1.20	5.17
-70	3.77	26.43

Table 8: Measured Maximum Standard Deviations in the UofO HP8410B ANA

the cables, hence excludes the effects of connector repeatability (see Table 8 for numerical values.) For S_{21} magnitude measurements less than -40dB, the standard deviation s is about 0.20dB. For S_{21} phase measurements, the standard deviation is about 0.64 degrees. This represents the best practical precision available for the object reconstruction system desired using the available equipment, i.e. $\alpha \approx 5$ dB/cm $\times$ 8 cm is about -40 dB.

Figure 29 shows the waveguide radiator (WG2) used to produce the incident fields. WG2 was immersed in a 1L beaker of untreated (tap) water and the frequency nearest 3 GHz to given a minimum reflection coefficient $|S_{11}|$ was found: $|S_{11}| \approx -16$ dB at $f = 3.156$ GHz. This corresponds to the maximum power output from the radiator. At this frequency, the far-field $r_{FF} = 2D^2/\lambda_A \approx 340$ mm where D is the largest dimension of the radiator cross-section. Due to the geometrical constraints of the water tank and the attenuation of water, the certer of the object was located nearer the aperture of WG2. Figures 32 and 33 show the $E-$field pattern (magnitude and phase) in the x-y plane measured by MSCAN with the waveguide about 135 mm from the tank center $(0,0)$. At each point in the measurement

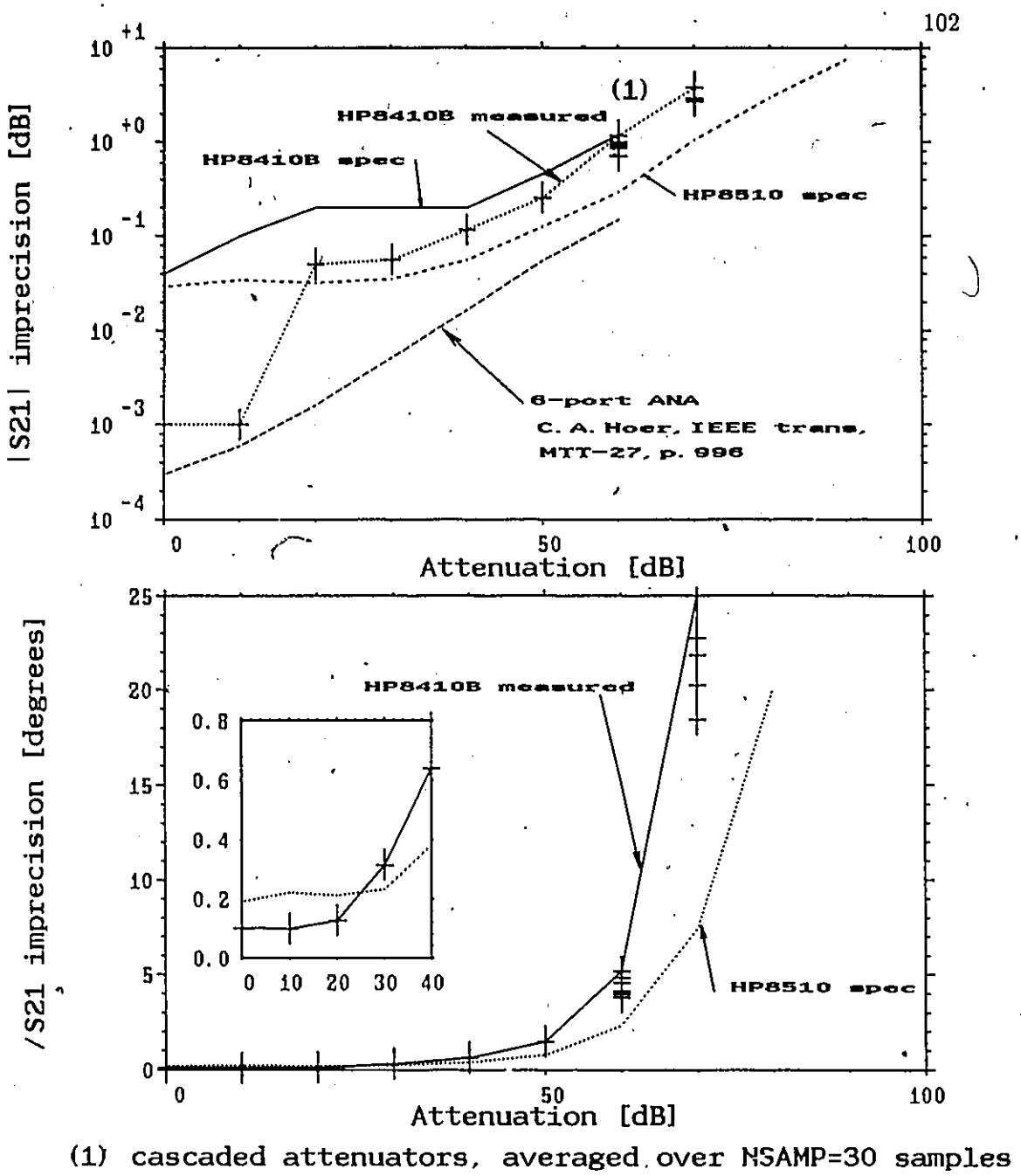


Figure 31: Comparison of Measurement Uncertainty in various ANA's
S_{21} Magnitude and Phase standard deviation (measured shown by a +) for the UofO HP8410B, and the uncertainties/accuracy for other network analyzers. Note that the measured standard deviation s is within the accuracies specified for the HP8410B. (Hoer did not give phase imprecisions.)

plane, NSAMP=10 samples of S_{21} were taken and averaged (in magnitude [dB] and phase [degrees]). The constant phase contours show that the wave front is approximately plane at the center of the tank at the z=z_0 cut shown.

B.2 Measurement of Scattered $E-$fields

To further test and establish the potential accuracy of the system, a 1/2 " (measured diameter is 12.8mm) acrylic rod of circular cross-section was used as a test object because its scattered field can be computed exactly (see Appendix A). Further, its high contrast dielectric constant $\kappa_D = 2.6 - j0.0$ (with respect to the ambient water $\kappa_A = 76.5 - j11.3$) provides a strong scattered field. This can be seen by Figure 34 which shows the computed scattered field normalized to the incident field $E_i(0,0)$ for three 12.8 mm diameter circular cylinders of various materials (acrylic, muscle, and sea water) representing the range of dielectric constants of interest. Also shown are the computed differences between the total $E-$field (with the object) EO and the incident $E-$field (without the object) Ei. Note that as the dielectric constant becomes closer to that of the ambient (water), i.e. the *contrast* decreases, the fields must be measured to greater precision in both magnitude [dB] and phase [degrees].

The geometry for the measurement of the acrylic rod's scattered field is as shown in Figure 26, except that the incident field E_i is propagating in the +y-direction. Figure 35 shows the exact (series expansion) solution for the scattered field $E_s\hat{z}$ normalized to the incident field at the center of the measurement plane, $E_i(0,0)$, compared to the field computed by $|E - E_i|$ with both quantities truncated to two (2) decimal places in [dB] and [degrees]. Figure 36 shows the same with truncation to one (1) decimal place. These figures show that it is not easy to measure E_s accurately even when the errors in both the incident field E_i (the field without the object) and the total field E (the field with the object) seem to be "small". As we have seen in Table 8, the standard deviation s for the ANA alone at $|S_{21}| = 40$ dB is about 0.64 degrees.

Figure 37 shows the scattered field measured by MSCAN, compared to the exact solution in a distilled water ($\kappa_A = 76.5 - j11.3$) ambient. In the region around $\phi = 180°$, the scattered field is weakest and $|E_i|$ and $|E|$ are approximately equal. The large error in E_s is due to differences in phase between E_i and E. (This problem due to phase errors has been verified by simulating both magnitude and phase errors. Indeed, the signals from

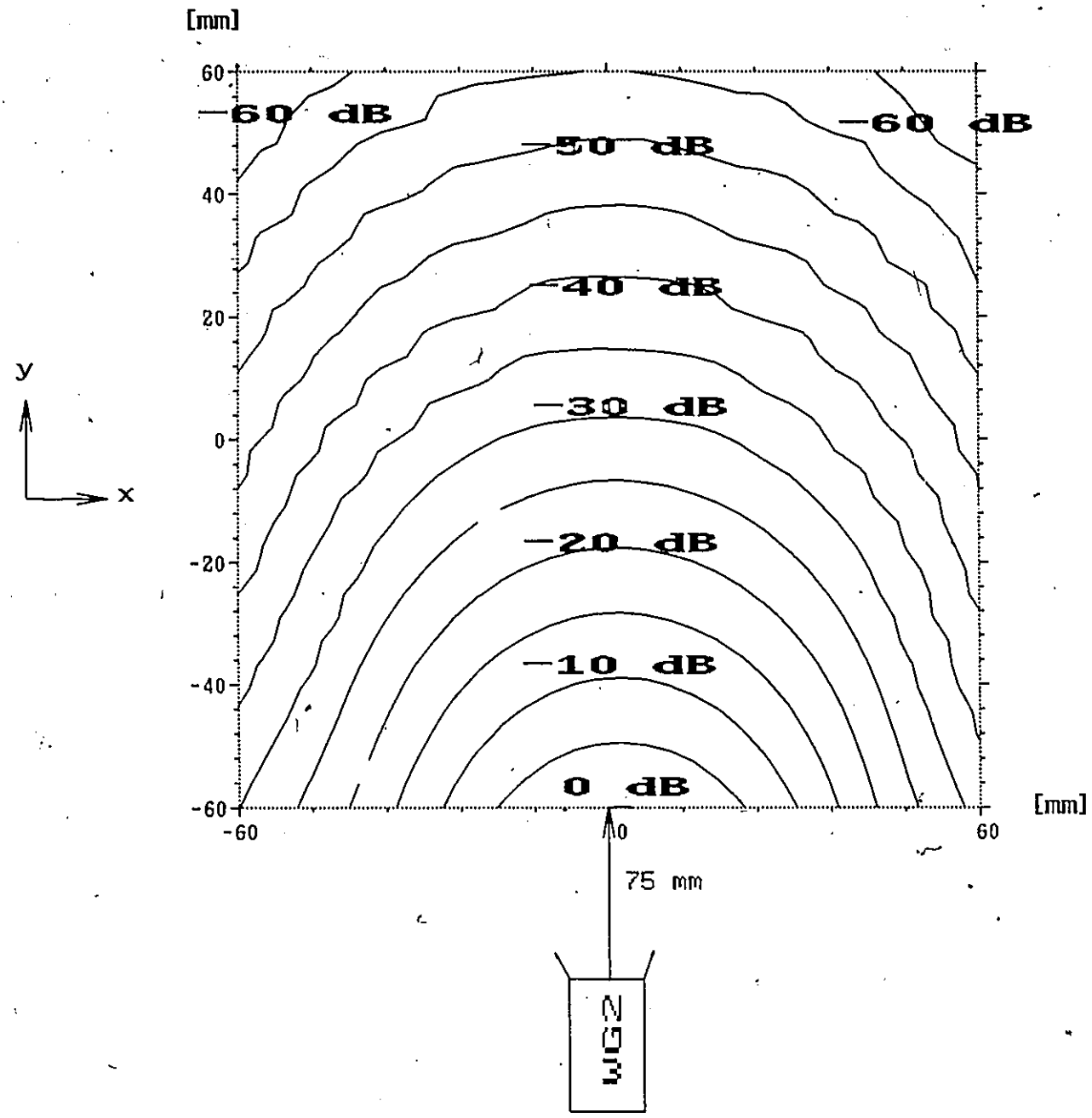


Figure 32: WG2 constant magnitude contours
Axes are labelled in [mm]. The waveguide aperture is located at about y = -135 mm. The field is sampled on a 32 × 32 grid.

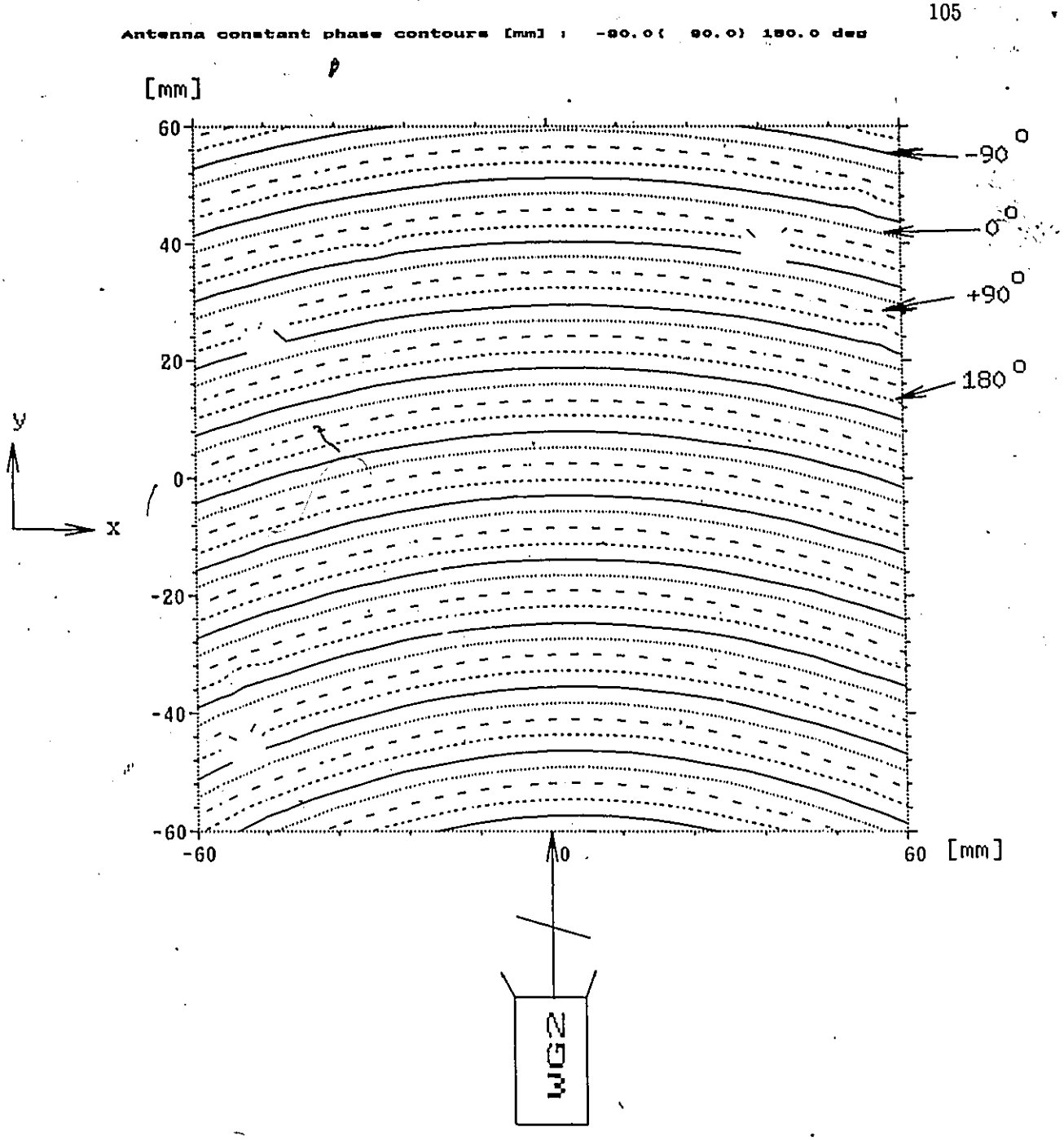


Figure 33: WG2 constant phase contours
See Figure 32 for details.

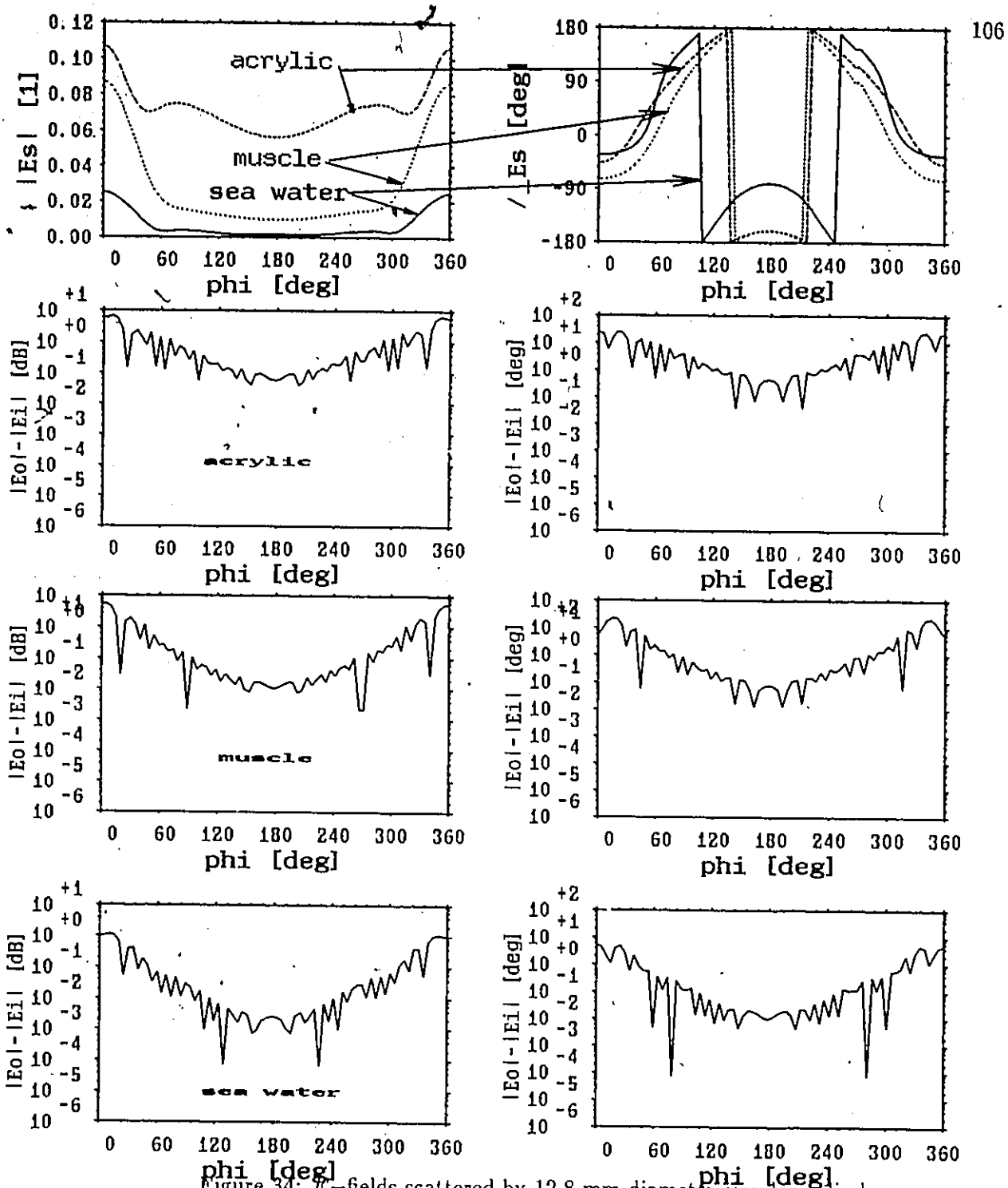


Figure 34: E–fields scattered by 12.8 mm diameter circular cylinders
The bottom six figures show the *difference* between the total (Eo) and incident fields (Ei) in both magnitude (on the left in [dB]) and in phase (on the right in [degrees]) plotted on a log scale.

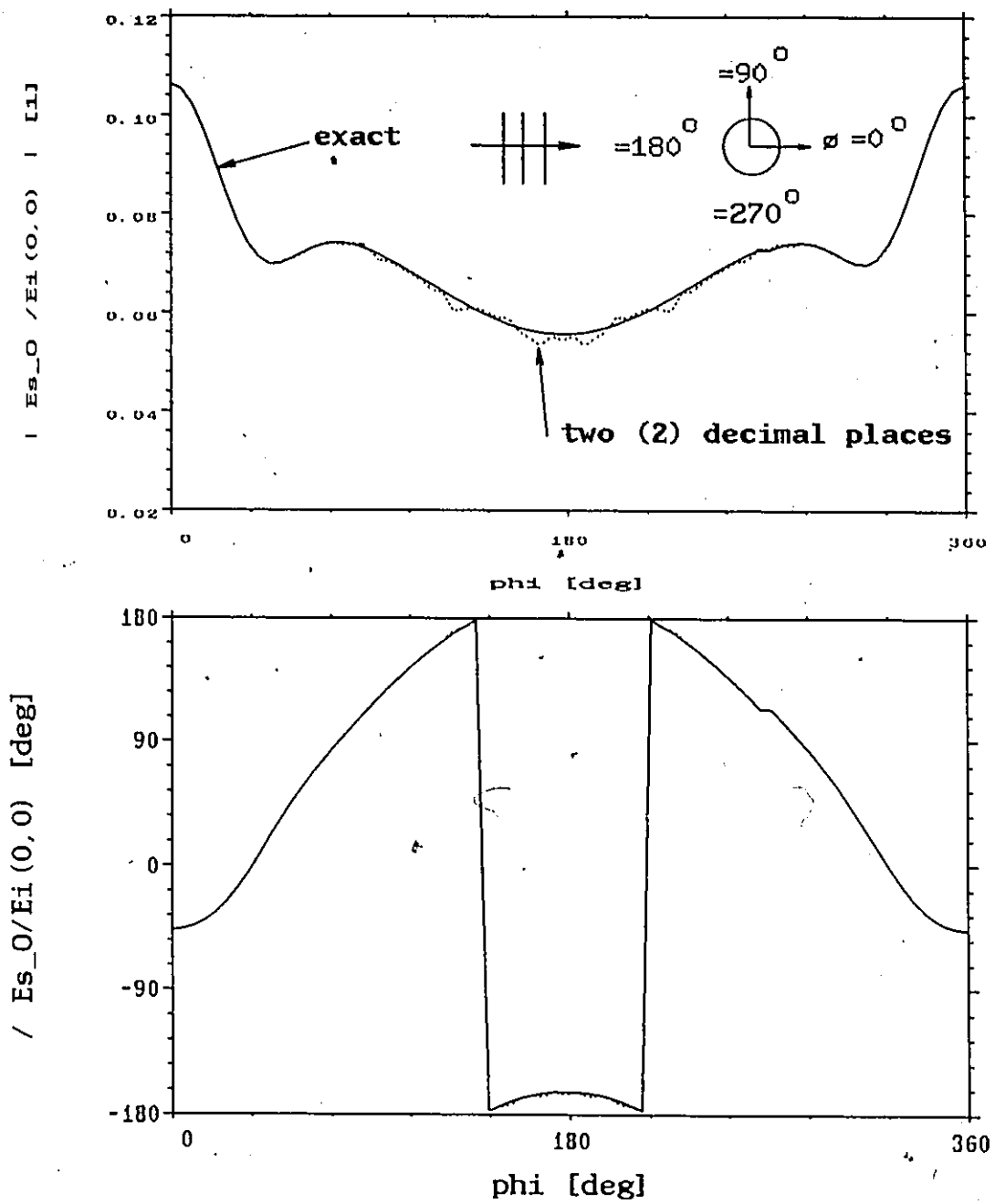


Figure 35: Exact Scattered Field for the 1/2 inch acrylic cylinder
The field is computed by the series expansion technique (see Appendix A), and is compared with the scattered field computed by the difference $|E - E_i|$ with both quantities truncated to two decimal places (dB, degrees).

HOMOCYL.R40 - "HOMOCYL" test : 1/2" circular cylinder of acrylic rod @ 40mm

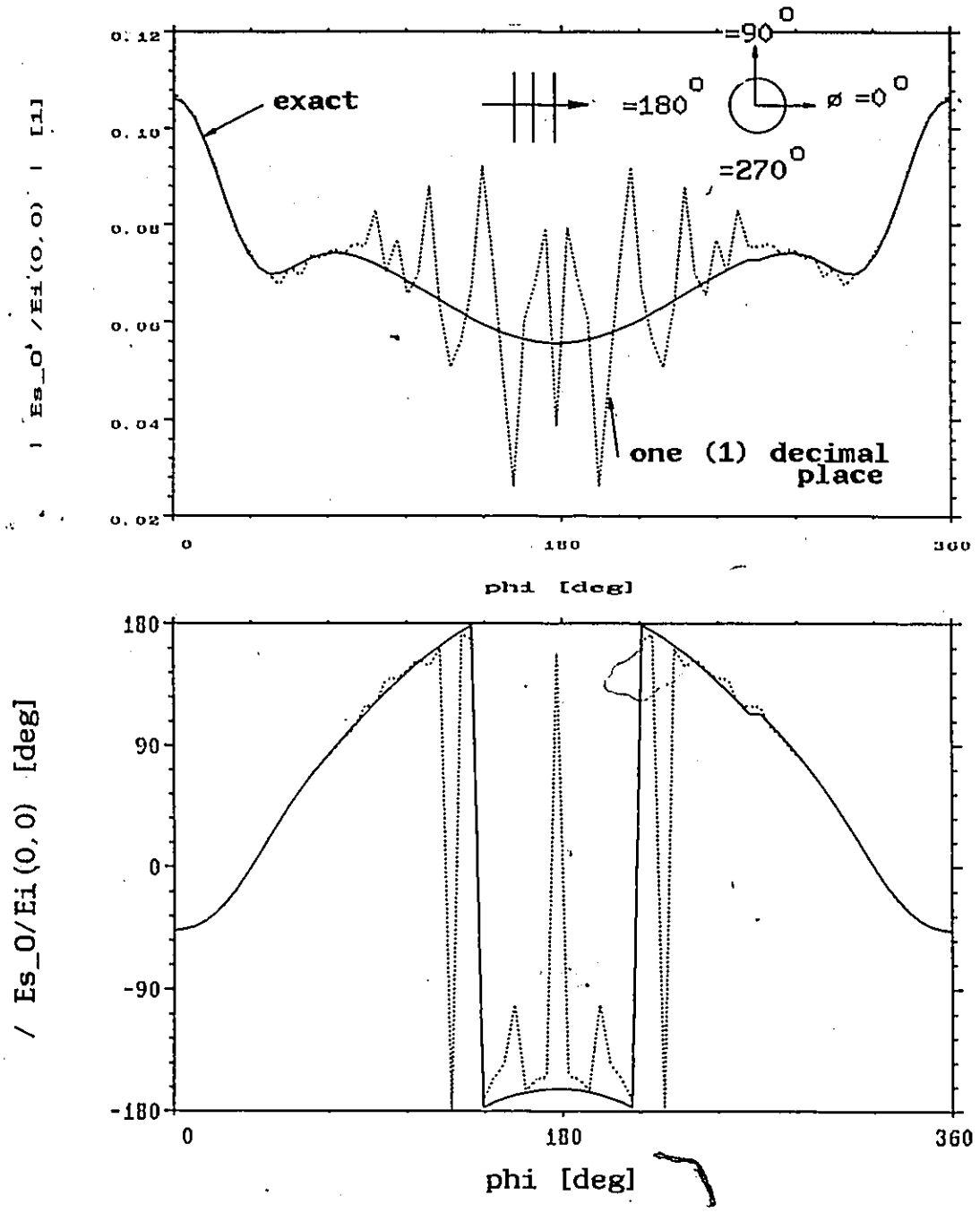


Figure 36: Exact Scattered Field for the 1/2 inch acrylic cylinder
See the previous figure. E, E_i are truncated to one decimal place (dB, degrees).

the HP8411A Harmonic Frequency Converter are downconverted to 20.278 MHz to avoid phase errors due to (1) the larger phase noise of the amplifiers in the GHz region, and (2) cable flexing at high frequencies.) We recall that $\lambda_A = 11.43$ mm which corresponds to $\beta_A = 31.5°/\text{mm}$. Hence, if we measured E_i at (x, y) and measured E at $(x, y) \pm 0.1$ mm, our scattered field could be very inaccurate because of the phase difference alone. Such large phase differences can easily occur by rotating the joints in the HP11065A Flexible Arm in the MSCAN system. Figure 38 shows the same measured values compared to the series solution for an ambient of $\kappa_A = 78.0 - j14.4$. In the region where the scattered field is the strongest, the measured and computed values of E_s are in fair agreement ($\phi \approx \pm 30°$). Since it is difficult to measure κ_A precisely enough to get good agreement with the scattered field measurement, a trial and error scheme should be used to find a suitable value for the ambient complex relative permittivity κ_A.

B.3 Network Analyzer Limitations

The theoretical limitations of the network analyzer can be evaluated by considering the quantization error of the A/D converters used to digitize the data. Assuming truncation to the nearest binary value, the smallest increment in magnitude $m_T = 20$ dB $/2^N$ where N is number of bits of magnitude in the A/D converter. In our HP8410B, N=10, hence m_T = 0.098 dB. Of course, this degree of accuracy is not usually attained because of noise in the source voltage of the A/D converter. The major problem encountered thus far, however, is the measurement of phase - a well-known problem with the HP8410B. Apart from requiring the division of two measured signals, phase [degrees] $= 180/\pi \tan^{-1}(r_2/r_1)$ where r_i is the reading [volts] on A/D channel i, the flexible arm does not give equi-phase shifts at all combinations of the bending of its joints. Phase problems are certainly the more serious.

Further investigation of the accuracy of the MSCAN system is warranted if it is to be used in a object reconstruction system.

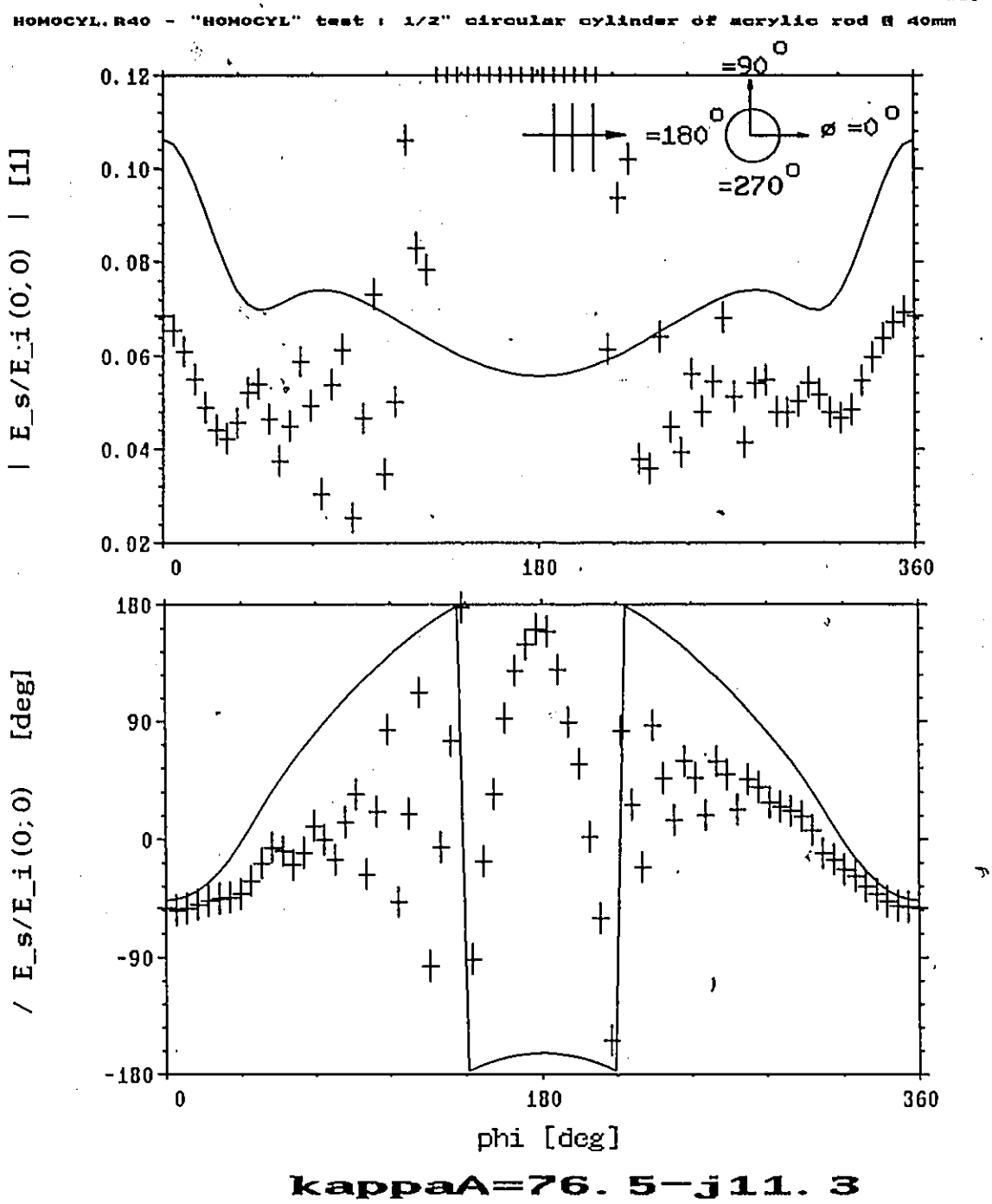


Figure 37: $E-$field scattered by a 12.8 mm (1/2") acrylic cylinder in water measured by MSCAN

The measured values (+) are compared with the series solution for $\kappa_A = 76.5 - j11.3$ (distilled water) around the cylinder $\phi = 0°$ to 360°. The magnitude data near $\phi = 180°$ are clipped on the upper part of the graph to show more detail in the regions where the measurements are good. (Measured, T=20C, f=3.156GHz, Pol=VV.)

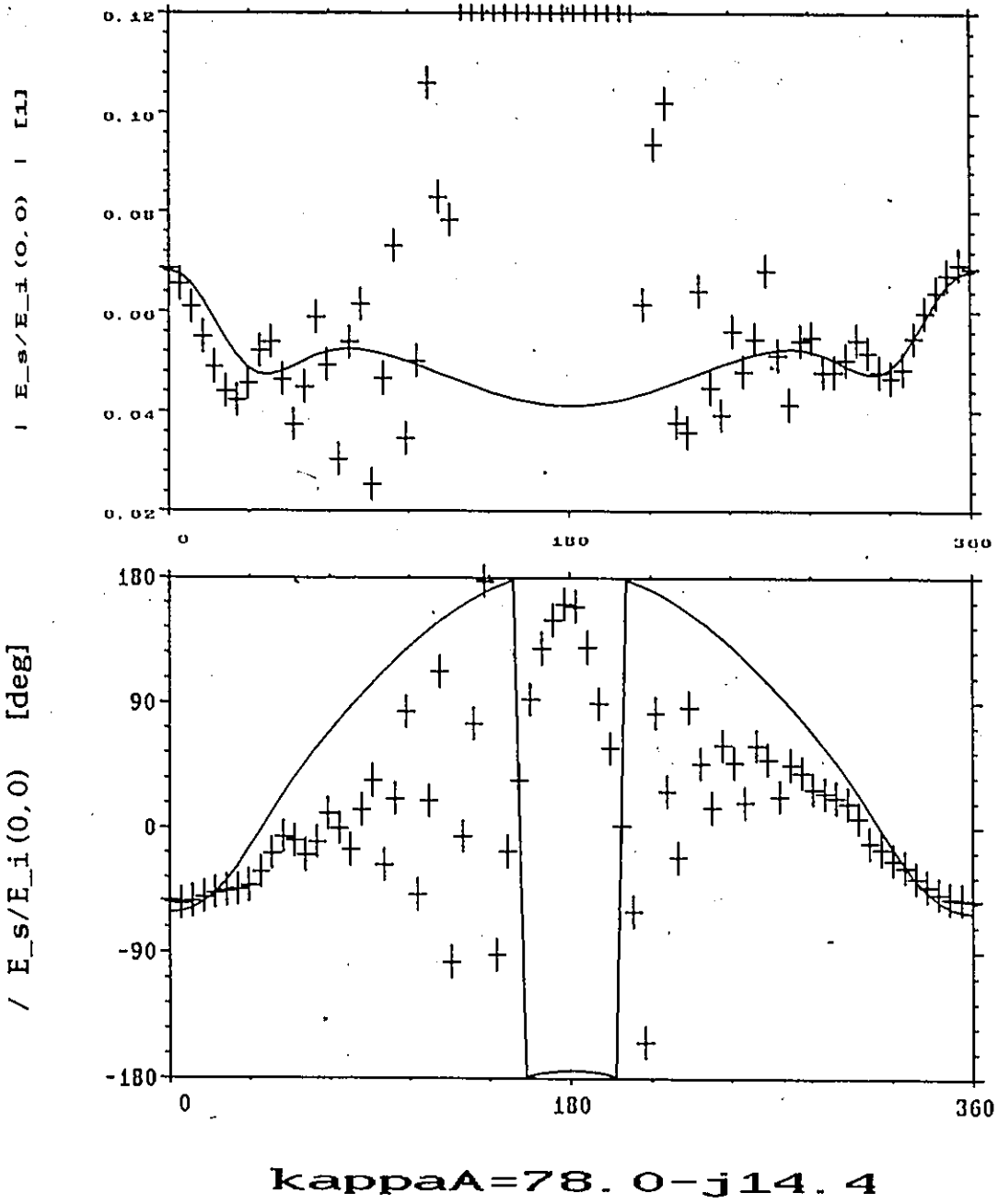


Figure 38: $E-$field scattered by a 12.8 mm (1/2") acrylic cylinder in water measured by MSCAN
The series solution for $\kappa_A = 78.8 - j14.4$ is superimposed. See the previous figure for details. (Measured, T=20C, f=3.156GHz, Pol=VV.)

Appendix C

OBJECT RECONSTRUCTION SOFTWARE AND COMPUTING MACHINE

Computations required for this thesis were performed on Digital Equipment Corporation (DEC) VAX-11/750 and a VAXstation II/RC computers running the VAX/VMS 4.5 operating system. All programs were written in VAX/FORTRAN-77 and data was computed in DOUBLE COMPLEX (D_floating complex number format) which can be expressed by the 4-tuple (see Chapter 3) $(\beta, t, L, U) = (2, 55, -128, 127)$. It convenient to consider the unit roundoff, u, which in base 10 is approximately 2.8×10^{-17}. (u can be approximated by finding the smallest number for which (1.+EPSM.EQ.1.).) This gives about sixteen (16) decimal digits of computing precision. (A computer has a *precision* of T digits is there are T digits in the mantissa of a floating point number.)

The software required for reconstruction is based on a number of algorithms and FORTRAN code obtained from other sources. In most cases, the code has been extensively modified by the author.

Code for the computation of Bessel functions $J_n(z), Y_n(z)$ and $H_n^{(}(z)$ with complex argument $z = Ae^{j\phi}$ are taken from Ardill and Moriarity [1979]. These subroutines were checked with tabulated values (Abramowitz and Stegun [1972], National Bureau of Standards [1947] and [1950]) for $\phi = 0°$ and $\phi = 5°$. This is the range required for propagation in the low loss ambient medium considered. These agree to ten decimal places in all cases,

except for a few instances where an error occurs in the tenth decimal places in the tables. The FORTRAN code was obtained from the National Research Council of Canada and was tuned by Mario Chénier.

LINPACK (see Dongarra et al. [1979]) subroutines ZGEFA and ZGESL were used to estimate the condition number RCOND $= 1/\kappa_2$ and solve the linear systems of equations. ZSVDC was used to compute the singular values of matrices, and hence the condition number of the matrices.

Pseudo-inverse computations were performed using subroutine PSEUDO (Ney [1983, pp. 135-139]) which is a Modified Gram-Schmidt (MGS) method. This finds the least-squares minimum norm (LSMN) solution of a problem.

Two-dimensional (2D) functions $(x, y = f(x))$ were plotted by routines developed by the author. y values which are outside the bounds of the graph are clipped to the bounds of the graph and marked with a '+'.

The plotting of three-dimensional (3-D) functions $z = f(x, y)$ was done using modified versions of the projection code in ACM Algorithm 483 (Watkins [1974]) and the contour plots from a heavily modified version of (Synder [1978]). Gray scale images were simulated by dithering. The value of the function at each pixel (x, y) was expanded to an $n \times n$ block with the number of "white" pixels corresponding to the "whiteness" of the value (0 = black, maximum = white).

The scattered fields outside the object were simulated by program GEOMEPS which computes these fields using Richmond's method [1965], and is capable of rounding data to various numbers of decimal places in order to simulate loss of accuracy. (The key subroutine SCAT is described in Appendix A.) Alternatively, these fields can be measured with program MSCAN as described in Appendix B. Either of these are compatible with program ISI1, which reconstructs the object using a number of possible procedures as described in Chapter 4 and 5. ISI1 is based on the program PROGEPS by M. Ney [1983]. This program is quite lengthy and is not reproduced here. Finally, program PLOTEPS produces graphical and image output of the reconstruction data on a DEC VAX Station II/RC workstation.

Appendix D

NUMERICAL COMPUTER CODE AND CODE TESTING

The following pages contain VAX/FORTRAN-77 source code for some of the methods of solving the linear systems of equations $Ax = y$ where $A \in C^{M \times N}$, $M \geq N$ which were used in chapter 5. Because the Gaussian Elimination (from LINPACK) and the least-squares minimum norm (LSMN, from Ney [1983]) can be found elsewhere only the Tikhonov regularization (subroutine TIKHREGD) and conjugate gradient method (subroutine CONJGRAD) are given here. The addition of the clipping constraints are can easily be handled outside this routine, and are omitted.

Subroutine TIKHREGD computes the function which minimizes the additive functional of eq. (69), where f is assumed to be a two-dimensional function. Parameter IFUNC specifies which regularizing functional $\Omega(f)$ to use, with the derivatives of f approximated by finite differences. Normally, IFUNC=4, or $\Omega(f) = \|f_{xx}\|_2^2$. The resulting linear system with input parameter ALPHA (α) is solved using LINPACK subroutines ZGECO and ZGESL.

Subroutine CONJGRAD is called with an initial guess, X. CONJGRAD then performs conjugate gradient iterations until the l_2 convergence criterion $\|x_i - x_{i-1}\|_2 \leq$ EPS is satisfied, restarting the algorithm every N iterations in order to reduce roundoff error. Input parameter MAXITER limits the maximum number of iterations. After each iteration, subroutine CGCST may optionally be called to find a better direction vector P and/or better direction vector scale factor ALPHA. As choosing direction vectors is a more involved procedure, usually the scale factor is "optimized". Table 9 shows the results of applying this

subroutine to the linear system

$$H_n x = y$$

where H_n is the Hilbert matrix of order n, $(h_{i,j} = 1/(i+j-1))$, and $x = (1, 2, \ldots, n)$.

Order of Hilbert matrix n	κ_2 (ZSVDC)	No. of iterations (EPS=1.0D-12)	$\|x - \hat{x}\|_2$	$\|Kf - g\|_2$
2	1.93×10^{1}	3	3.17×10^{-16}	0.00
3	5.24×10^{2}	5	1.26×10^{-15}	5.55×10^{-17}
4	1.55×10^{4}	9	2.61×10^{-13}	8.33×10^{-17}
5	4.77×10^{5}	11	2.70×10^{-7}	8.87×10^{-13}
6	1.50×10^{7}	14	3.04×10^{-2}	3.29×10^{-9}
7	4.75×10^{8}	22	8.34×10^{-3}	8.55×10^{-10}
8	1.53×10^{10}	23	2.08×10^{-2}	3.73×10^{-10}
9	4.93×10^{11}	25	4.01×10^{-2}	2.14×10^{-9}
10	1.60×10^{13}	21	6.69×10^{-2}	8.11×10^{-9}
11	5.23×10^{14}	24	1.91×10^{-2}	1.09×10^{-10}
12	1.70×10^{16}	29	1.01×10^{-1}	3.94×10^{-8}
13	5.77×10^{17}	26	4.41×10^{-2}	1.62×10^{-9}
14	2.13×10^{18}	20	6.16×10^{-2}	2.12×10^{-9}
15	1.60×10^{18}	26	8.29×10^{-2}	4.47×10^{-9}
16	2.77×10^{18}	25	1.08×10^{-1}	8.70×10^{-9}
19	4.38×10^{18}	17	2.10×10^{-1}	4.40×10^{-8}
20	8.38×10^{18}	17	2.53×10^{-1}	6.89×10^{-8}

Table 9: Testing of subroutine CONJGRAD with the Hilbert matrix, H_n
The conditions numbers κ_2 computed by LINPACK subroutine ZSVDC are meaningful up to about 3.6×10^{16}, which is approximately the reciprocal of machine precision, u. $\hat{x}$ is the computed solution.

Subroutine CONJGRAD

```
C MATH_CG.FOR - Solution of the linear system Ax=y by conjugate gradients.
C Location        : [KOZLOWSKI.LIBRARY]@UOTELG01.BITNET
C Class           : linear equations
C Source for code : A.J. Kozlowski, Univ. of Ottawa, Dept. of Elec. Eng.,
C                   Ottawa, K1N 6N5    CANADA.
C Reference       :
C Date            : 1986
C Last mod        : March 9, 1987 (AJK) - retcode added.
C                   May 25, 1987 (AJK) - Extend norms to (6).
C Language        : VAX FORTRAN-77 (DOUBLE: COMPLEX)
C Type            : library routine (MATHLIB)
C External requ.  : Subroutines: CSTCG (user supplied, for constrained solutions
C      only).
C                   Libraries: FEXLIB(getopt),
C      MATHLIB(INNERPRODUCT,ZMPAHB,ZMPROD,ZMSUBR)
C                   Data: none
C                   Packages: none
C Contents        : GLOBAL = .
C                   LOCAL = .
C Description     : Solves a rectangular linear system of equations
C                                A X = Y
C     where A is an M by N matrix, and both X and Y are vectors,
C     using the method of conjugate gradients. Constraints may be added
C     through a user-supplied subroutine CGCST.
C      TEST PROGRAM :
C Compile/link    :
C Backup          : tape
C-


      SUBROUTINE CONJGRAD (iopt, A,m,im,n, x, y, maxiter,eps, norms,
     +      r,p,t,u, retcode )
C Solves the operator equation Ax=y by the method of conjugate gradients,
C converging when the l2-norm ||x(i)-x(i-1)||_2 of the x update
C vector <= EPS.
C Note that the l_2 norm is tested against EPS and ***NOT*** the square
C of the norm.
C See:
C    M.R. Hestenes, E. Stiefel,
C    "Method of Conugate Gradients for Solving Linear Systems",
C    J. Research NBS,1952.
C Notes:
C   1. This procedure requires the following work areas:
C      R - max(m,n) -vector
C      P - n-vector
C      T - m vector
C      u - n -vector
C   2. This code "restarts" the algorithm every "N" iterations, i.e. the
C      current solution vector "X" is used to recompute the residuals, etc.
C      from their initial expressions. This reduces the roundoff error
C      accumulation.
C
C   IOPT -
C     Bit #/add-value      Description
```

```
C      0/1            =0 No (terminal) output during iterations.
C       ,             =1 After each iterations, output the solution NORMS.
C      1/2            =1 Use X as the initial guess. Otherwise, use the
C                        0-vector.
C      2/4            =1 Find a constrained CG solution. The user must supply
C                        a subroutine CGCST (see below) to apply the
C                        constraints.
C   A(C16:(m/im,n):INPUT) .
C   IM is the row DIMENSION of A.
C   M,N the number of rows and columns respectively in A.
C      N is also the row DIMENSION of X.
C   X(C16(n/n,1 or m):OUTPUT)  contains the LSMN solution of [A]X = B.
C   Y(C16(m):INPUT) - contains the LHS vector.
C   R(C16():OUTPUT) - residual vector
C   P(C16():OUTPUT) - work area
C   T(C16(max(m,n)):OUTPUT) - work area.
C   U(C16:n) - work area.
C
C   MAXITER(I:INPUT) - Maximum number of iterations. On output, is the
C         the number of iterations performed.
C   EPS(R8:INPUT) - Maximum l_2 difference between successive iterations.
C   NORMS(6:OUTPUT) - The l_1,l_2 and l_infinity norms of the last two
C      solution vectors, and the l_1, l_2 and l_infinity norms of
C      A x - y on output.
C   RETCODE(I:OUTPUT) = 0 The algorithm terminated successfully.
C      =+1 The algorithm terminated because ||Ap||_2^2 = 0.
C Local work areas limits the number of columns to "MAXCOL".
      INTEGER           iopt
      DOUBLE COMPLEX    A(im,1)
      DOUBLE COMPLEX    x(n)
      DOUBLE COMPLEX    y(m)
      DOUBLE COMPLEX    r(n),p(n)
      DOUBLE COMPLEX    t(m)
      DOUBLE COMPLEX    u(n)
      INTEGER           maxiter
      DOUBLE PRECISION eps
      DOUBLE PRECISION norms(6)
C-
C Local .....
C
      DOUBLE COMPLEX    alpha,beta
      DOUBLE COMPLEX    normAp2,normr2
      DOUBLE COMPLEX    ipApr
      DOUBLE COMPLEX    cjunk
      DOUBLE PRECISION sum
      INTEGER           mt,nt,mr,nr
      LOGICAL           opt(16),done
      LOGICAL           constrained
      DOUBLE COMPLEX    DREAL,DCMPLX
C 0 - Execution begins by initialization.
C     Check if the algorithm terminated in a previous call.
      retcode = 0
      CALL GETPAR (iopt, opt)
      constrained = opt(3)
      IF(.not.opt(2)) THEN
```

```
      DO 1 i=1,n
1       x(i) = (0.0D0,0.0D0)
      ENDIF
      DO 2 i=1,6
2       norms(i) = 0.D0
C*************************    Main Iteration Loop    ****************************
      iternumber = 0
      done = .FALSE.
      DO WHILE (.not.done)
        iternumber = iternumber + 1
        IF(MOD(iternumber,n).EQ.1) THEN
C         Restart the algorithm every "N" iterations to reduce round-off.
C         Process the initial guess: x_0: t0=Ax, r0=(Ax-y),p0=A* r0
          CALL ZMPROD (A,m,im,n, x,n,n,1,  t,mt,m,nt)
          CALL ZMSUBR (y,m,m,1,  t,m,m,1,  r,mr,m,nr)
          CALL ZMPAHB (A,m,im,n, r,m,m,1,  p,mp,n,np)
        ENDIF
C       Compute alpha, and update x.
        CALL ZMPROD       (A,m,im,n, p,n,n,1, t,mt,m,nt)
        CALL INNERPRODUCT (t,r, m,1,1, ipApr)
        CALL INNERPRODUCT (t,t, m,1,1, normAp2)
        IF(CDABS(normAp2).EQ.0.0D0) GOTO 800
        alpha = ( ipApr/normAp2 )

C       - Update the solution vector, x_{n+1} = x_n + alpha * p_n.
        CALL CG_NORMS(0,c,norms)
        IF(constrained) THEN
          CALL CGCST (1,iternumber,A,m,im,n,x,y,r,alpha,p)
        ENDIF
        DO 10 i=1,n
          CALL CG_NORM(1,alpha*p(i), norms)
10        x(i) = x(i) + alpha * p(i)
        norms(2) = DSQRT(norms(2))

C       - Update r
        DO 14 i=1,m
14        r(i) = r(i) - alpha * t(i)
        CALL INNERPRODUCT (r,r, m,1,1, normr2)

        CALL CG_BETA (A,m,im,n, r,p,t,u, normAp2, alpha,beta)

          DO 11 i =1,n
11          p(i) = u(i) - beta * p(i)
C ------> End of ITERATION computations.
        IF(opt(1)) CALL CG_OUT(iternumber,A,m,im,n,x,y, norms)
        done = (iternumber.GE.MAXITER) .OR. (norms(2).LE.EPS) .OR.
     +         (CDABS(normr2).EQ.0.0D0)
        ENDDO
      maxiter = iternumber
C     Finally, compute the ||A x - y|| norms
      CALL ZMPROD (A,m,im,n, x,n,n,1, t,mt,m,nt)
      CALL ZMSUBR (y,m,m,1,  t,m,m,1, r,mr,m,nr)
C     ... r = A x - y at this point.
80    CONTINUE
      CALL CG_NORMS (0,c,norms(4))
```

```
      DO 50 i=1,m
50      CALL CG_NORMS (1,r(i), norms(4))
      norms(5) = DSQRT(norms(5))
90    CONTINUE
      RETURN
C ------------------ / Error Handling ----------------------------
C     Can not proceed further because ||Ap||^2 = 0, therefore
C     alpha will be infinite.
800   CALL ZMPROD (A,m,im,n, x,n,n,1, t,mt,m,nt)
      CALL ZMSUBR (y,m,m,1, t,m,m,1, r,mr,m,nr)
      CALL INNERPRODUCT (r,r, m,1,1, normr2)
      TYPE801,iternumber,CDABS(normr2)
801   FORMAT(' CONJGRAD - ||Ap||^2 = 0 at iteration #',I6,
     +   ' with ||Ax-b||_2 = ',1PE10.2)
      maxiter = iternumber - 1
      retcode = 1
C     Compute ||Ax-y|| norms before leaving ...
      GOTO 80
      END


      SUBROUTINE CG_BETA (A,m,im,n, r,p,t,u, normAp2, alpha,beta)
C Compute "BETA", with t = A p on entry.
      DOUBLE COMPLEX   A(im,n)
      INTEGER          m,im,n
      DOUBLE COMPLEX   r(n)
      DOUBLE COMPLEX   p(n)
      DOUBLE COMPLEX   t(m)
      DOUBLE COMPLEX   u(n)
      DOUBLE COMPLEX   normAp2
      DOUBLE COMPLEX   alpha,beta
C-
      DOUBLE COMPLEX   ip,s

      CALL ZMPAHB(A,m,im,n, r,m,m,1, u,mu,n,nu)

C     Now, compute beta. ip = ( Ap = t, A A* r = A u )
      ip = (0.0D0,0.0D0)
      DO 11 i=1,n
        s = (0.0D0,0.0D0)
        DO 10 j=1,m
10        s = s + a(i,j) * u (j)
        ip = ip + t(i) * DCONJG(s)
11      CONTINUE

      beta =  ( ip / normAp2 )
      RETURN
      END


      SUBROUTINE CG_NORMS (ifunc, dx, norms)
C Updates the 1-,1_2 and 1_infinity norms of the difference vector:
C    x - xi. This routine is called before each component of 'x'
C    is updated.
```

```
C     IFUNC = 0, initialized the norms-vector, and =1 updates it.
      INTEGER          ifunc
      DOUBLE COMPLEX   dx
      DOUBLE PRECISION norms(3)
C-
      DOUBLE PRECISION s
      DOUBLE COMPLEX   CDABS
      IF(ifunc.eq.0) THEN
        norms(1) = 0.0D0
        norms(2) = 0.0D0
        norms(3) = 0.0D0
       ENDIF
C
      IF(ifunc.eq.1) THEN
        s = CDABS( dx )
        norms(1) = norms(1) + s
        norms(2) = norms(2) + s**2
        IF(s.gt.norms(3)) norms(3) = s
       ENDIF

      RETURN
      END



      SUBROUTINE CG_OUT (iternumber,a,m,im,n,x,y, norms)
C Output the data for this iteration.
C Local switches:
C     OUTKfg = .TRUE. -> Output the values K f - g.
      INTEGER          iternumber
      DOUBLE COMPLEX   a(im,1)
      DOUBLE COMPLEX   x(n)
      REAL*8           norms(3)
      DOUBLE COMPLEX   y(n)
C-
      DOUBLE COMPLEX   s
      DOUBLE PRECISION NAxmy2
      LOGICAL          OUTxi,OUTAxmy,CUTKfmg,OUTNKfmg
      DATA OUTxi,OUTAxmy,OUTKfmg,OUTNKfmg
     +                 /.FALSE.,.FALSE.,.FALSE.,.TRUE./
      TYPE100,iternumber,norms
      IF(OUTxi) TYPE102,(i,x(i),i=1,n)
      NAxmy2 = 0.0D0
      DO 20 i=1,m
        s = (0.D0,0.D0)
        DO 21 j=1,n
21        s = s + a(i,j) * x(j)
        IF(OUTAxmy) THEN
          TYPE520,y(i),s,i,(DREAL(a(i,j)*x(j)),
     +                     DIMAG(a(i,j)*x(j)),j=1,n)
         ENDIF
        NAxmy2 = NAxmy2 + CDABS(s-y(i))**2
20      CONTINUE
      TYPE528,DSQRT(NAxmy2)
      RETURN
```

```
100   FORMAT(' CG Alg. Iter number =',I5/
     +            ' Norms (l_1,l_2,l_infinity) =',3d14.4)
102   FORMAT(', x(',i3,') = ',D14.7,1X,D14.7)
520   FORMAT(' A x = y = ',G11.3,1X,G11.3,
     +       '. Sig_Kf = ',G11.3,1X,G11.3,
     +            ': row # ',I4/
     +(' -> ',4(1X,1PE8.1,'|',1PE8.1)/)
     +)
528   FORMAT(' ||Ax-y||_(2) = ',D10.2)
      END


      SUBROUTINE CGCST (iopt,iternumber,n,x,alpha,p)
C User supplied subroutine to add constraints to CG:
C   - step size, ALPHA
C   - search direction, P
C On INPUT,
C   X - the last updated value of the unknown vector
C   ALPHA -
C   P -
C On OUTPUT,
C   X - Same as input.
C   ALPHA - Determined by this subroutine.
C   P - Determined by this subroutine, normally same as on INPUT.
C On RETURN,
C   The new "X" is determined by :        X <- X + ALPHA . P
      INTEGER          iopt
      INTEGER          iternumber
      INTEGER          n
      DOUBLE COMPLEX   x(n)
      DOUBLE COMPLEX   alpha
      DOUBLE COMPLEX   p(n)
C-
      STOP 'CGCST - This subroutine must be user-supplied.'
      END


      SUBROUTINE ZMPROD(A,ma,ima,na, B,mb,imb,nb, C,mc,imc,nc)
C Computes the matrix product A B -> C.
C A(I) -
C MA,IMA,NA(I) -
C B(I) -
C MB,IMB,NB(I) -
C C(O) -
C MC,NC(O) -
C IMC(I) -
      COMPLEX*16 a(ima,1)
      COMPLEX*16 b(imb,1)
      COMPLEX*16 c(imc,1)
C-
      IF(na.ne.mb) GOTO 80
      mc = ma
      nc = nb
```

```
      DO 2 i=1,mc
        DO 2 j=1,nc
          c(i,j) = (0.0d0,0.0d0)
          DO 1 k=1,na
1           c(i,j) = c(i,j) + a(i,k)*b(k,j)
2       CONTINUE
      GOTO 90
80    TYPE*,'ZMPROD - Dimensions of A,B are incomptable.'
90    CONTINUE
      RETURN
      END


      SUBROUTINE ZMPAHB (A,ma,ima,na, B,mb,imb,nb, C,mc,imc,nc)
C Computes the matrix product A* B -> C, where A* = conjugate transpose of A.
C Notes:
C   1. A and C must necessarily be different matrices.
C A(I) -
C MA,IMA,NA(I) -
C B(I) -
C MB,IMB,NB(I) -
C C(O) -
C MC,NC(O) -
C IMC(I) -
      DOUBLE COMPLEX a(ima,1)
      INTEGER        ma,na
      DOUBLE COMPLEX b(imb,1)
      INTEGER        mb,nb
      DOUBLE COMPLEX c(imc,1)
      INTEGER        mc,nc
C-
C Local .....
      DOUBLE COMPLEX DCONJG
      IF(ma.ne.mb) GOTO 80
      mc = na
      nc = nb
      DO 2 i=1,mc
        DO 2 j=1,nc
          c(i,j) = (0.0d0,0.0d0)
          DO 1 k=1,ma
1           c(i,j) = c(i,j) + DCONJG(a(k,i))*b(k,j)
2       CONTINUE
      GOTO 90
80    TYPE*,'ZMPAHB - Dimensions of A,B are incomptable.'
90    CONTINUE
      RETURN
      END


      SUBROUTINE INNERPRODUCT(A,B, m,inca,incb, C)
C Computes the inner product (C) of two vectors:
C     C = A . B*, where
C     INCA and INCB are the increments for the elements in A and B
```

```
C     respectively.
      COMPLEX*16 a(m),b(m)
      INTEGER m
      COMPLEX*16 C
C-
C Local .....
      COMPLEX*16 DCONJG
      c = (0.0d0,0.0d0)
      DO 1 i=1,m
1       c = c + a(1+(i-1)*inca)*DCONJG(b(1+(i-1)*incb))
      RETURN
      END
```

Subroutine TIKHREGD

```
C MATH_REG.FOR - Solution of linear system by (Tiknonov) regularization.
C Location        : [KOZLOWSKI.LIBRARY]@UOTELG01.BITNET
C Class           : linear equations
C Source for code : A.J. Kozlowski, Univ. of Ottawa, Dept. of Elec. Eng.,
C                   Ottawa, K1N 6N5    CANADA.
C Reference       :
C Date            : January    , 1987.
C Last mod        : March 3, 1987 (AJK) - TIKHREGD.
C Language        : VAX FORTRAN-77 (DOUBLE COMPLEX, REAL)
C Type            : library routine (MATHLIB)
C External requ.  : Subroutines: none
C                   Libraries: none
C                   Data: none
C                   Packages: none
C Contents        : GLOBAL = .
C                   LOCAL = .
C Description     :
C      TEST PROGRAM :
C Compile/link    :
C Backup          : tape
C-
```

```
      SUBROUTINE TIKHREGD (iopt, ifunc,  alpha,
     +     K,mk,dimk,nk, f,mf,nf,h, g,
     +     rcond, A,L,LAL, ipvt,  retcode)
C Solves the MK by NK linear system
C     K                f = g
C where f is a two-dimensional MF by NF function with cell separation
C of H (in both the M and N directions).
C The NK by NK linear system
C     (K* K + alpha L*L) f = g
C This linear system is stored as
C            A f = g,      where A is nk x nk (nk=n).
C and solved by finding the unique solution which minimizes
C
C     M(f) =   ||K f - g ||_2^2 + alpha. Omega(f)
C
C where Omega is defined by adding f and its derivatives according to IFUNC.
C
C IFUNC=      Add value  Omega(f) = Sum(||_||_2^2) of
C         1:     1          f
C         2:     2          f_x
C         3:     4          f_y
C         4:     8          f_xx
C         5:    16          f_yy
C         6:    32          f_xy
C This system has the following properties:
C    1. It is (in theory) better conditioned than the original system for
C       an appropriate ALPHA.
C INPUT:
C   K((mk,nk);I)
C   F(nk;I) - is a vectorized array of size nk = mf(rows).nf(columns) stored
```

```
C       row by row.
C
C INPUT/OUTPUT:
C   G(nk;IO) - Destroyed on output.
C
C WORK AREAS:
C   A((mk,mk);O) - work area.
C   L((mk,nk);O) - work area.
C   LAL((mk,mk);O) - work area.
C   IPVT(nk;O) - work area of size NK.
C
C OUTPUT:
C   RCOND(;O) - condition number estimate.
C
C Notes:
C    1. The linear system is solved by LINPACK subroutines.
C    2. Three (3) large work areas are required:
C       - A must be at least of size NK by NK.
C       - L
C       - LAL
C    3. The derivatives are approximated numerically (see code for details).
C    4. This is a modified version of subroutine TIKREGI.
C External requ.  : Subroutines: none
C                   Libraries: FEXLIB(getpar),
C                    MATHLIB(zmpahb(1),zgeco,zgesl)
C                   Data: none
C                   Packages: none
C     (1) zmpahb performs the matrix computatation of A^H B where A^H is the
C         complex conjugate transpose of A.
      INTEGER           iopt
      DOUBLE PRECISION alpha
      INTEGER           mk,dimk,nk
      DOUBLE COMPLEX    K(dimk,1)
      DOUBLE COMPLEX    f(1)
      INTEGER           mf,nf
      DOUBLE PRECISION h
      DOUBLE COMPLEX    g(1)
C
      DOUBLE PRECISION rcond
      DOUBLE COMPLEX    A(mk,1)
      DOUBLE COMPLEX    L(mk,nk)
      DOUBLE COMPLEX    LAL(nk,nk)
      INTEGER           ipvt(1)
      INTEGER           retcode
C-
      INTEGER           dimwa
      LOGICAL           Omega(16)
      if(i,j) = (j-1)*mf + i
C
      CALL GETPAR (ifunc,Omega)
C 1 - Create the new linear system, A f = g.
C     - LHS : K*K + alpha I -> A
      CALL ZMPAHB (K,mk,dimk,nk, K,mk,dimk,nk,    A,n,nk,n)
C ------------------------  D e r i v a t i v e s  ------------------------
      DO 8 i=1,mk
```

```
      DO 8 j=1,nk
8       L(i,j) = (0.0D0,0.0D0)
C     f ...
      IF(Omega(1)) THEN
        DO 1 i=1,n
1         L(i,i) = (1.0D0,0.0D0)
       ENDIF
C     f_x ...
      IF(Omega(2)) THEN
        DO 2 j=1,nf
        DO 2 i=1,mf
          jj = if(i,j)
          IF((2.LE.i) .AND. (i.LE.mf-1)) THEN
C            (F(i+1,j) - F(i-1,j))/(2.D0*h) + L(i,j)
           L(jj,if(i+1,j)) = +1.0D0/h + L(jj,if(i+1,j))
           L(jj,if(i-1,j)) = -1.0D0/h + L(jj,if(i-1,j))
           ELSEIF (i.EQ.1) THEN
C            (F(2,j) - F(1,j))/h + L(i,j)
           L(jj,if(2,j)) = +1.0D0/h + L(jj,if(2,j))
           L(jj,if(1,j)) = -1.0D0/h + L(jj,if(1,j))
           ELSEIF (i.EQ.mf) THEN
C            (F(mf,j) - F(mf-1,j))/h + L(i,j)
           L(jj,if(mf,j)  ) = +1.0D0/h + L(jj,if(mf,j))
           L(jj,if(mf-1,j)) = -1.0D0/h + L(jj,if(mf-1,j))
           ENDIF
2         CONTINUE
       ENDIF
C     f_y ...
      IF(Omega(3)) THEN
        DO 3 j=1,nf
        DO 3 i=1,mf
          jj = if(i,j)
          IF((2.LE.j) .AND. (j.LE.nf-1)) THEN
C             (F(i,j+1) - F(i,j-1))/(2.D0*h) + L(i,j)
           L(jj,if(i,j+1)) = +1.0D0/h + L(jj,if(i,j+1))
           L(jj,if(i,j-1)) = -1.0D0/h + L(jj,if(i,j-1))
           ELSEIF (j.EQ.1) THEN
C            (F(i,2) - F(i,1))/h + L(i,j)
           L(jj,if(i,2)) = +1.0D0/h + L(jj,if(i,2))
           L(jj,if(i,1)) = -1.0D0/h + L(jj,if(i,1))
           ELSEIF (j.EQ.mf) THEN
C            (F(i,nf) - F(i,nf-1))/h + L(i,j)
           L(jj,if(i,nf  )) = +1.0D0/h + L(jj,if(i,nf))
           L(jj,if(i,nf-1)) = -1.0D0/h + L(jj,if(i,nf-1))
           ENDIF
3         CONTINUE
       ENDIF
C     f_xx ...
      IF(Omega(4)) THEN
        DO 4 j=1,nf
        DO 4 i=1,mf
          jj = if(i,j)
          IF((2.LE.i) .AND. (i.LE.mf-1)) THEN
C             (F(i+1,j) - 2.D0*F(i,j) + F(i-1,j))/(h**2)
           L(jj,if(i+1,j)) = +1.0D0/h**2 + L(jj,if(i+1,j))
```

```
            L(jj,if(i  ,j)) = -2.0D0/h**2 + L(jj,if(i  ,j))
            L(jj,if(i-1,j)) = +1.0D0/h**2 + L(jj,if(i-1,j))
            ENDIF
4         CONTINUE
        ENDIF
C     f_yy ...
      IF(Omega(5)) THEN
        DO 5 j=1,nf
        DO 5 i=1,mf
          jj = if(i,j)
          IF((2.LE.j) .AND. (j.LE.nf-1)) THEN
C            (F(i,j+1) - 2.D0*F(i,j) + F(i,j-1))/(h**2)
           L(jj,if(i,j+1)) = +1.0D0/h**2 + L(jj,if(i,j+1))
           L(jj,if(i,j  )) = -2.0D0/h**2 + L(jj,if(i,j  ))
           L(jj,if(i,j-1)) = +1.0D0/h**2 + L(jj,if(i,j-1))
           ENDIF
5         CONTINUE
        ENDIF
C     f_xy ...
      IF(Omega(6)) THEN
        DO 6 j=1,nf
        DO 6 i=1,mf
          jj = if(i,j)
          IF((2.LE.i) .AND. (i.LE.mf-1)) THEN
C           (F(i+1,j+1) - F(i+1,j-1) -F(i-1,j+1) + + F(i-1,j-1))/
C                              (4.D0*h**2)
           L(jj,if(i+1,j+1)) = +1.0D0/(4.D0*h**2) + L(jj,if(i+1,j+1))
           L(jj,if(i+1,j-1)) = -1.0D0/(4.D0*h**2) + L(jj,if(i+1,j-1))
           L(jj,if(i-1,j+1)) = -1.0D0/(4.D0*h**2) + L(jj,if(i-1,j+1))
           L(jj,if(i-1,j-1)) = +1.0D0/(4.D0*h**2) + L(jj,if(i-1,j-1))
           ENDIF
6         CONTINUE
        ENDIF

      dimwa = n
      CALL ZMPAHB (L,mk,dimk,nk, L,mk,dimk,nk,   LAL,mg,dimwa,ng)

      DO 9 j=1,n
      DO 9 i=1,n
9       a(i,j) = a(i,j) + alpha * lal(i,j)
C ------------------------------------------------------------------
C     - RHS : K*g              -> g
      dimwa = n
      CALL ZMPAHB (K,mk,dimk,nk, g,mk,mk,1,   lal,mg,dimwa,ng)
      DO 10 i=1,n
10      g(i) = lal(i,1)
C 2 - Solve the linear system using LINPACK.
      CALL ZGECO (A,nk,n, ipvt,rcond,lal)
      CALL ZGESL (A,nk,n, ipvt,g,0)
      DO 20 i=1,n
20      f(i) = g(i)
C
      retcode = 0
      RETURN
      END
```

Appendix E

MATHEMATICAL PRELIMINARIES

Some concepts from functional analysis which are needed to study the properties of the integral equation eq. (33) are reviewed here. This theory is needed to explain the difficulties involved in inverting the equation to find the "solution". The contents of this chapter should not be treated as obvious, as many authors do not define these concepts precisely in their work. More detail can be found in books on functional analysis, e.g. Stakgold [1979].

Consider the abstract form of equation (33)

$$Kf = g \tag{127}$$

where $f \subset X$ and $g \subset Y$. Alternatively, it can be expressed as

$$K : X \mapsto Y \tag{128}$$

In either case, K is a linear operator which assigns to each element of a metric space X an element in a metric space Y. X is called the *domain* $\mathrm{dom}(K)$ of the operator K, and Y is called the *range* $\mathrm{ran}(K)$ of the operator K.

$$\mathrm{ran}(Y) = \{g | g = Kf \quad \text{for some} \quad f \in X\} \tag{129}$$

A subspace which is of interest is the null space or *kernel* of K, and is the set of elements mapped to zero

$$\ker(K) = \{f \in X | Kf = 0\} \subseteq X. \tag{130}$$

E.1 Linear Operators

Some properties of function spaces (a linear space whose elements are functions) are now reviewed. A *metric space* is a space upon which a metric (distance function) is defined (see Naylor and Sell [1982, p.48+]).

INNER PRODUCT. An inner product $(,)$ is a scalar defined on a space X, which satisfies the following conditions.

1. $(x, x) \geq 0$ and $(x, x) = 0$ if and only if $x = 0$.
2. $(x + y, z) = (x, z) + (y, z)$ for all x, y, z in X
3. $(\lambda x, y) = \lambda(x, y)$ for all x, y in X and $\lambda \in C$
4. $(x, y) = \overline{(y, x)}$ for all x, y in X
5. If $\{x_n\} \subset X$ and $\lim(x_n - x_m, x_n - x_m) \to 0$ as $n, m \to \infty$, then there is an element $x \in X$ such that $\lim_{n\to\infty}(x_n - x, x_n - x) = 0$. The number $\|x\|_2 = \sqrt{(x, x)}$ is called the *norm* of x.

For analytical simplicity, we shall deal with the inner product

$$(u, v) = \int_\Omega u\bar{v}d\Omega \tag{131}$$

which gives the norm

$$\|u\|_2 = \sqrt{\int_\Omega |u|^2 d\Omega} \tag{132}$$

and metric

$$d(u, v) = \|u - v\|_2 \tag{133}$$

where the integrations are over the domain of u and v.

This space of functions (called L_2) is a special case of a Hilbert space. A non-empty set H is called a Hilbert space if H is a complex linear space on which an inner product is defined (see Friedman [1982, p. 201]). H is complete in the metric generated by the inner product. If in addition H contains a countable set of vectors which span H (called the *spanning set*), then the Hilbert space is said to be *separable* (Stakgold [1979, p. 276]).

A useful auxiliary operator is the adjoint.

ADJOINT. The adjoint of a bounded operator K (denoted K^* is defined by the relation

$$(Ku, v) = (u, K^*v) \tag{134}$$

where $u, v \in C$. An operator is said to be formally self-adjoint if $K^* = K$. If in addition, the boundary conditions for K and K^* define the same space, then the operator is said to be self-adjoint.

Operators have a number of properties which give their solutions different properties.

BOUNDED. A linear operator $K : X \mapsto Y$ where X and Y are metric spaces is bounded if there exists a real number $M \geq 0$ such that

$$\|Kf\|_Y \leq M\|f\|_X$$

CLOSED SET. A set S is called a closed set if its complement is an open set. (A linear subspace S of any finite-dimensional normed linear space X is always closed.)

CONTINUOUS. A function f is continuous at $\underline{p}_0$ if for each $\varepsilon > 0$ there exists a $\delta > 0$ such that $|f(\underline{p}) - f(\underline{p}_0)| < \varepsilon$ for all $\underline{p} \in \text{dom}(f)$ with $\|\underline{p} - \underline{p}_0\| < \delta$ (Todd [1977, p. 21]).

COMPACT. An operator K is compact if every sequence $\{x_n\}(x_n \in X)$ has a subsequence $\{x_{n_j}\}$ such that the sequence $\{Kx_{n_j}\}$ converges in Y. (Note that finite-dimensional operators are compact (see Bronshtein and Semendyavev [1985, p. 772]).)

DENSE. A set S is dense in a set $T \supset S$, if for each element $u \in T$ and each $\varepsilon > 0$, there exists an element $v \in S$ such that

$$d(u, v) < \varepsilon$$

Thus, every element $u \in T$ can be approximated to arbitrary precision by elements of S (Stakgold [1979, p. 240]).

OPEN SET. A set E is open if E contains each local neighborhood of each of its points.

POSITIVE/POSITIVE DEFINITE. A bounded linear self-adjoint operator $K : X \mapsto Y$ is said to be *positive definite* if there exists a real constant M such that $\overline{(Kf, f)} \geq M(f, f)$ for all $f \in X$. If $(Kf, f) \geq 0$ for all $f \in X$, then K is said to be *positive*.

Another concept is dimensionality. Let X be a metric space. A sequence of elements $\{x_i\}_{i=1}^n$ is said to be a *basis* of X if every $x \in X$ can be uniquely represented in the form

$$x = \sum_{i=1}^{n} \alpha_i x_i$$

The smallest n is called the *dimension* of X and is denoted as $\dim(X)$. If $n = \infty$, then the space X is said to be *infinite dimensional*, otherwise it is *finite dimensional*.

ORTHOGONAL. Two elements are orthogonal, denoted $x \perp y$, if

$$(x, y) = 0$$

E.2 Finite Dimensional Operators

Although the operators we deal with are continuous, it is necessary to discretize them somehow in order to compute with them. In this case, we transform them into finite dimensional operators, and, perhaps change their properties. The most convenient form to consider discretized operators is matrix form. Hence, linear algebra concepts must also be considered.

A^H is the conjugate transpose of $A \in C^{m \times n}$, and the elements are given by

$$a_{ij}^H = \overline{a_{ji}}$$

Consider a linear operator K which is equivalent to a matrix operator A. In the case that A is the matrix of K relative to an orthonormal basis of the space X, then the conjugate transpose of A, A^H, is the matrix of K^* in the same orthonormal basis, i.e.

$$K^* = K^H \tag{135}$$

(see Lipschutz [1968, p. 284]).

The *rank* of matrix A, denoted by rank(A), is the common dimension of its row space and its column space, i.e. dim(ran(A)).

There are some simplifications when dealing with finite dimensional systems. A finite dimensional space is always separable since any basis can serve as a finite spanning set.

Bibliography

[1] M. Abramowitz, I. A. Stegun (eds.), *Handbook of Mathematical Functions* (9th printing), Dover Publ., 1972.

[2] R. W. B. Ardill, K. J. M. Moriarty, "Accurate Bessel Functions $J_n(z)$, $Y_n(z)$, $H_n^1(z)$ and $H_n^2(z)$ of Integer Order and Complex Argument", *Comp Phy Comm*, vol. 17, 321-326, 1979.

[3] H. M. Assenheim, G. W. Hartsgrove, E. Preston, "A Diathermy Applicator for the Irradiation of Rat Brain", *IEEE Trans*, vol. BME-27, no. 8, 476-479, Aug 1980.

[4] M. Baribaud, F. Dubois, R. Floyrac, M. Kom, S. Wang, "Tomographic image reconstitution of biological objects from coherent microwave diffraction data", *IEE Proc*, vol. 129, no. H, 356-359, Dec 1982.

[5] R. H. T. Bates, F. L. Ng, "Polarisation-source formulation of electromagnetism and dielectric-loaded waveguides", *IEE Proc*, vol. 119, no. 11, 1568-1574, Nov 1972.

[6] M. Bertero, "Regularization Methods for Linear Inverse Problems" in *Lecture Notes in Mathematics*, vol. 1225, G. Talenti , ed., 1986, Springer-Verlag, Germany, 52-112, 1986.

[7] A. Bjorck, "Iterative Refinement of Linear Least Squares Solutions I", *BIT*, vol. 7, 257-278, 1967.

[8] A. Bjorck, "Iterative Refinement of Linear Least Squares Solutions II", *BIT*, vol. 8, 8-30, 1968.

[9] N. Bleistein, R. A. Handelsman, *Asymptotic Expansions of Integrals*, , Dover Publ., U. S. A., 1986.

[10] N. N. Bojarski, "Inverse scattering inverse source theory", *J Math Phys*, vol. 22, no. 8, 1647-1650, Aug 1981.

[11] M. Born, E. Wolf, "Principles of Optics", 6th ed., *Pergamon Press*, G. B., 1983.

[12] R. N. Bracewell, *The Fourier Transform and its Applications*, 2nd ed., McGraw-Hill, U. S. A., 1978.

[13] I. N. Bronshtein, K. A. Semendyayev, *Handbook of Mathematics*, K. A. Hirsch, translator, Van Nostrand Reinhold Co., G. D. R, 1985.

[14] J. M. Cameron, "Statistics" in *Fundamental Formulas of Physics*, vol. 1, D. H. Menzel, ed., Dover Publ., U. S. A., 1960.

[15] K. M. Chen, "Interaction of Electromagnetic Fields with Biological Bodies" in *Research Topics in Electromagnetics*, J. A. Kong, ed., J. Wiley & Sons, U. S. A. , 290-347, 1981.

[16] C. H. Chen, C. -D. Lien, "The Variational Principle for Non-Self-Adjoint Electromagnetic Problems", *IEEE Trans*, vol. MTT-28, no. 8, 878-886, Aug 1980.

[17] A. K. Cline, A. R. Conn, C. F. Van Loan, "Generalizing the LINPACK Condition Estimator" in *Lecture Notes in Mathematics*, Springer-Verlag , vol. 909, 73-83, 1982.

[18] S. D. Conte, C. deBoor, *Elementary Numerical Analysis: An Algorithmic Approach*, 3rd ed., McGraw-Hill, U. S. A., 1980.

[19] B. Davies, *Integral Transforms and their Applications* , 2nd ed., Springer-Verlag, U. S. A., 1985.

[20] J. J. Dongarra, C. B. Moler, J. R. Bunch, G. W. Stewart, *LINPACK Users' Guide*, SIAM, U. S. A., 1979.

[21] C. H. Dorny, *A Vector Space Approach to Models and Optimization*, R. E. Krieger Co., U. S. A., 1980.

[22] M. Foster, "An Application of the Wiener-Kolmogorov Smoothing Theory to Matrix Inversion", *J Soc Indus Appl Math*, vol. 9, no. 3, 387-392, Sep 1961.

[23] J. N. Franklin, "On Tikhonov's Method for Ill-Posed Problems", *Math Comput*, vol. 28, no. 128, 889-907, Oct 1974.

[24] A. Friedman, *Foundations of Modern Analysis*, Dover Publications, U. S. A., 1982.

[25] G. Golub, W. Kahan, "Calculating the Singular Values and Pseudo-Inverse of a Matrix", *J SIAM Numer Anal*, Ser. B, vol. 2, no. 2, 205-224, 1965.

[26] G. H. Golub, C. F. van Loan, *Matrix Computations*, Johns Hopkins Univ Press, 1984.

[27] J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, 1968.

[28] R. C. Gonzalez, P. Wintz, *Digital Image Processing*, Addison-Wesley, 1977.

[29] J. Graves, P. M. Prenter, "Numerical Iterative Filters Applied to First Kind Fredholm Integral Equations", *Numer Math*, vol. 30, 281-299, 1978.

[30] J. Greenstadt, "On the Reduction of Continuous Problems to Discrete Form", *IBM J Res Dev*, vol. 3, 355-363, Oct 1959.

[31] J. L. Guerquin-Kern, M. Gautherie, G. Peronnet, L. Jofre, J. C. Bolomey, "Active Microwave Tomographic Imaging of Isolated, Perfused Animal Organs", *Bioelectromagnetics*, vol. 6, 145-156, 1985.

[32] M. J. Hagmann, R. L. Levin, "Criteria for Accurate Usage of Block Models", *J Micro Power*, 19-27, 1987.

[33] R. F. Harrington, *Time-Harmonic Electromagnetic Fields*, Mc-Graw Hill, 1961.

[34] R. F. Harrington, "Matrix Methods for Field Problems", *IEEE Proc*, vol. 55, no. 2, 136-149, Feb 1967.

[35] M. R. Hestenes, E. Stiefel, "Methods of Conjugate Gradients for Solving Linear Systems", *J Res NBS*, vol. 49, no. 6, 409-436, Dec 1952.

[36] M. Hestenes, *Conjugate Direction Methods in Optimization*, Springer-Verlag, U. S. A., 1980.

[37] C. A. Hoer, "Performance of a Dual Six-Port Automatic Network Analyzer", *IEEE Trans*, vol. MTT-27, no. 12, 993-998, Dec 1979.

[38] A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Dover Publ., U. S. A., 1975.

[39] IEEE, *IEEE Standard Dictionary of Electrical and Electronics Terms*, 2nd ed., Wiley & Sons, 1977.

[40] W. A. Imbriale, R. Mittra, "The Two-Dimensional Inverse Scattering Problem", *IEEE Trans*, vol. AP-18, no. 5, 633-642, Sep 1970.

[41] F. John, *Partial Differential Equations*, 4th ed., Springer Verlag, U. S. A., 1982.

[42] D. S. Jones, *The Theory of Electromagnetism*, Pergamon Press, 1964.

[43] W. J. Kammerer, M. Z. Nashed, "Iterative Methods for Best Approximate Solutions of Linear Integral Equations of the First and Second Kinds", *J Math Anal Appl*, vol. 40, 547-573, 1972.

[44] R. Kastner, R. Mittra, "A Spectral-Iteration Technique for Analyzing Scattering from Arbitrary Bodies, Part I: Cylindrical Scatterers with E-Wave Incidence", *IEEE Trans*, vol. AP-31, no. 3, 499-506, May 1983.

[45] M. Kom, *imagerie micro-onde en milieu biologique: application au diagnostic medical*, thèse, Doctorat de 3ème cycle, Univ. de Paris, 1981.

[46] G. Kristensson, C. R. Vogel, "Inverse Problems for acoustic waves using the penalised likelihood method", *Inverse Problems* , vol. 2, 461-479, 1986.

[47] H. J. Landau, H. O. Pollack, "Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty - II", *BSTJ*, vol. 40, 65-84, Jan 1961.

[48] H. J. Landau, H. O. Pollak, "Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty - III: The Dimension of the Space of Essentially Time- and Band-Limited Signals", *BSTJ*, vol. 41, 1295-1336, Jul 1962.

[49] N. N. Lebedev, I. P. Skalskaya, Y. S. Uflyand, *Worked Problems in Applied Mathematics*, transl. by R. A. Silverman, U. S. A., 1965.

[50] J. C. Lin, "Frequency Optimization for Microwave Imaging of Biological Tissues", *IEEE Proc*, vol. 73, no. 2, 374-375, Feb 1985.

[51] S. Lipschutz, *Linear Algebra* (Schaum's Outline Series), McGraw-Hill, U. S. A., 1968.

[52] D. E. Livesay, K. M. Chen, "Electromagnetic Fields Induced Inside Arbitrarily Shaped Biological Bodies", *IEEE Trans*, vol. MTT-22, no. 12, 1273-1280, Dec 1974.

[53] J. R. Mentzer, *Scattering and Diffraction of Radio Waves*, Pergamon Press, U. S. A., 1955.

[54] S. F. McCormick, G. H. Rodrigue, "A Uniform Approach to Gradient Methods for Linear Operator Equations", *J Math Anal Appl*, vol. 49, 275-285, 1975.

[55] R. Mittra, ed. , *Computer Techniques for Electromagnetics*, Pergamon Press, 1973.

[56] C. Müller, "Foundations of the Mathematical Theory of Electromagnetic Waves", Springer-Verlag, Germany, 1969.

[57] M. Z. Nashed, "Aspects of Generalized Inverse in Analysis and Regularization" in *Generalized Inverses and Applications*, M. Z. Nashed, ed., Academic Press, U. S. A., 193-243, 1976.

[58] M. Z. Nashed, "Approximate Regularized Solutions to Improperly Posed Linear Integral and Operator Equations" in *Lecture Notes in Mathematics*, vol. 430, D. L. Colton, R. P. Gilbert, eds., Springer-Verlag, U. S. A., 289-332, 1974.

[59] M. Z. Nashed, "Perturbations and Approximations for Generalized Inverses and Linear Operator Equations" in *Generalized Inverses and Applications*, M. Z. Nashed, ed., Academic Press, U. S. A., 325-396, 1976.

[60] M. Z. Nashed, "Operator-Theoretic and Computational Approaches to Ill-posed Problems with Applications to Antenna Theory", *IEEE Trans* , vol. AP-29, no. 2, 220-231, Mar 1981.

[61] National Bureau of Standards (U. S. A.), *Table of the Bessel Functions* $J_0(z)$ *and* $J_1(z)$ *for Complex Arguments* (2nd ed.), Columbia Univ. Press, U. S. A. , 1947.

[62] National Bureau of Standards (U. S. A.), *Table of the Bessel Functions* $Y_0(z)$ *and* $Y_1(z)$ *for Complex Arguments*, Columbia Univ. Press, U. S. A., 1950.

[63] A. W. Naylor, G. R. Sell, "Linear Operator Theory in Engineering and Science" (Applied Mathematical Sciences; 40), Springer-Verlag, U. S. A., 1982.

[64] M. M. Ney, *Electromagnetic Imaging*, Ph. D. thesis, Univ. of Ottawa, Canada, 1983.

[65] M. M. Ney, A. M. Smith, S. S. Stuchly, "A Solution of Electromagnetic Imaging Using Pseudoinverse Transformation", *IEEE Trans* , vol. MI-3, no. 4, 155-162, Dec 1984.

[66] D. L. Phillips, "A Technique for the Numerical Solution of Certain Integral Equations of the First Kind", *JACM*, vol. 9, 84-96, 1962.

[67] C. Pichot, L. Jofre, G. Peronnet, J. C. Bolomey, "Active Microwave Imaging of Inhomogenous Bodies", *IEEE Trans*, vol. AP-33, no. 4, 416-425, Apr 1985.

[68] A. G. Ramm, *Scattering by Obstacles*, D. Reidel, U. S. A., 1986.

[69] H. J. Reinhardt, "Analysis of Approximation Methods for Differential and Integral Equations" in *Applied Mathematical Science Series*, Springer-Verlag, vol. 57, 1985.

[70] K. Rektorys, *Variational Methods in Mathematics, Science and Engineering*, 2nd ed., D. Reidel Co. , U. S. A., 1980.

[71] J. H. Richmond, "Scattering by a Dielectric Cylinder of Arbitrary Cross Section Shape", *IEEE Trans*, vol. AP-13, 334-341, May 1965.

[72] J. H. Richmond, "TE-Wave Scattering by a Dielectric Cylinder of Arbitrary Cross-Section Shape", *IEEE Trans*, vol. AP-14, no. 4, 460-464, Jul 1966.

[73] V. G. Romanov, *Integral Geometry and Inverse Problems for Hyperbolic Equations*, Springer-Verlag, Germany, 1974.

[74] T. K. Sarkar, K. R. Siarkiewicz, R. F. Stratton, "Survey of Numerical Methods for Solution of Large Systems of Linear Equations for Electromagnetic Field Problems", *IEEE Trans*, vol. AP-29, no. 6, 847-856, Nov 1981.

[75] T. K. Sarkar, "The application of the conjugate gradient method for the solution of operator equations arising in electromagnetic scattering from wire antennas", *Rad Sci*, vol. 19, no. 5, 1156-1172, Sep-Oct 1984.

[76] T. K. Sarkar, "The Conjugate Gradient Method as Applied to Electromagnetic Field Problems", *IEEE AP-Society Newsletter*, pp. 5-14, Aug 1986.

[77] T. B. A. Senior, "Approximate Boundary Conditions", *IEEE Trans* , vol. AP-29, no. 5, 826-829, Sep 1981.

[78] C. Sibbald, *The Automated Network Analyzer*, Dept. of Electrical Eng. , University of Ottawa, May 1985.

[79] M. Slaney, A. C. Kak, L. E. Larsen, "Limitations of Imaging with First-Order Diffraction Tomography", *IEEE Trans*, vol. MTT-32, 860-873, Aug 1984.

[80] F. Smithies, *Integral Equations*, Cambridge Univ. Press, G. B., 1958.

[81] Special Issue on Inverse Methods in Electromagnetics, *IEEE Trans*, vol. AP-29, no. 2, Mar 1981 .

[82] I. Stakgold, *Green's Functions and Boundary Value Problems*, J. Wiley & Sons, U. S. A. , 1979.

[83] W. R. Stone, "An exact, closed-form, analytical solution to the general synthesis problem", *Proc SPIE*, San Diego, CA., vol. 294, 90-95, Aug 25-26, 1981.

[84] W. V. Synder, "Algorithm 531: Contour Plotting", *ACM Trans Math Soft*, vol. 4, no. 3, 290-294, Sep 1978.

[85] F. Szidarovsky, S. Yakowitz, *Principles and Procedures of Numerical Analysis*, Plenum Press, U. S. A., 212-220.

[86] G. A. Thiele, "Wire Antennas" in *Computer Techniques for Electromagnetics*, R. Mittra, ed., Pergamon Press, 7-, 1973.

[87] A. N. Tihonov, "Solution of Incorrectly Formulated Problems and the Regularization Method", *Soviet Math Dokl*, vol. 4, 1035-1038, 1963.

[88] A. N. Tihonov, "Regularization of Incorrectly Posed Problems", *Soviet Math Dokl*, 1624-1627, 1963.

[89] A. N. Tikhonov, V. Y. Arsenin, *Solutions of Ill-Posed Problems*, V. H. Winston & Sons, U. S. A., 1977.

[90] J. Todd, *Basic Numerical Mathematics*, vol. 2: Numerical Algebra , Academic Press, U. S. A., 1978.

[91] G. Tyras, *Radiation and Propagation of Electromagnetic Waves*, Academic Press, 1969.

[92] J. van Bladel, "Some Remarks on Green's Dyadic for Infinite Space", *IRE Trans*, vol. AP-9, no. 6, 563-566, Nov 1961.

[93] S. L. Watkins, "Masked Three-Dimensional Plot Program with Rotations", *Comm ACM*, vol. 17, no. 9, 520-523, Sep 1974.

[94] W. E. Williams, *Partial differential equations*, Clarendon Press, Oxford, 1980.

[95] E. Wolf, "Three-dimensional Structure Determination of Semi-transparent Objects from Holographic Data", *Opt Comm*, vol. 1, no. 4, 153-156, Sep 1969.

[96] T. K. Wu, L. L. Tsai, "Numerical Analysis of Electromagnetic Fields in Biological Tissue", *IEEE Proc*, 1167-1168, Aug 1974.