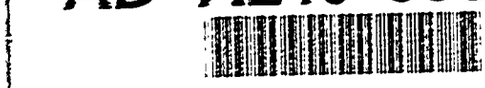
Pitman Research Notes in Mathematics Series 245



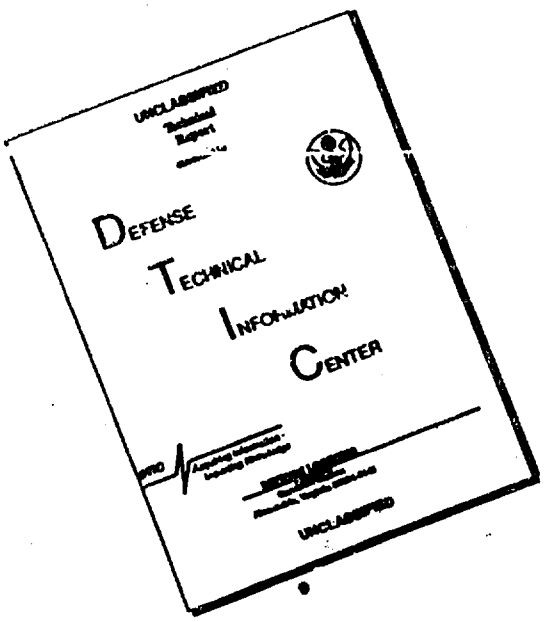
G F Roach (Editor)

Inverse problems and imaging

R & D No. 6105-MA-02
DAJA45-89-M-0086

Longman Scientific & Technical

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

Pitman Research Notes in Mathematics Series

Submission of proposals for consideration
Suggestions for publication, in the form of outlines and representative samples, are invited by the Editorial Board for assessment. Intending authors should approach one of the main editors or another member of the Editorial Board, citing the relevant AMS subject classifications. Alternatively, outlines may be sent directly to the publisher's offices. Refereeing is by members of the board and other mathematical authorities in the topic concerned, throughout the world.

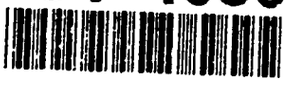
Preparation of accepted manuscripts
On acceptance of a proposal, the publisher will supply full instructions for the preparation of manuscripts in a form suitable for direct photo-lithographic reproduction. Specially printed grid sheets are provided and a contribution is offered by the publisher towards the cost of typing. Word processor output, subject to the publisher's approval, is also acceptable.

Illustrations should be prepared by the authors, ready for direct reproduction without further improvement. The use of hand-drawn symbols should be avoided wherever possible, in order to maintain maximum clarity of the text.

The publisher will be pleased to give any guidance necessary during the preparation of a typescript, and will be happy to answer any queries.

Important note
In order to avoid later retyping, intending authors are strongly urged not to begin final preparation of a typescript before receiving the publisher's guidelines and special paper. In this way it is hoped to preserve the uniform appearance of the series.

Longman Scientific & Technical
Longman House
Burnt Mill
Harlow, Essex, UK
(tel (0279) 426721)



Titles in this series

1 Improperly posed boundary value problems
A Carasso and A P Stone

2 Lie algebras generated by finite dimensional ideals
I N Stewart

3 Bifurcation problems in nonlinear elasticity
R W Dickey

4 Partial differential equations in the complex domain
D L Colton

5 Quasilinear hyperbolic systems and waves
A Jeffrey

6 Solution of boundary value problems by the method of integral operators
D L Colton

7 Taylor expansions and catastrophes
T Poston and I N Stewart

8 Function theoretic methods in differential equations
R P Gilbert and R J Weinacht

9 Differential topology with a view to applications
D R J Chillingworth

10 Characteristic classes of foliations
H V Pittie

11 Stochastic integration and generalized martingales
A U Kussmaul

12 Zeta-functions: An introduction to algebraic geometry
A D Thomas

13 Explicit *a priori* inequalities with applications to boundary value problems
V G Sigillito

14 Nonlinear diffusion
W E Fitzgibbon III and H F Walker

15 Unsolved problems concerning lattice points
J Hammer

16 Edge-colourings of graphs
S Fiorini and R J Wilson

17 Nonlinear analysis and mechanics: Heriot-Watt Symposium Volume I
R J Knops

18 Actions of fine abelian groups
C Kosniowski

19 Closed graph theorems and webbed spaces
M De Wilde

20 Singular perturbation techniques applied to integro-differential equations
H Grabmüller

21 Retarded functional differential equations: A global point of view
S E A Mohammed

22 Multiparameter spectral theory in Hilbert space
B D Sleeman

24 Mathematical modelling techniques
R Aris

25 Singular points of smooth mappings
C G Gibson

26 Nonlinear evolution equations solvable by the spectral transform
F Calogero

27 Nonlinear analysis and mechanics: Heriot-Watt Symposium Volume II
R J Knops

28 Constructive functional analysis
D S Bridges

29 Elongational flows: Aspects of the behaviour of model elasticoviscous fluids
C J S Petrie

30 Nonlinear analysis and mechanics: Heriot-Watt Symposium Volume III
R J Knops

31 Fractional calculus and integral transforms of generalized functions
A C McBride

32 Complex manifold techniques in theoretical physics
D E Lerner and P D Sommers

33 Hilbert's third problem: scissors congruence
C-H Sah

34 Graph theory and combinatorics
R J Wilson

35 The Tricomi equation with applications to the theory of plane transonic flow
A R Manwell

36 Abstract differential equations
S D Zaidman

37 Advances in twistor theory
L P Hughston and R S Ward

38 Operator theory and functional analysis
I Erdelyi

39 Nonlinear analysis and mechanics: Heriot-Watt Symposium Volume IV
R J Knops

40 Singular systems of differential equations
S L Campbell

41 N-dimensional crystallography
R L E Schwarzenberger

42 Nonlinear partial differential equations in physical problems
D Graffi

43 Shifts and periodicity for right invertible operators
D Przeworska-Rolewicz

44 Rings with chain conditions
A W Chatters and C R Hajarnavis

45 Moduli, deformations and classifications of compact complex manifolds
D Sundararaman

46 Nonlinear problems of analysis in geometry and mechanics
M Atteia, D Bancel and I Gumowski

47 Algorithmic methods in optimal control
W A Gruver and E Sachs

48 Abstract Cauchy problems and functional differential equations
F Kappel and W Schappacher

49 Sequence spaces
W H Ruckle

50 Recent contributions to nonlinear partial differential equations
H Berestycki and H Brezis

51 Subnormal operators
J B Conway

52 Wave propagation in viscoelastic media
F Mainardi

53 Nonlinear partial differential equations and their applications: Collège de France Seminar. Volume I
H Brezis and J L Lions

54 Geometry of Coxeter groups
H Hiller

55 Cusps of Gauss mappings
T Banchoff, T Gaffney and C McCrory

56 An approach to algebraic K-theory
A J Berrick

57 Convex analysis and optimization
J-P Aubin and R B Vinter

58 Convex analysis with applications in the differentiation of convex functions
J R Giles

59 Weak and variational methods for moving boundary problems
C M Elliott and J R Ockendon

60 Nonlinear partial differential equations and their applications: Collège de France Seminar. Volume II
H Brezis and J L Lions

61 Singular systems of differential equations II
S L Campbell

62 Rates of convergence in the central limit theorem
Peter Hall

63 Solution of differential equations by means of one-parameter groups
J M Hill

64 Hankel operators on Hilbert space
S C Power

65 Schrödinger-type operators with continuous spectra
M S P Eastham and H Kalf

66 Recent applications of generalized inverses
S L Campbell

67 Riesz and Fredholm theory in Banach algebra
B A Barnes, G J Murphy, M R F Smyth and T T West

68 Evolution equations and their applications
F Kappel and W Schappacher

69 Generalized solutions of Hamilton-Jacobi equations
P L Lions

70 Nonlinear partial differential equations and their applications: Collège de France Seminar. Volume III
H Brezis and J L Lions

71 Spectral theory and wave operators for the Schrödinger equation
A M Berthier

72 Approximation of Hilbert space operators I
D A Herrero

73 Vector valued Nevanlinna Theory
H J W Ziegler

74 Instability, nonexistence and weighted energy methods in fluid dynamics and related theories
B Straughan

75 Local bifurcation and symmetry
A Vanderbauwhede

76 Clifford analysis
F Brackx, R Delanghe and F Sommen

77 Nonlinear equivalence, reduction of PDEs to ODEs and fast convergent numerical methods
E E Rosinger

78 Free boundary problems, theory and applications. Volume I
A Fasano and M Primicerio

79 Free boundary problems, theory and applications. Volume II
A Fasano and M Primicerio

80 Symplectic geometry
A Crumeyrolle and J Grifone

81 An algorithmic analysis of a communication model with retransmission of flawed messages
D M Lucantoni

82 Geometric games and their applications
W H Ruckle

83 Additive groups of rings
S Feigelstock

84 Nonlinear partial differential equations and their applications: Collège de France Seminar. Volume IV
H Brezis and J L Lions

85 Multiplicative functionals on topological algebras
T Husain

86 Hamilton-Jacobi equations in Hilbert spaces
V Barbu and G Da Prato

87 Harmonic maps with symmetry, harmonic morphisms and deformations of metrics
P Baird

88 Similarity solutions of nonlinear partial differential equations
L Dresner

89 Contributions to nonlinear partial differential equations
C Bardos, A Damlamian, J I Díaz and J Hernández

90 Banach and Hilbert spaces of vector-valued functions
J Burbea and P Masani

91 Control and observation of neutral systems
D Salamon

92 Banach bundles, Banach modules and automorphisms of C*-algebras
M J Dupré and R M Gillette

93 Nonlinear partial differential equations and their applications: Collège de France Seminar. Volume V
H Brezis and J L Lions

94 Computer algebra in applied mathematics: an introduction to MACSYMA
R H Rand

95 Advances in nonlinear waves. Volume I
L Debnath

96 FC-groups
M J Tomkinson

97 Topics in relaxation and ellipsoidal methods
M Akgül

98 Analogue of the group algebra for topological semigroups
H Dzinotyiweyi

99 Stochastic functional differential equations
S E A Mohammed

100 Optimal control of variational inequalities
V Barbu

101 Partial differential equations and dynamical systems
W E Fitzgibbon III

102 Approximation of Hilbert space operators. Volume II
C Apostol, L A Fialkow, D A Herrero and D Voiculescu

103 Nondiscrete induction and iterative processes
V Ptak and F-A Potra

104 Analytic functions – growth aspects
O P Juneja and G P Kapoor

105 Theory of Tikhonov regularization for Fredholm equations of the first kind
C W Groetsch

106 Nonlinear partial differential equations and free boundaries. Volume I
J I Díaz

107 Tight and taut immersions of manifolds
T E Cecil and P J Ryan

108 A layering method for viscous, incompressible L_p flows occupying R^n
A Douglis and E B Fabes

109 Nonlinear partial differential equations and their applications: Collège de France Seminar. Volume VI
H Brezis and J L Lions

110 Finite generalized quadrangles
S E Payne and J A Thas

111 Advances in nonlinear waves. Volume II
L Debnath

112 Topics in several complex variables
E Ramírez de Arellano and D Sundararaman

113 Differential equations, flow invariance and applications
N H Pavel

114 Geometrical combinatorics
F C Holroyd and R J Wilson

115 Generators of strongly continuous semigroups
J A van Casteren

116 Growth of algebras and Gelfand–Kirillov dimension
G R Krause and T H Lenagan

117 Theory of bases and cones
P K Kamthan and M Gupta

118 Linear groups and permutations
A R Camina and E A Whelan

119 General Wiener–Hopf factorization methods
F-O Speck

120 Free boundary problems: applications and theory, Volume III
A Bossavit, A Damlamian and M Fremond

121 Free boundary problems: applications and theory, Volume IV
A Bossavit, A Damlamian and M Fremond

122 Nonlinear partial differential equations and their applications: Collège de France Seminar. Volume VII
H Brezis and J L Lions

123 Geometric methods in operator algebras
H Araki and E G Effros

124 Infinite dimensional analysis–stochastic processes
S Albeverio

125 Ennio de Giorgi Colloquium
P Krée

126 Almost-periodic functions in abstract spaces
S Zaidman

127 Nonlinear variational problems
A Marino, L Modica, S Spagnolo and M Degiovanni

128 Second-order systems of partial differential equations in the plane
L K Hua, W Lin and C-Q Wu

129 Asymptotics of high-order ordinary differential equations
R B Paris and A D Wood

130 Stochastic differential equations
R Wu

131 Differential geometry
L A Cordero

132 Nonlinear differential equations
J K Hale and P Martínez-Amores

133 Approximation theory and applications
S P Singh

134 Near-rings and their links with groups
J D P Meldrum

135 Estimating eigenvalues with *a posteriori/a priori* inequalities
J R Kuttler and V G Sigillito

136 Regular semigroups as extensions
F J Pastijn and M Petrich

137 Representations of rank one Lie groups
D H Collingwood

138 Fractional calculus
G F Roach and A C McBride

139 Hamilton's principle in continuum mechanics
A Bedford

140 Numerical analysis
D F Griffiths and G A Watson

141 Semigroups, theory and applications. Volume I
H Brezis, M G Crandall and F Kappel

142 Distribution theorems of L-functions
D Joyner

143 Recent developments in structured continua
D De Kee and P Kaloni

144 Functional analysis and two-point differential operators
J Locker

145 Numerical methods for partial differential equations
S I Hariharan and T H Moulden

146 Completely bounded maps and dilations
V I Paulsen

147 Harmonic analysis on the Heisenberg nilpotent Lie group
W Schempp

148 Contributions to modern calculus of variations
L Cesari

149 Nonlinear parabolic equations: qualitative properties of solutions
L Boccardo and A Tesei

150 From local times to global geometry, control and physics
K D Elworthy

151 A stochastic maximum principle for optimal control of diffusions
U G Haussmann

152 Semigroups, theory and applications. Volume II
H Brezis, M G Crandall and F Kappel

153 A general theory of integration in function spaces
P Muldowney

154 Oakland Conference on partial differential equations and applied mathematics
L R Bragg and J W Dettman

155 Contributions to nonlinear partial differential equations. Volume II
J I Díaz and P L Lions

156 Semigroups of linear operators: an introduction
A C McBride

157 Ordinary and partial differential equations
B D Sleeman and R J Jarvis

158 Hyperbolic equations
F Colombini and M K V Murthy

159 Linear topologies on a ring: an overview
J S Golan

160 Dynamical systems and bifurcation theory
M I Camacho, M J Pacifico and F Takens

161 Branched coverings and algebraic functions
M Namba

162 Perturbation bounds for matrix eigenvalues
R Bhatia

163 Defect minimization in operator equations: theory and applications
R Reemtsen

164 Multidimensional Brownian excursions and potential theory
K Burdzy

165 Viscosity solutions and optimal control
R J Elliott

166 Nonlinear partial differential equations and their applications. Collège de France Seminar. Volume VIII
H Brezis and J L Lions

167 Theory and applications of inverse problems
H Haario

168 Energy stability and convection
G P Galdi and B Straughan

169 Additive groups of rings. Volume II
S Feigelstock

170 Numerical analysis 1987
D F Griffiths and G A Watson

171 Surveys of some recent results in operator theory. Volume I
J B Conway and B B Morrel

172 Amenable Banach algebras
J-P Pier

173 Pseudo-orbits of contact forms
A Bahri

174 Poisson algebras and Poisson manifolds
K H Bhaskara and K Viswanath

175 Maximum principles and eigenvalue problems in partial differential equations
P W Schaefer

176 Mathematical analysis of nonlinear, dynamic processes
K U Grusa

177 Cordes' two-parameter spectral representation theory
D F McGhee and R H Picard

178 Equivariant K-theory for proper actions
N C Phillips

179 Elliptic operators, topology and asymptotic methods
J Roe

180 Nonlinear evolution equations
J K Engelbrecht, V E Fridman and E N Pelinovski

181 Nonlinear partial differential equations and their applications. Collège de France Seminar. Volume IX
H Brezis and J L Lions

182 Critical points at infinity in some variational problems
A Bahri

183 Recent developments in hyperbolic equations
L Cattabriga, F Colombini, M K V Murthy and S Spagnolo

184 Optimization and identification of systems governed by evolution equations on Banach space
N U Ahmed

185 Free boundary problems: theory and applications. Volume I
K H Hoffmann and J Sprekels

186 Free boundary problems: theory and applications. Volume II
K H Hoffmann and J Sprekels

187 An introduction to intersection homology theory
F Kirwan

188 Derivatives, nuclei and dimensions on the frame of torsion theories
J S Golan and H Simmons

189 Theory of reproducing kernels and its applications
S Saitoh

190 Volterra integrodifferential equations in Banach spaces and applications
G Da Prato and M Iannelli

191 Nest algebras
K R Davidson

192 Surveys of some recent results in operator theory. Volume II
J B Conway and B B Morrel

193 Nonlinear variational problems. Volume II
A Marino and M K Murthy

194 Stochastic processes with multidimensional parameter
M E Dozzi

195 Prestressed bodies
D Iesan

196 Hilbert space approach to some classical transforms
R H Picard

197 Stochastic calculus in application
J R Norris

198 Radical theory
B J Gardner

199 The C* – algebras of a class of solvable Lie groups
X Wang

200 Stochastic analysis, path integration and dynamics
D Elworthy

201 Riemannian geometry and holonomy groups
S Salamon

202 Strong asymptotics for extremal errors and polynomials associated with Erdös type weights
D S Lubinsky

203 Optimal control of diffusion processes
V S Borkar

204 Rings, modules and radicals
B J Gardner

205 Numerical studies for nonlinear Schrödinger equations
B M Herbst and J A C Weideman

206 Distributions and analytic functions
R D Carmichael and D Mitrović

207 Semicontinuity, relaxation and integral representation in the calculus of variations
G Buttazzo

208 Recent advances in nonlinear elliptic and parabolic problems
P Bénilan, M Chipot, L Evans and M Pierre

209 Model completions, ring representations and the topology of the Pierce sheaf
A Carson

210 Retarded dynamical systems
G Stepan

211 Function spaces, differential operators and nonlinear analysis
L Paivarinta

212 Analytic function theory of one complex variable
C C Yang, Y Komatu and K Niino

213 Elements of stability of visco-elastic fluids
J Dunwoody

214 Jordan decompositions of generalised vector measures
K D Schmidt

215 A mathematical analysis of bending of plates with transverse shear deformation
C Constanda

216 Ordinary and partial differential equations Vol II
B D Sleeman and R J Jarvis

217 Hilbert modules over function algebras
R G Douglas and V I Paulsen

218 Graph colourings
R Wilson and R Nelson

219 Hardy-type inequalities
A Kufner and B Opic

220 Nonlinear partial differential equations and their applications. College de France Seminar Volume X
H Brezis and J L Lions

221 Workshop on dynamical systems
E Shiels and Z Coelho

222 Geometry and analysis in nonlinear dynamics
H W Broer and F Takens

223 Fluid dynamical aspects of combustion theory
M Onofri and A Tesei

224 Approximation of Hilbert space operators. Volume I. 2nd edition
D Herrero

225 Operator Theory: Proceedings of the 1988 GPOTS–Wabash conference
J B Conway and B B Morrel

226 Local cohomology and localization
J L Bueso Montero, B Torrecillas Jover and A Verschoren

227 Sobolev spaces of holomorphic functions
F Beatrous and J Burbea

228 Numerical analysis. Volume III
D F Griffiths and G A Watson

229 Recent developments in structured continua. Volume III
D De Kee and P Kaloni

230 Boolean methods in interpolation and approximation
F J Delvos and W Schempp

231 Further advances in twistor theory, Volume 1
L J Mason and L P Hughston

232 Further advances in twistor theory, Volume 2
L J Mason and L P Hughston

233 Geometry in the neighborhood of invariant manifolds of maps and flows and linearization
U Kirchgraber and K Palmer

234 Quantales and their applications
K I Rosenthal

235 Integral equations and inverse problems
V Petkov and R Lazarov

236 Pseudo-differential operators
S R Simanca

237 A functional analytic approach to statistical experiments
I M Bomze

238 Quantum mechanics, algebras and distributions
D Dubin and M Hennings

239 Hamilton flows and evolution semigroups
J Gzyl

240 Topics in controlled Markov chains
V S Borkar

241 Invariant manifold theory for hydrodynamic transition
S Sritharan

242 Lectures on the spectrum of $L^2(\Gamma\backslash G)$
F L Williams

Inverse problems and imaging

G F Roach (Editor)

University of Strathclyde

Inverse problems and imaging

Longman Scientific & Technical

Copublished in the United States with
John Wiley & Sons, Inc., New York

Longman Scientific & Technical,
Longman Group UK Limited,
Longman House, Burnt Mill, Harlow,
Essex CM20 2JE, England
and Associated Companies throughout the world.

Copublished in the United States with
John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158

First published 1991

AMS Subject Classification: 35R30, 76Q05, 78A45

ISSN 0269-3674

British Library Cataloguing in Publication Data
Inverse problems and imaging.
1. Tomography. Applications of inversion methods
I. Roach, G.F. (Gary Francis)
621.3673

ISBN 0-582-06424-4

Library of Congress Cataloging-in-Publication Data
Inverse problems and imaging / G.F. Roach, editor.
p. cm.— (Pitman research notes in mathematics series. ISSN 0269-3674; 245)
1. Tomography—Mathematics—Congresses. 2. Inverse problems
(Differential equations)—Congresses. I. Roach, G.F. (Gary
Francis) II. Series.
RC78.7.T6I58 1991
616.07'57'0151535—dc20 90-49351
CIP

Printed and bound in Great Britain
by Biddles Ltd, Guildford and King's Lynn

Contents

Preface

As a result of informal discussions during the 1975 Methoden und Verfahren der mathematischen Physik meeting at Oberwolfach it was apparent that there was an outstanding need for an international seminar group which concerned itself with problems which had been attempted but which still stubbornly resisted solutions rather than with problems which had been solved.

The first such seminar was held at the University of Bonn in September 1976 under the heading of Mixed Boundary Value Problems. Since then a number of similar meetings on a variety of different topics have been held at either the University of Bonn or the University of Strathclyde. There is now every indication that similar meetings will be held regularly and, moreover, that they will become increasingly multidisciplinary. Consequently, it has been decided to make the work of these meetings more readily available by publishing proceedings. This volume comprises the proceedings of the most recent meeting held at Ross Priory, University of Strathclyde. We are most grateful to Longman Group Limited for agreeing to publish these proceedings.

These various meetings could not take place without a considerable amount of support and assistance from a number of sources. In this connection we are particularly grateful to

The Royal Society of Edinburgh

The Edinburgh Mathematical Society

The U.S. Army

for their financial support. Thanks are also due to all those colleagues who helped with the refereeing of the papers. Last but by no means least, we record our appreciation of the work done by the secretarial staff of the Department of Mathematics in the University of Strathclyde, especially Mrs. Mary Sergeant whose quiet efficiency helped in so many ways to ensure the smooth running of the conference.

University of Strathclyde
Glasgow G1 1XH

G. F. Roach

August 1989

List of contributors

W. R. Breckon, Department of Computing and Mathematical Sciences, Oxford Polytechnic, Oxford OX3 0BP, England, U.K.

S R. Cloude, University of Dundee, Department of Mathematics, Dundee DD1 4HN, Scotland, U.K.

now at University of York, Electronics Department, Heslington, York YO1 5DD, England, U.K.

G. Dassios, Division of Applied Mathematics, Department of Chemical Engineering, University of Patras, and Institute of Chemical Engineering, and High Temperature Chemical Processes, GR 261 10 Patras, Greece.

A. Faridani, Westfaelische Wilhelms-Universitaet, Institut fuer Numerische und instrumentelle Mathematik, Einsteinstr. 62, D-4400 Muenster, West Germany.

now at Department of Mathematics, Oregon State University, Corvallis, OR 97331, U.S.A.

K. Kiriaki, Department of Mathematics, National Technical University of Athens, GR 157 73 Athens, Greece.

V. Kostopoulos, Department of Applied Mechanics, University of Patras, GR 261 10 Patras, Greece.

R. F. Millar, Department of Mathematics, University of Alberta, Edmonton, Alberta, Canada T6G 2G1.

A. D. Parry, Department of Mathematics, University of Strathclyde, Glasgow G1 1XH, Scotland, U.K.

M. K. Pidcock, Applied Analysis Research Group, Department of Computing and Mathematical Sciences, Oxford Polytechnic, Headington, Oxford, England, U.K.

B. D. Sleeman, Department of Mathematical Sciences, University of Dundee, Dundee DD1 4HN, Scotland, U.K.

W. W. Symes, Department of Mathematical Sciences, Rice University, Houston, Texas 77005, U.S.A.

D. G. H. Tan, D.A.M.T.P., University of Cambridge, Silver Street, Cambridge CB3 9EW, England, U.K.

A. Zinn, Institut fur Numerische und Angewandte Mathematik, George-August-Universitat Gottingen, Lotzestrasse 16-18, D-3400 Gottingen, West Germany.

W.R. BRECKON

Measurement and reconstruction in electrical impedance tomography

INTRODUCTION

In this paper we consider the inverse problem of Electrical Impedance Tomography (EIT). In this medical imaging technique a pattern of current is applied to the surface $\partial\Omega$ of the body Ω and measurements of the resulting electric potential are made on $\partial\Omega$. A survey of clinical application of this technique can be found in Brown, Barber and Seagar [1] and comprehensive information on various aspects of the technique in the proceedings [2] [3].

Mathematically the problem can be formulated as follows. Let $\gamma \in L^\infty(\Omega)$ be the electrical conductivity, which satisfies $\gamma(x) > c > 0$ for (almost) all $x \in \Omega$. The potential u then satisfies

$$\nabla.\gamma\nabla u = 0$$

being a combination of Ohm's law and Kirchoff's current law. The current density j on $\partial\Omega$ is given by

$$j = -\gamma\nabla_n u$$

where n is the outward unit normal on $\partial\Omega$. Where convenient we will use the operator $L_\gamma = \nabla.\gamma\nabla$. To solve the equation $L_\gamma u = 0$ it is sufficient either to specify the Neumann condition j (together with an additional condition on the potential, such as $u(p) = 0$ for some p) or the Dirichlet conditions $u|\partial\Omega$. When one of the *sufficient* conditions is specified the other, or *complimentary*, boundary condition is determined also for a given γ.

Whilst Ω is clearly a domain in R^3 the case of $\Omega \subset R^2$ is also considered. The question of the possibility of identification of γ from a knowledge of all pairs $(j,u|\partial\Omega)$, that is the uniqueness of solution of the inverse problems,

has been considered by Kohn and Vogelius [4], who proved uniqueness for a pieccwise analytic conductivity and Sylvester and Uhlman [5] who proved the same for $\gamma \epsilon\ C^{\infty}(\Omega)$.

We will consider two aspects of the inverse problem here We present a suggestion for a system of optimal measurements and a linearisation approach to the solution of the inverse problem. In the latter we present an example to show that the linearised inverse problem is not the inversion of a Generalised Radon Transform as some authors have assumed.

Two-norm optimal measurement

The ideal case considered by Kohn and Vogelius and by Sylvester and Uhlman assume that we have a perfect set of measurements, that is all pairs $(j,u|\partial\Omega)$ are known. This is equivalent to knowing the transfer impedance operator R_γ where $R_\gamma j=u_{|\partial\Omega}$. If we consider the finite energy case $u \epsilon H^1(\Omega)$ then $R_\gamma: H^{-1/2}(\partial\Omega) \rightarrow H^{1/2}(\partial\Omega)$ is a pseudo differential operator which is compact and self adjdoint as a map $H^0(\partial\Omega) \rightarrow H^0(\partial\Omega)$. The mapping $\gamma \rightarrow R_\gamma$ is non-linear and it is this mapping we seek to invert.

In practice we can only apply a finite number of current patterns and measure voltage only at a finite number of points. The question arises: which are the best measurements to make?

Isaacson [6] gives one answer to this question which he frames in terms of a measure he calls distinguishability. If $\gamma 1$ and $\gamma 2$ are two conductivities then they are *distinguishable* by measurements of precision e if there is a current density $j \epsilon H^0(\partial\Omega)$ for which

$$\delta(j) = ||R_{\gamma 1} j - R_{\gamma 2} j|| / ||j|| > \epsilon$$

the number $\delta(j)$ is called the *distinguishability*. The best currents in the sense of Isaacson are those which maximise $\delta(j)$

$$\delta(j) = \sup_{||j||=1} ||R_{\gamma 1} j - R_{\gamma 2} j|| = \sup_{||j||=1} \langle j, D^2 j \rangle$$

where

$$D = |R_{\gamma 1} - R_{\gamma 2}|^{1/2}$$

The map D is a compact, self adjoint pseudo differential operator $H^0(\partial\Omega) \to H^0(\partial\Omega)$. It has a complete set of orthonormal eigen functions $j_k \in C^\infty(\partial\Omega)$, with eigen values $\lambda_1, \lambda_2, \ldots$, with $\lambda_k \to 0$ as $k \to \infty$. From the min-max principle one can deduce that the largest distinguishability possible is λ_1 which is achieved when j is an eigen function with this eigen value.

Isaacson's algorithm for calculating this optimal current is based on the power method (see for example [7]). We will take $\gamma 1$ to be the (unknown) conductivity of the body and $\gamma 2$ as the best available guess for the conductivity. The method is an iterative process which involves repeated measurement and can be expressed as follows:

Guess $j^{(0)}$ (where $||j^{(0)}|| = 1$)

Repeat

Apply $j^{(n)}$ and measure $v^{(n)} = R_{\gamma 1} j^{(n)}$;

Compute $v^{\sim(n)} = R_{\gamma 2} j^{(n)}$;

Set $\lambda_n = ||v^{(n)} - v^{\sim(n)}||$

Set $j^{(n+1)} = v^{(n)} - v^{\sim(n)})/\lambda_n$

Until $||j^{(n+1)} - j^{(n)}|| < \epsilon$

The reasoning behind making repeated measurements of v rather than taking a basis of currents and working with the resulting matrix for $R_{\gamma 1}$, is that the measurement process involves error and so is not truly linear. However the power method only produces the largest eigen function, which would give only one measurement with which to estimate $\gamma 1$. It would be desirable to have a

basis of currents at least spanning the same space as the eigen functionns j_k with $\lambda_k<\epsilon$. For this reason it is desirable to have a procedure which computes all the eigen functions jk with $\lambda_k>\epsilon$. The following procedure has been found to work in computer simulations with pseudo random noise in the measurement stage.

Guess $j1^{(0)}, j2^{(0)} \ldots, j_m^{(0)}$ (an orthonormal basis with $\int_{\partial\Omega} j1 = 0$)

Repeat

Measure $v_k^{(n)} = R_{\gamma 1} jk^{(n)}$, compute $v\tilde{}_k^{(n)} = R_{\gamma 2} jk^{(n)}$ for all k

Compute $r_{1k} = \langle jk, vk^{(n)} - v\tilde{}_k^{(n)} \rangle$

Calculate the eigen system for $R = \{r_{1k}]$, $RU = U\Lambda$

Set $j_k^{(n+1)} = S\ U_{k1} j_1$

Until $||R-\Lambda||<\epsilon$

In numerical experiments little was gained after two iterations.

Point optimal currents

Isaacson's criterion for best currents gives the best currents only in the sense of optimising the two norms of the voltage data measured for a given current pattern. In practical systems each measurement of voltage is made separately. It is of interest therefore to find a current pattern which optimises the voltage difference at a point on $\partial\Omega$ between bodies of conductivity $\gamma 1$ and $\gamma 2$.

Let $p\epsilon\ \partial\Omega$ be the point at which we make the measurment $v(p) - v\tilde{}(p) = R_{\gamma 1} j)(p) - (R_{\gamma 2} j)(p)$. We seek a j such that

$$\tau(j) = |v(p) - v\tilde{}(p)|$$

is optimised subject to $||j|| = 1$ and $\int_{\partial\Omega} j = 0$.

We can express the current pattern in terms of the eigen functions as $j = \Sigma_k j_k$. The optimisation problem is then

maximise $\Sigma a_k \lambda_k j_k(p)$

subject to $\Sigma a_k^2 = 1$.

Using Lagrangian procedure we obtain

$$a_k = C\lambda_k j_k(p)$$

where

$$C = 1/(\Sigma(\lambda_k j_k(p))^2).$$

The measurement procedure would then be first to calculate the eigen functions using the method of the previous section then to make a measurement at p apply the current

$$j = \Sigma\ C\lambda_k j_k(p) j_k$$

to give the voltage $v(p)-v\sim(p)$.

In the simple case $\Omega = \{x \in R^2 : |x| \leq 1\}$ with $\gamma 2 = 1$,

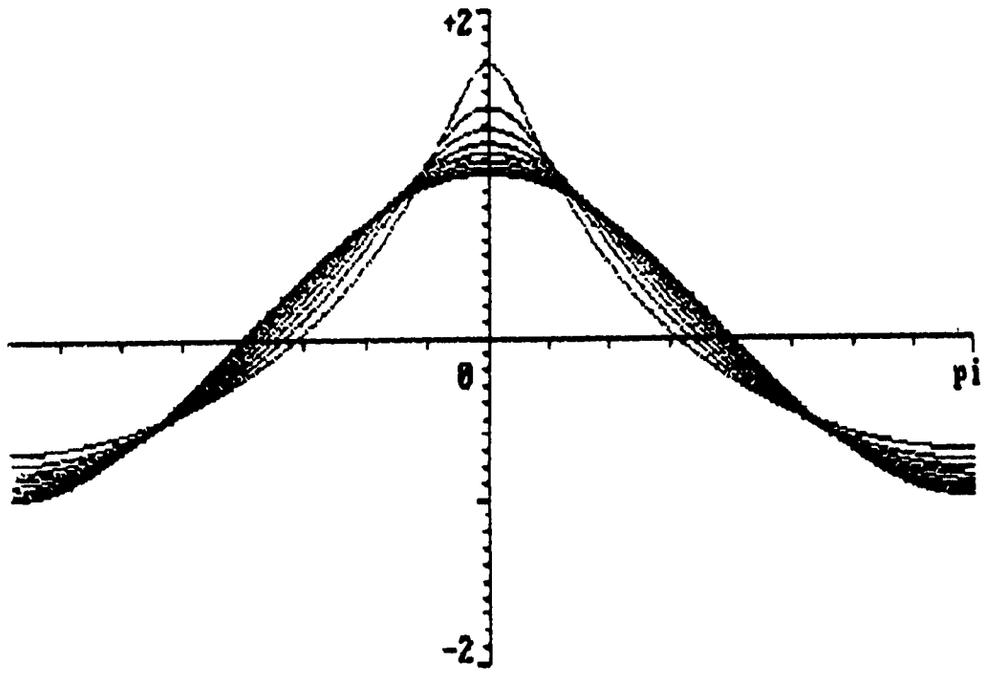
$$\gamma 1(x) = \begin{cases} \sigma & |x|<p \\ 1 & |x|>p \end{cases}$$

the eigen values can be explicitly calculated as

$$\lambda_k = -2\mu\rho^{2k}/k(1+\mu\rho^{2k})$$

where $\mu = (1-\sigma)/(1+\sigma)$ and the eigen functions are $\cos \kappa\theta$ and $\sin k\theta$.

A plot of the point optimal current for this case with p = 0 and various values of ρ is shown in Fig 1. In the case of small objects in the centre this procedure deviates least from that of Isaacson.



Point Optimal Curent sigma = 0.300 r = 0.1 to 0.9.

Figure 1

Linearisation

The mapping $\gamma \to R_\gamma$ is non-linear. To solve the inverse problem it is desirable to linearise this mapping. The original statement of the linearisation appears in Calderon [8]. In this chapter Calderon's techniques are used and elaborated upon to give linearised forms of the forward problem

both in its direct and integral form. Since the original reference is rather hard to find it is hoped that this will help readers who have not yet located a copy of this legendary Brazilian conference paper. Calderon's result is extended to give a 'Neumann conditions constant' formulation. All published reconstruction algorithms rely on some form of linearisation and yet the approximations used are not always justified. With this in mind the subject is treated in some detail here. We find, reassuringly, that they are all simply the Frechet derivative of appropriately defined forward mappings.

If the conductivity γ is perturbed to $\gamma+\delta\gamma$ and yet one form of sufficient boundary data for u is kept constant, the complimentary boundary data will change. For example, if a current density j is applied resulting in a potential u, that is $L_\gamma u = 0$, with $-\gamma\nabla_n u = j$ on $\partial\Omega$, then when γ is changed to $\gamma+\delta\gamma$, the potential u will change to $u+\delta u$, hence $L_{\gamma+\delta\gamma}(u+\delta u) = 0$ with $-(\gamma+\delta\gamma)\nabla_n(u+\delta u) = j$. The voltage difference on the boundary $\delta u|_{\partial\Omega}$ will be the data we measure in an attempt to detect this conductivity change so we want a formula for δu in terms of $\delta\gamma$ (the reverse would be too optimistic!), neglecting higher order terms in $\delta\gamma$. This is achieved by writing δu as a series in $\delta\gamma$ and truncating after the linear term. This series involves the linear operators $L_{\delta\gamma}$ which depends on $\delta\gamma$ in a linear way and G_γ which is the inverse of L_γ,(the G stands for Green's function of course). The non-specialist reader, if daunted by manipulations of differential and integral operators as though they were numbers, may like to think of them as matrices, as they would be if we passed to some discrete approximation to the operators.

CHOICE OF SPACE FOR γ

Standard elliptic theory, such as that presented in Gilbarg and Trudinger

[[9], requires the coefficients of the partial differential equations to be in $L^\infty(\Omega)$ which is a fairly weak assumption. To guarantee ellipticity of the operator L_γ we certainly need γ to be bounded away from zero, $\gamma > c > 0$ (almost everywhere). In the following results we will need to be able to evaluate $||\gamma|_{\partial\Omega}||$, that is we need to estimate the magnitude of the conductivity on the boundary. In $L^\infty(\Omega)$ there is no natural restriction mapping as $\partial\Omega$ is a set of measure zero. While $L^\infty(\Omega)$ contains extremely nasty functions it has an extremely strong convergence criterion. We would certainly be unwise to compare images on the basis of their L^∞ distance.

Natterer [10] suggests that the appropriate norms with which to compare two dimensional images is $H^{1/2}$. This space just fails to include the characteristic functions of domains with sufficiently regular boundaries. The weighting of high frequency terms (or if you like the inclusion of the 1/2th derivative) weights edges more strongly than the simple L^2 norm and this is consistent with the importance of edges in medical images. If we are to include continuous images in our space, then by the Sobolev Embedding Theorem we must use H^s for $s > 1$. This is also sufficient to guarantee $M^s \subseteq L^\infty$.

The requirement of this section, that of a bounded restriction mapping, corresponds to the existence of a natural trace operator in Sobolev spaces. In the space $H^s(\Omega)$ there is a natural trace operator $H^s(\Omega) \to H^{s-1/2}(\partial\Omega)$. If we require $\gamma|_{\partial\Omega}$ to be in $H^0(\partial\Omega)$, then that too would point to using $\gamma \in H^s(\Omega)$ for s greater than $\frac{1}{2}$.

DIRECT FORM

In this section the Frechet derivative of the potential as a function of conductivity is calculated. First the inverse operators are defined as follows:

$$G_\gamma^D: H^{-1}(\Omega) \to H_0^1(\Omega)$$

is the inverse of $L_\gamma: H_0^1(\Omega) \to H^{-1}(\Omega)$. Since the Dirichlet problem

$$L_\gamma u = s, \quad u|_{\partial\Omega} = 0$$

has a unique solution, G is well defined and the spaces have been chosen so that the inverse is bounded. Similarly

$$G_\gamma^N: H^{-1}(\Omega) \to H_N^1(\Omega)$$

solve

$$L_\gamma u = s, \quad \gamma\nabla_n u = 0, \quad \int_{\partial\Omega} u = 0$$

In addition we will need the mapping

$$G_\gamma^B: H^{-1/2}(\Omega) \to H^1(\Omega)$$

which solves the homogeneous Neumann problem

$$L_\gamma u = 0, \quad \gamma\nabla_n u = h, \int_{\partial\Omega} u = 0$$

In estimating the norm of L_γ which depends linearly on γ, the following Lemma will be used.

Lemma 1.1 $\downarrow ||L_\gamma||_{S,H^{-1}} \leq ||\gamma||$ where S is either H_0^1 or H_N^1

Proof

$$\begin{aligned} ||L_\gamma u||_{H^{-1}(\Omega)} &= \sup_{w \in S \subset H^1} \frac{|\int_\Omega w L_\gamma u|}{||w||_{H^1}} \\ &= \sup_{w \in S} \frac{|\int_{\partial\Omega} w\gamma\nabla u.n - \int_\Omega \gamma\nabla w.\nabla u|}{||w||_{H^1}} \\ &= \sup_{w \in S} (\int_\Omega \gamma\nabla w.\nabla u)/||w||_{H^1} \\ &\leq ||\gamma||\ ||u||_{H^1} \end{aligned}$$

We are now ready to prove the following theorem:

Theorem If $L_\gamma u = 0$ and $L_{\gamma+\delta\gamma}(u+\delta u) = 0$ and the Dirichlet or Neumann data for u and $u + \delta u$ agree then

$$L_{\delta\gamma} u + L_\gamma \delta u = o(||\delta\gamma||^2)$$

For the Neumann constant case we also have

$$\delta\gamma \nabla_n u + \gamma \nabla_{\underline{n}} \delta u = o(||\delta\gamma||^2).$$

The forward mapping defined by $F(\gamma) = u$ (which solves a given Neumann or Dirichlet problem) is C^∞. For the Dirichlet case

$$F'(\gamma)\delta\gamma = -G^D_\gamma L_{\delta\gamma} u$$

and the Neumann case

$$F'(\gamma)\delta\gamma = -G^N_\gamma L_{\delta\gamma} u - G^B_\gamma \delta\gamma \nabla_n u$$

Proof

In all cases we have

$$L_{\gamma+\delta\gamma}(u+\delta u) = L_{\delta}\gamma u + L_\gamma \delta u + L_{\delta\gamma}\delta u = 0$$

First consider the Dirichlet conditions constant case $\delta u|_{\partial\Omega} = 0$. Applying $G = G^D_\gamma$ gives

$$(1+G^D L_{\delta\gamma})\delta u = -G L_{\delta\gamma} u$$

Formally

$$\delta u = \sum_{j=1}^{\infty} (-G L_{\delta\gamma})^j u$$

which converges for $||GL_{\delta\gamma}|| < 1$. Using Lemma 1.1

$$||GL_{\delta\gamma}|| \leq ||G||.||\delta\gamma||$$

so convergence is achieved by requiring

$$\frac{||\delta\gamma||}{||G||} < 1$$

(the norm used for linear operators being the standard linear operator norm). This constitutes a Taylor series for F, as $L_{\delta\gamma}$ is linear in $\delta\gamma$ the j^{th} term is homogeneous of order j in $\delta\gamma$. In particular we have

$$F'(\gamma)\delta\gamma = -G^D_\gamma L_{\delta\gamma} u$$

and

$$L_{\delta\gamma} u + L_\gamma \delta u = o(||\delta\gamma||^2)$$

as claimed. It can readily be seen that F is C^∞ as the higher derivatives can be extracted from the Taylor series.

In the Neumann constant case we will treat the two components of $H^s(\Omega) = H^s_0(\Omega) \oplus H^{s-1/2}(\partial\Omega)$ separately writing $\gamma = \gamma_0 + \gamma_\partial$. The two partial derivatives can be calculated separately. In addition to the equation

$$L_\delta \gamma u + L_\gamma \delta u + L_{\delta\gamma} \delta u = 0$$

we have its boundary equivalent

$$(\gamma+\delta\gamma)\nabla_n(u+\delta u) = \gamma\nabla_n u.$$

To calculate $\partial F/\partial\gamma_0$ we assume $\delta\gamma|_{\partial\Omega} = 0$.

The boundary condition now reduces to $\gamma\nabla_n \delta u = 0$ and the proof proceeds as before with $G = G^N_\gamma$ and we have

$$\delta u = \sum_{j=1}^{\infty} (-GL_{\delta\gamma})^j u$$

and thus

$$\frac{\partial F}{\partial \gamma_0}\,\delta\gamma_0 = -G^N_\gamma L_{\delta\gamma 0} u.$$

On the other hand $\partial F/\partial\gamma_\partial$ can be calculated by assuming that $\gamma = 0$ in Ω^0. This leads to

$$L_\gamma \delta u = 0$$

and

$$\gamma\nabla_n \delta u = -\delta_{\gamma\partial}\nabla_n u - \delta_{\gamma\partial}\nabla_n \delta u.$$

The proof proceeds in a similar way to the interior case. Applying G^B_γ

$$\delta u = G^B_\gamma(-\delta\gamma_\partial \nabla_n u - \delta\gamma\partial\, \nabla_n \delta u)$$

rearranging

$$(1+G^B_\gamma \delta\gamma_\partial \nabla_n)\delta u = -G^B_\gamma \delta\gamma_\partial \nabla_n u.$$

The series

$$\delta u = \sum_{j=1}^{\infty} (-G^B_\gamma \delta\gamma_\partial \nabla_n)^j u$$

converges for $||G^B_\gamma \delta\gamma_\partial \nabla_n|| < 1$ which is ensured by $||\delta\gamma_\partial|| < 1/(||G^B_\gamma||.||\nabla_n||)$. Thus we have the derivative

$$\frac{\partial F}{\partial \gamma_\partial}\delta\gamma_\partial = -C_\gamma^B \delta\gamma \nabla_n u$$

and the result

$$\delta\gamma_\partial \nabla_n u + \gamma \nabla_n \delta u = o(||\delta\gamma_\partial||^2).$$

From the chain rule

$$F'(\gamma)\delta\gamma = -C_\gamma^N L_{\delta\gamma 0} u - C_\gamma^B \delta\gamma_\partial \nabla_n u.$$

The function can be seen to be C^∞ as the partial derivatives exist of all order.

□

In the Dirichlet constant case F is equal to its Taylor expansion in a neighbourhood of γ and is therefore an analytic mapping (this was pointed out by Calderon using $\gamma = 1$). (In the Neumann conditions constant case the proof gives the slightly weaker result that F is an analytic in each component separately). This indicates that the forward mapping could hardly be better behaved; however it is the inverse of this mapping which is required for EIT reconstruction and as we shall see that is not nearly as nice.

INTERPRETATION AS A SOURCE

The essential result of section 3.4 is that to first order we can assume

$$\nabla.\gamma\nabla\delta u = -\nabla.\delta\gamma\nabla u.$$

One interpretation of this is that the perturbed field δu is 'caused' by a distribution of current sources $-\nabla.\delta\gamma\nabla u$. Equivalently one could say that adding a source field $s = \nabla.\delta\gamma\nabla u$ would cancel the effect of changing the conductivity by $\delta\gamma$.

An interesting case to consider is to take $\delta\gamma = \delta_p$ the dirac delta distribution at a point p. We will assume that $u \in C^1(\Omega)$ and that p is not a critical point, $\nabla u(p) \neq 0$. For simplicity we will take $\gamma = 1$. The source

term is now $-\nabla\delta_p.\nabla u$ which is a dipole with dipole moment $|\nabla u(p)|$ oriented in the direction of the current vector at p. The perturbation δu to first order is then the electric field associated with the dipolee. The function $\delta u|_{\partial\Omega}$ is the point response of the system (up to first order), in optics this would be called the point spread function. In contrast to ideal optical systems the response is position dependent falling off dramatically as p gets further from the boundary. The field from a dipole is asymptotically $(2\mu \cos \theta)/r^2$. Here μ is the dipole moment, r the distance from the dipole and θ the angle relative to the dipole orientation. In this case $\mu = |\nabla u|$ which is at best constant and typically decreases away from $\partial\Omega$. Hence we find $||\delta u|_{\partial\Omega}||_{H^{1/2}(\partial\Omega} = o(1/p^2)$ where $p = \text{dist}(p,\partial\Omega)$. This means that the ability of an EIT system to detect a small object of high conductivity contrast will fall off at best proportionately to the inverse square of distance from the boundary.

INTEGRAL FORM

It is useful to reformulate the linearised problem in an integral form. In Chapter 4 the finite element method will be used to represent the electric potential and to solve the forward problem. In the finite element method differential equations are formulated as variational problems, this is equivalent to the weak form of the differential equation. Since this is essentially an integral formulation it will be advantageous to express the linearised conductivity-to-voltage mappin as an integral operator.

First notice that for any U with $L_\gamma U = 0$ and any V (U and V in $H^1(\Omega)$ say)

$$\nabla.(U\gamma\nabla V) = \gamma\nabla U,\nabla V$$

and so using the divergence theorem

$$\int_{\partial\Omega} V\gamma\nabla_n U = \int_\Omega \gamma\nabla V.\nabla U.$$

From above

$$L_{\delta\gamma}u + L_\gamma \delta u = o(||\delta\gamma||^2)$$

where $L_\gamma u = L_{(\gamma+\delta\gamma)}(u+\delta u) = 0$ and $\gamma\nabla_n u = (\gamma+\delta\gamma)\nabla_n u$. Choose any v with $L_\gamma w = 0$ then

$$\nabla.w(\gamma\nabla\delta u+\delta\gamma\nabla u) = \gamma\nabla w.\nabla\delta u+\delta\gamma\nabla w.\nabla u+o(||\delta\gamma||^2).$$

Applying the divergence theorem

$$\int_{\partial\Omega} w(\gamma\nabla_n\delta u+\delta\gamma\nabla_n u) = \int_\Omega \gamma\nabla w.\nabla\delta u+\delta C\nabla w.\nabla u+o(||\delta\gamma||^2)$$

Since $L_\gamma w = 0$ we know that

$$\int_{\partial\Omega} \delta u\gamma\nabla_n w = \int_\Omega \gamma\nabla\delta u\nabla w$$

also, from the boundary conditions $\gamma\nabla\delta u\nabla_n + \delta\gamma\nabla u\nabla_n = o(||\delta\gamma||^2)$, the result is

$$\int_{\partial\Omega} \delta u\gamma\nabla_n w = -\int_\Omega \delta\gamma\nabla w.\nabla u + o(||\delta\gamma||^2).$$

Let us now consider the implications of this formula for the reconstruction process. One has some initial estimate of the conductivity γ and wishes to correct this using the best linear approximation. Some known current patterns j_i are applied to the surface of the body $\partial\Omega$. Measurements of voltage u_i are made between various electrodes. Since measurement is an averaging process over the electrode we will assume that the measurements are of the form

$$V_{i,k} = \int_{\partial\Omega} u, w_k$$

where w_k is characteristic of the geometry and electrical properties of the electrode pair k. We have an *a priori* model of the body with conductivity γ which we compare with the real body which has conductivity $\gamma + \delta\gamma$. The discrepancy between the two is measured by the data

$$\delta V_{i,k} = \int_{\partial\Omega} \delta u w_k$$

We then solve the Neumann problem $L_\gamma w = 0$ subject to $\nabla_n w_k = w_k$. To find a linear approximation to the conductivity error $\delta\gamma$ we solve the system of linear equations

$$\delta V_{i,k} = -\int_\Omega \delta\gamma \nabla w_k . \nabla u_i$$

In this formulation the Neumann conditions were kept constant. This is the most useful formulation for impedance measurement for both theoretic and practical reasons. For completeness we must compare this to the problem investigated by Calderon in which the Dirichlet conditions were fixed and a difference of Neumann conditions (that is boundary current densities) measured. In this case the boundary conditions are $j = -\gamma\nabla_n u$ and

$$\delta j = -\delta(\gamma\nabla_n u) = \delta\gamma\nabla_n u + \gamma\nabla_n \delta u$$

assuming now that $\delta u|_{\partial\Omega} = 0$. This leads to the result

$$\int_{\partial\Omega} w\delta(\gamma\nabla_n w) = \int_\Omega \delta\gamma \nabla w . \nabla u + o(||\delta\gamma||^2)$$

which does not have the curious minus sign

Not a radon transform

It has been assumed by a number of authors (Schaumberg [11], Barber and Brown [12], Vogelius and Santoza [13] that the mapping $F:\delta\gamma \to \delta u|_{\partial\Omega}$ can be approximated by a Generalised Radon Transform (GRT), in the sense of Quinto [14]. Schaumbert assumed that the fibres were the current streamlines whereas Barber and Brown and Vogelius and Santoza assumed them to be the equipotential lines (in the case $\Omega \subseteq \mathbb{R}^2$). Such GRTs are linear maps so it is reasonable to compare them with the Frechet derivative $F(\gamma)$ which is in a specific sense the best linear approximation.

Consider as before the unit disc. Take $J = \cos\theta$, $\gamma = 1$. The solution of the Neumann problem for Laplace's equation is clearly $u(r,\theta) = r\cos\theta$ or

equivalent $u(x,y) = x$. If the perturbation $\delta\gamma$ is circularly symmetric then $\cos\theta$ will still be an eigen funtion of $R_1 + \delta\gamma$. For simplicity take

$$\gamma(\rho,x) = \begin{cases} 1 & |x|>\rho \\ \sigma & |x|<\rho \end{cases}$$

then $R_{\gamma(\rho,x)}\cos\theta = \lambda_1\cos\theta$ and $\delta u = \delta_1'(0)\cos\theta + o(|\delta\rho|^2)$. Thus δu has support on $\partial\Omega - \{\pi/2.3\pi/2\}$. If $F'(1)$ were a GRT then the support of δu would be contained in the 'shadow' of the set $S = \{x:|x|<\rho\}$, that is all the $\theta \in \partial\Omega$ such that $\mathcal{F}_\theta \cap S \neq \emptyset$ where $\mathcal{F}_\theta$ is the fibre through θ. It is clear that this is not the case in this example either for the current streamlines $\mathcal{F}_\theta = \{(x,y):y=\cos\theta\}$ nor for equipotential lines $\mathcal{F}_\theta = \{(x,y):x=\cos\theta\}$.

Reconstruction algorithms

Some of the reconstruction algorithms suggested by this work have been implemented and numerical results are reported in Breckon and Pidcock [15]. We will only give brief details here. The general procedure is as follows.

Make an initial guess $\gamma^{(0)}$ to γ

Repeat

Choose current patterns $J_1^{(n)}, j_2^{(n)}, \ldots, j_m^{(n)}$, possible using an optimal procedure and make measurements δV_{ik}

Find a least squares solution to

$$\delta V_{ik} = -\int_\Omega \gamma \nabla w_k . \nabla u_i \text{ for all } k,i$$

Update $\gamma^{(n+1)} = \gamma^{(n)} + \delta\gamma$

Until $\delta V_{ik} \leq \epsilon$

The linearised inverse problem is extremely ill posed as can be seen from the singular values given in [16]. Consequently regularisation must be used in solving the linear system. If only one iteration were used, the small number of useful data values given by the optimal methods (i.e. only those

which can distinguish $\gamma^{(n)}$ from γ) would be a severe problem. However in both of the adaptive methods discussed new currents are found at each iteration to distinguish best between what is really there and the latest guess. The map $R_\gamma^{(n)} = R_\gamma$ will be completely different at each iteration giving rise to completely different current patterns and voltage data. While the data set measured at each iteration is typically smaller than that used by non-adaptive methods, the information content of the data is maximised at each step.

REFERENCES

[1] B. H. Brown, D. C. Barber, and A. D. Seagar, 'Applied Potential Tomography: Possible clinical applications', Clin. Phys. Physiol. Meas. 109-121 1985.

[2] Proc EEC Workshop on Electrical Impedance Imaging/APT, Sheffield, Clin. Phys. Physiol. Meas. 8 Suppl A 1987.

[3] Proc 2nd EEC workshop on Electrical Impedance Tomography, Lyon, Clin. Phys. Physiol. Meas. 9 Suppl A 1988.

[4] R. V. Kohn and M. Vogelius, 'Determining the conductivity by boundary measurement II, interior results'. Comm. Pure Appl. Math. 38 643-667 1985.

[5] J. Sylvester and G. Uhlman, 'A Global Uniqueness Theorem for an Inverse Boundary Value Problem', Ann. Math. 125 643-667 1987.

[6] D. Isaacson, 'Distinguishability of conductivities by electric current computer tomography', IEEE Trans. on Medical Imaging, MI-5, No 2 p 91 1986.

[7] E. L. Isaacson and H. B. Keller, 'Analysis of numerical methods', Wiley 1966.

[8] A. P. Calderon, 'On an Inverse Boundary Value Problem', Seminar on Numerical Analysis and its Applications, W. H. Meyet and M. A. Ranpp, Eds. Rio de Janerio, Brazil: Brazilian Math. Soc. 1980.

[9] D. Gilbarg and N. S. Trudinger, 'Elliptic partial differential equations of second order', Springer, New York, 1970.

[10]F. Natterer, 'A Sobolev space analysis of picture reconstruction', SIAM J. Appl. Math. vol 39 402-411 1980.

[11]H. Schomberg 'Nonlinear image reconstructions from projections of ultra sonic travel times and electric current densities', in Mathemaical Aspects of Computed Tomography, Proceedings G. T. Herman and F. Natterer, eds. Berlin, Springer Verlag, 270-291, 1981.

[12]D. C. Barber, B. H. Brown, 'Applied Potential Tomography', I Phys. E. Instrum., 17 723-733 1984.

[[13]M. Vogelius and F. Santoza, 'A backprojection algorithm for electrical impedance imaging', Technical Note BN-1081, Inst. for Physical Science and Technology, Univ. of Maryland, 1988.

[14]E. T. Quinto, 'The dependence of the generalised Radon Transform on defining measure', Trans. Amer. Math. Soc., 257 No 2 1980 231-246.

[15]W. R. Breckon and M. K. Pidcock 'Data errors and reconstruction algorithms in electrical impedance tomography', Clin. Phys. Physiol. Meas. vol 9 Suppl A, 105-109 1988.

[16]W. R. Breckon and M. K. Pidcock 'Some mathematical aspects of electrical impedance tomography', in Mathematics and Computing in Medical Imaging, M. A. Viergever and A. E. Todd-Pokropek, Springer, Berlin, 351-362 1987.

W. R. Breckon
Department of Computing and Mathematical Sciences
Oxford Polytechnic
Oxford
England, U.K.

S.R. CLOUDE

Polarisation in electromagnetic inverse problems

ABSTRACT

With recent advances in measurement technology, full vector information is now available from electromagnetic scattering experiments. This paper presents a review of techniques developed for the inclusion of this vector information in inverse scattering theory.

The magnetic field integral equation is used to develop three inverse scattering models and to demonstrate the importance of full polarisation information for 3D reconstruction. We also consider two other inversion techniques, vector diffraction tomography and inverse boundary conditions, and discuss the polarisation aspects of each.

1. INTRODUCTION

Very often the vector nature of electromagnetic waves is ignored in direct scattering problems. Solutions are then formed for a complex scalar wavefield satisfying the scalar Helmholtz equation. This approximation is justified only when 2-dimensional problems are considered (in which case we can treat TE and TM waves separately) or when paraxial solutions are adequate and crosspolarisation is not of interest. A further advantage of adopting a scalar approach is the unity it provides with acoustic scattering theory (see Jones 1986).

For 3-D inverse problems however, this scalar approximation is neither justified nor desirable. Measurement techniques have been developed (mainly as a result of interest in frequency re-use in communications) to accurately

measure vector field quantities (see the review by Cox 1981). Here, we address the problem of how such information can be used to help in the inversion process. This provides a great challenge to inverse scattering theory; a complete inversion of the vector problem still evades us (see Langenberg 1989) but many of the concepts required have been identified. They involve advanced ideas from vector scattering theory, as well as such unlikely elements as Lie algebra and group theory. While the subject is still in the early stages of development, we present here a selective review of techniques of use in the vector inverse problem. For the sake of brevity, we concentrate only on those features which relate directly to polarisation of the scattered field, leaving other details to the references. This may be considered an update of a similar review published in 1981 by Boerner.

With this in mind we develop the paper in three main stages: first we review the matrix algebra used to describe polarisation effects and consider general symmetry properties of the scattered field which follow from generic properties of the scatterer. These are important because they may be used to impose global or local symmetry on the reconstruction. We then consider a set of inversion techniques based on the magnetic field integral equation (MFIE), with the high frequency physical optics and extended physical optics theories as special limiting cases. In particular, we review the important Kennaugh-Cosgriff inversion formula and show how we may add a polarisation correction term (first derived by Bennet) to yield information on specular point curvature. We then describe an exact inversion method based on the MFIE and discuss polarisation aspects of two other inversion schemes; vector diffraction tomography using the dyadic Greens function (Langenberg 1989) and the inverse boundary condition method as developed by Imbriale and Mittra (1970).

2 MATRIX FORMALISM

Under a far field assumption, we may write the vector scattered field as a multiple convolution between the incident field and object impulse response as

$$e_M = S_{MN} * e_N \quad 0 \leq M,N \leq 1 \tag{1}$$

By a Fourier transform we obtain a 2×2 complex scattering matrix [S] at angular frequency ω as

$$\underset{\sim}{E}_S = [S].\underset{\sim}{E}_I \tag{2}$$

where $\underset{\sim}{E}_S$ and $\underset{\sim}{E}_I$ are spinor quantities representing the scattered and incident electric fields respectively. The spinor nature of the electric field follows from its transformation properties under a change of (complex) orthogonal base states (see Cloude 1986). We may write such a transformation as

$$\underset{\sim}{E}' = [U_2].\underset{\sim}{E} \tag{3}$$

where $[U_2]$ is a 2×2 unitary matrix with unit determinant. We may interpret this geometrically as a rotation in a real three-dimensional space by writing [U2] as

$$\begin{aligned}[U_2] &= \cos\,\theta\,\sigma_0 - i\,\sin\,\theta\,(\sigma_X\cos\,\alpha + \sigma_Y\cos\,\beta + \sigma_Z\cos\,\gamma) \\ &= \exp(-i\theta[\underline{\sigma}.\underline{n}])\end{aligned} \tag{4}$$

representing a rotation of 2θ about an axis specified by angles α, β, γ. The set $\underline{\sigma} = (\sigma_0,\sigma_X,\sigma_Y,\sigma_Z)$ are the Pauli matrices. From spinor algebra we know that we may associate with (3) a spin matrix or quaternion

$$[X] = x\sigma_X + y\sigma_Y + z\sigma_Z \tag{5}$$

where the real 3 vector r = (x,y,z) maintains unit modulus under a transformation

$$[X] = [U_2][X][U_2]^+ \tag{6}$$

or

$$r' = [O_3].r \tag{7}$$

where $[0_3]$ is a 3×3 real orthogonal matrix. The locus formed under unit modulus is termed the Poincare sphere (see Born and Wolf 1965), which sits not in physical space, but in an abstract polarisation space of the spinor E. Many measurement systems operate by measuring the vector r rather than the complex spinor E (Cloude 1989). The two are related by the spin matrix X as

$$x = E_X E_X^* - E_Y E_Y^*$$
$$y = \mathrm{Re}(E_X E_Y^*)$$
$$z = \mathrm{Im}(E_X E_Y^*)$$

If we further define $m = E_X E_X^* + E_Y E_Y^*$ then $g = (m,r)$ is called the Stokes vector of the wave spinor and $[X] = m\sigma_0 + r.\sigma$ is called the wave coherency matrix.

In scattering theory, we generalise the above by considering (2) as a spinor transformation

$$[S] = \det([S])[U][H]$$
$$= d.\exp[(-i\theta\underline{n}+\lambda\underline{p}).\underline{\sigma}] \qquad (8)$$

where [U] is 2×2 unitary and [H], 2×2 hermitian (we shall assume for the moment that det([S]) is nonzero, it becomes zero only for degenerate scattering systems such as a linear dipole or helix). This transformation corresponds to a combination of boost and rotation of the Stokes vector and has 8 degrees of freedom; 2 direction vectors $\underline{n}$ and $\underline{p}$ lying in the space of the vector r_1 and 2 angles θ and λ. The direction vectors $\underline{n}$ and $\underline{p}$ figure prominently in the theory of null polarisations as developed by Kennaugh (see Kennaugh 1952 and Boerner '81). He showed that for backscatter, [S] has two orthogonal eigenstates and two copolar null states (where the incident wave is scattered into an orthogonal state). Since these states display special symmetries of the scattering volume they have been proposed as important features in the inverse problem (Boerner 1981), but to date no successful use

has been made of these null states, mainly because of the complexity of their dynamics.

By using the Stokes vector g instead of the spinor E we may express (2) and (8) in the form

$$g_S = [M].g_I \tag{9}$$

where

$$[M] = \frac{1}{2} \mathrm{Tr}(\underline{\sigma}\, S\, \underline{\sigma}\, S^{+}) = A^{-1} S \otimes S^{+} A \tag{10}$$

[M] is called the Mueller or Stokes reflection matrix, Tr stands for the trace of the matrix and A is a 4×4 unitary matrix of the form

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -i \\ 0 & 0 & 1 & i \\ 1 & -1 & 0 & 0 \end{bmatrix}$$

The matrix [M] is very important in vector scattering theory; the key problem in inverse scattering is to relate the elements of [M] to shape and material properties of the scattering volume. We also note that the Stokes vector g may also account for partially polarised waves where $m^2 > r.r.$ Such states lie inside the Poincare sphere (with random polarisation at the origin). This extra degree of freedom in g means that [M] is more general than [S], it contains all the relative phase information of [S], together with information on correlations between various elementary scattering mechanisms. To see this we outline an alternative matrix formalism based on a 4×4 hermitian covariance or coherency matrix (see Cloude 1986, 1989). The covariance matrix [Σ] is related to m_{ij}, the elements of [M], by 16 Dirac matrices, formed from direct products of the Pauli matrices as

$$[\Sigma] = m_{ji}\, \sigma_i \otimes \sigma_j \tag{11}$$

The scattering coherency matrix [T] is based on complexification of the spin matrix X and is related to [M] by a set of 16 traceless hermitian matrices β

as

$$[T] = m_{ij}\,\beta_{4i+j} \tag{12}$$

where

$$\beta_{4i+j} = A^{+}_{\sigma j} \otimes \sigma_{iA} \tag{13}$$

The matrices $[\Sigma]$ and $[T]$ transform under a similarity transformation involving a 4×4 unitary matrix $[U_4]$, representing a change in complex matrix basis for the expansion of $[S]$, i.e. if $[S] = \underset{\sim}{k}.\sigma$ then

$$\underset{\sim}{k}' = [U_4].\underset{\sim}{k} \tag{14}$$

and

$$[T'] = [U_4][T][U_4]^{+} \tag{15}$$

where $[T] = \underset{\sim}{k}.\underset{\sim}{k}^{+}$. For example, we may determine the form of $[S]$ under backscatter when the object has N fold symmetry in a plane perpendicular to the line of sight. From reciprocity it follows that $[S]$ is symmetric, i.e. $S_{HV} = S_{VH}$, $[M]$ is symmetric and T and Σ are 3×3 hermitian. Under a plane rotation τ, $\underset{\sim}{k}$ becomes $\underset{\sim}{k}' = [U_4].\underset{\sim}{k}$ where $[U_4]$ is given by

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\tau & -\sin 2\tau & 0 \\ 0 & \sin 2\tau & \cos 2\tau & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Since $k_3 = 0$ for backscatter it follows that for objects with N fold symmetry (and $N \geq 3$) $[S]$ must have the form of a complex scalar, i.e.

$$[S] = k_0\,\sigma_0 \tag{16}$$

This is a simple example of what can be done by considering symmetry properties of matrix descriptors (see Van de Hulst 1981 for further discussion of symmetry constraints on $[M]$). We can generalise such arguments to arbitrary unitary transformations of k, when we need to involve Lie algebra to ascertain invariant features under elementary rotations. Further details may be found in the references (Cloude 1986, 1989). We now turn to scattering

theory in an attempt to relate these matrix observables to shape and material properties of the scattering volume.

3. ELECTROMAGNETIC INVERSE SCATTERING

In this section we consider the problem of reconstruction of a three-dimensional perfectly conducting body using the magnetic field integral equation, the wave equation with inverse boundary conditions and vector diffraction tomography. In particular, we show how full polarisation information is needed for accurate reconstruction when the object has unequal principal curvatures, and show how such curvature information is contained within the elements of the Mueller matrix [M].

We begin with Maxwells equations in the time domain,

$$\nabla \times \underset{\sim}{E}(r,t) = -\mu \, \delta \underset{\sim}{H}(r,t)/\delta t$$

$$\nabla \times \underset{\sim}{H}(r,t) = \epsilon \, \delta \underset{\sim}{E}(r,t)/\delta t + \underset{\sim}{J}(r,t)$$

$$\nabla . \underset{\sim}{H}(r,t) = 0$$

$$\nabla . \underset{\sim}{E}(r,t) = \rho(r,t)/\epsilon$$

where

$$\underset{\sim}{H}(r,t) = 1/\mu \, \nabla \times \underset{\sim}{A}(r,t)$$

and

$$\underset{\sim}{A}(r,t) = \mu/4\pi \int_V \underset{\sim}{J}(r',t-\tau)/R \, dV$$

By defining the total magnetic field as $\underset{\sim}{H}^T = \underset{\sim}{H}^i + \underset{\sim}{H}^s$ and imposing the boundary condition $\underset{\sim}{n} \times \underset{\sim}{H}^T = \underset{\sim}{J}s$, we obtain an integral equation (the MFIE) for surface current $\underset{\sim}{J}s$ as (see Poggio and Miller 1973)

$$\underset{\sim}{J}s(r,t) = 2\underset{\sim}{n} \times \underset{\sim}{H}^i + 1/2\pi \, \underset{\sim}{n} \times \int_S L(\underset{\sim}{J}s) \times \underset{\sim}{R} \, ds \qquad (17)$$

where L is a differential operator

$$L = \frac{1}{R^2 c} \frac{\delta}{\delta t} + \frac{1}{R^3}$$

Once the current is known, the scattered far field may be determined from

$$H_S(r,t) = \frac{1}{4\pi Rc} \int_S \frac{\delta \underset{\sim}{J}s}{\delta \tau} \times \underset{\sim}{r} \, ds \qquad (18)$$

We may write the MFIE (17) as the sum of three contributions:

$$\underset{\sim}{J}(r,t) = 2\underset{\sim}{n} \times \underset{\sim}{H}^i + \rho \underset{\sim}{J}(r,t) + \underset{\sim}{S} \qquad (19)$$

where the first term represents the direct influence of the incident field, the second represents a self patch contribution to the current and the third, $\underset{\sim}{S}$, is an integral over other patches with currents evaluated at earlier times. The factor ρ has been shown by Mieras and Bennet (1982) to be (Marx (1985) derived a more general expression for the self patch term which reduces to the Bennet/Mieras result for constant current across the patch)

$$\rho = \frac{1}{2} \sqrt{\frac{\Delta A}{4\pi}} \, (K_U - K_V) \qquad (20)$$

where K_U and K_V are principal curvatures and ΔA is the area of the patch.

We may now derive inverse scattering identities by considering various approximations to the surface current. In the simplest case we ignore the self patch term and the integral S and consider only the forcing term. If we further assume the current to be zero on the shadow side of the object we obtain the well-known physical optics assumption. The expression for the scattered far field (18) then becomes

$$\underset{\sim}{H}_S(r,t) = A \int \delta/\delta t \, (\underset{\sim}{n} \times \underset{\sim}{H}_i) \times \underset{\sim}{r} \, ds \qquad (21)$$

If we assume an incident impulse plane wave with E polarised in the x direction and travelling in the positive z direction then the vector triple product reduces to a simple area projection funtion which, when Fourier transformed and integrated by parts, yields the following expression for the impulse response (see Kennaugh and Moffat 1965)

$$rH_s(r,t) = \frac{1}{2\pi c} \frac{d^2 A(t_s)}{dt_s^2} \qquad (22)$$

where $A(t_s)$ is the projected area funtion of the target, i.e. the cross sectional area of the target in a plane perpendicular to the z axis, delineated by a plane moving with a speed of one half light velocity. For example, for axial incidence on a rotationally symmetric target $A = \pi\rho^2$, where ρ is the target contour function. This important result was first derived by Kennaugh and Moffat in 1965 and has been used by several authors as the basis for electromagnetic imaging (see Young 1976, Shubert 1977). Note that while the early time impulse response is well approximated by this technique, late time ringing due to damped current flow on the object is ignored (it is contained in the integral S which we chose to ignore). This technique has been extended to account for late time damping using GTD, moments of the impulse response and phenomenological damping terms (Kennaugh and Moffat 1965).

The technique known as ramp response imaging (Young 1976) is based on integrating the impulse response twice to yield a direct relationship between a measured waveform and the area function $A(t_s)$. This has been used to derive information on object cross-section as a function of distance along the line of sight as well as object length and total volume (from the integral of the ramp response). By using multiple looks, limiting contours have been found for object shapes such as cubes and concentric cylinders. Note that this time domain result is related to the well-known Bojarski identity by a Fourier-Radon transform (as shown by Boerner 1978).

For our puposes we note that under physical optics there is no depolarisation of the backscattered signal. However, this does not mean that

there is zero depolarisation in the high frequency limit of the MFIE, as we not only ignored the integral S but also the self patch term in deriving this result. Since the latter is proportional to the radius of the self patch, we might expect it to be of only second order importance (very often the self patch term is ignored in numerical calculations using the MFIE). However, this is not the case for general scattering bodies. When determining the scattering matrix for the object, the self patch term makes an important contribution to the phase difference between copolar terms, as well as generating finite cross polar terms. In fact, the self patch term is only zero for surfaces which have equal radii of curvature, such as a flat plate or sphere.

If we assume the current over the self patch to be of constant magnitude and equal to the physical optics current then we may write a better approximation to the surface current as (Mieras and Bennet 1982, Bennet 1978)

$$\underset{\sim}{J} = 2\underset{\sim}{n} \times \underset{\sim}{H}^{i} + \epsilon\,(K_u - K_V)(J_{POU}\underset{\sim}{u} - J_{POV}\underset{\sim}{V})/4 \qquad (23)$$

where we have assumed a circular patch of radius $\epsilon = ct - z$ and the unit vectors $\underset{\sim}{u}$ and $\underset{\sim}{v}$ are aligned with the principal coordinate axes of the surface at the point of interest. When we integrate this expression to obtain the scattered field, the ϵ factor yields a term proportional to the first instead of second derivative of the area function. The final result for the impulse response is of the form (see Foo 1984, Chaudhuri 1986, 1977)

$$r\,\underset{\sim}{H}_S(r,t) = \frac{1}{2\pi}\frac{\delta^2 A}{\delta t^2}\,\underset{\sim}{a}_H + \frac{K_U K_V}{4\pi}\frac{\delta A}{\delta t}\{(\underset{\sim}{a}_H\cdot\underset{\sim}{u})u - (\underset{\sim}{a}_H\cdot\underset{\sim}{v})\underset{\sim}{v}\} \qquad (24)$$

where we have assumed a horizontally polarised incident magnetic field with unit vector $\underset{\sim}{a}_H$. We can now determine the elements of the scattering matrix [S] as

$$S_{HH} = \frac{1}{2\pi}\frac{\delta^2 A}{\delta t^2} + \frac{K_U K_V}{4\pi}\frac{\delta A}{\delta t}\cos 2\alpha \qquad (25)$$

$$S_{HV} = \frac{K_U K_V}{4\pi}\frac{\delta A}{\delta t}\sin 2\alpha \qquad (26)$$

$$S_{VV} = \frac{1}{2\pi}\frac{\delta^2 A}{\delta t^2} - \frac{K_U K_V}{4\pi}\frac{\delta A}{\delta t}\cos 2\alpha \qquad (27)$$

where α is the angle between the wave coordinates and principal coordinates for the specular point. If we expand this matrix in terms of the Pauli matrices and then perform a Fourier transform with respect to k_o, the free space wavenumber, we obtain a complex vector $\underset{\sim}{k}$ of the form

$$\underset{\sim}{k} = (k_0^2 A(k_0)/2\pi,\ \Delta\ ik_0 A(k_0)\cos 2\alpha,\ \Delta\ ik_0 A(k_0)\sin 2\alpha,\ 0) \qquad (28)$$

where Δ is the difference in curvature at the specular point and $A(k_0)$ is the value of the Fourier transform of the area function at spatial wavenumber k_0. Note that Δ is real, so the second two elements are in phase quadrature with the first. This implies that S_{HH} and S_{VV} have equal magnitude. Their phase difference, however, is directly related to the factor Δ. This means that we can estimate curvature difference at the specular point by measurement of the phase difference between copolar terms (measured at a wavelength much shorter than body dimensions, see Chaudhuri 1986, Boerner 1981). Note we also have information on the size of the object through the factor $A(k_0)$ and on local orientation through the angle α.

We can avoid the need for coherent measurements (accurate phase measurements are very difficult to achieve in practice) by calculating the coherency and Mueller matrices. The three pieces of information $A(k_0)$, Δ and α are then encoded in the diagonal elements of [M] as

$$k_0^4 |A(k_0)|^2 = 4\pi^2 \mathrm{Tr}([M]) = 4\pi^2(m_{00}+m_{11}+m_{22}+m_{33}) \qquad (29)$$

$$\Delta^2 = 2k_0^2(m_{00} - m_{33})/Tr([M]) \qquad (30)$$

$$\tan 2\alpha \doteq \sqrt{\frac{m_{00}-m_{11}+m_{22}-m_{33}}{m_{00}+m_{11}-m_{22}-m_{33}}} \qquad (31)$$

The 12 off-diagonal terms of the Mueller matrix relate to correlations between curvature difference Δ, amplitude weighting $A(k_0)$ and orientation α and hence for single point scattering provide no extra target information. For complex objects however (e.g. rough surfaces) these additional terms are useful for assessing variation of surface topology.

The next logical stage is to include the whole MFIE in the determination of surface current. Unfortunately we cannot then obtain analytical solutions but must resort instead to numerical techniques. Nonetheless, Bennet (1981) has used the full time domain MFIE for object reconstruction using two numerical techniques, the first based on iteration using successive estimates of the surface current and measured ramp response and a second based on a straightforward extension of the time stepping procedure used in direct scattering implementations of the MFIE. However, examples are only worked for axial incidence on bodies of revolution where the area function may be parameterised in terms of a contour function ρ and depolarisation does not occur. Although extension to non-symmetrical bodies was attempted, this area still remains to be fully investigated.

In summary, we have seen that we can obtain three inverse schemes based on the MFIE, the first yielding the physical optics Kennaugh-Cosgriff formula which is a time domain version of the Bojarski identity. We then saw that by including the self patch term we obtained a polarisation dependent correction to the inversion formula whereby, in the high frequency limit, the phase difference between copolar terms is simply related to curvature difference and

depolarisation is caused by misalignment of wave and surface coordinates. Finally we saw how the full MFIE may be used for object reconstruction by using numerical techniques akin to the time stepping procedures well known from direct scattering applications of the MFIE.

3.1 Inverse boundary conditions and vector diffraction tomography

As an alternative to the MFIE formulation of electromagnetic inverse scattering, Imbriale and Mittra (1970) devised a technique based on an inverse boundary condition for the wave equation. Weston and Boerner (1969) showed that this condition (namely that the total tangential electric field must be zero at a surface of a perfect conductor) is sufficient to reconstruct the object.

In this technique, complex measured data is required over an enclosing sphere at one frequency. This data is then Fourier transformed to obtain weighting coefficients in a Fourier expansion in terms of the angular variable o. These coefficients are then used together with the appropriate Greens function (a Hankel function for the 2-D problemr considered) to search for a point at which the total field (i.e. incident plus scattered) is zero (or some minimum if limited aspect data are obtained). The technique effectively involves analytic continuation of the field up to the surface of the scatterer and can be modified for concave bodies and multiple scattering (see Mittra 1973).

Unfortunately, although the technique is perfectly general and can in principal be used for the three-dimensional case, only two-dimensional examples have been published. Since 2-D problems can always be decoupled into TE and TM waves, the significance of polarisation for the 3-D case has not been clearly developed. Note that this technique is a point by point

reconstruction and such tends to be computationally intensive. However, for many applications only a limited number of surface points may be required and, with the widespread availability of super computer power, this technique may become feasible for low to intermediate frequency applications (the Kennaugh-Cosgriff formula being more efficient for high frequency problems). Ahluwalia and Boerner (1973) considered a generalisation of this technique to lossy dielectric bodies and showed that, while uniqueness is no longer assured, useful reconstructions can still be made, given limited *a priori* target information.

The third technique of current interest for 3-D inverse scattering is that based on a vector extension of linearised diffraction tomography (see Langenberg 1989). In this technique, the vector Greens theorem is used to obtain an electromagnetic version of the Porter-Borjarski integral equation for a vector holographic field θ as

$$\theta = -2\omega\mu \iiint_{-\infty}^{\infty} Jc(R,\omega)G_1(R-R',\omega)d^3R \qquad (32)$$

Where G_1 is the imaginary part of the vector Greens Function given by

$$G = (I+\nabla\nabla)g \qquad (33)$$

where I is the identity dyadic and g the scalar free space Greens function. In order to invert the Porter-Bojarski formula we must perform two operations; an inversion of the convolution (as in the scalar case) plus the added problem of inversion of the dyadic G_1. The former can be handled by linearising, assuming physical optics, and integrating with respect to the frequency variable ω. The latter, however, is difficult because the imaginary part of the Greens dyadic has zero determinant (in far field scattering it effectively projects 3-D fields onto a transverse plane). The result is that only projections of the solution can be calculated, leading to two simultaneous

scalar inversion integrals for a dyadic $\nabla\gamma$, where $\gamma(R)$ is the object singular function which is related to the characteristic function considered in the traditional scalar Bojarski theory. The net result is that differential geometric manipulations are needed to obtain reconstructions using this vector theory (see Langenberg 1989). The relationship between diffraction tomography and the extended physical optics theory of Bennet has yet to be evaluated.

A SUMMARY AND CONCLUSIONS

In this paper we have reviewed techniques of use in the formulation of vector inverse scattering problems. We have seen that matrix algebra may be used to provide a formalism suitable for measurement and analysis of polarisation problems and, in particular, that transformation properties of the matrix descriptors may be used to establish symmetries of the scattering volume. We have also seen a clear relationship between the coherent scattering matrix and real Mueller matrix, with the latter providing information of correlation properties in complicated scattering scenarios.

The magnetic field integral equation may be used to establish three important inversion techniques by using three approximations for the surface current. The importance of polarisation is clearly demonstrated by considering the effect of the self patch term in the MFIE. It leads to a simple but important relationship between the copolar phase difference and surface curvature.

Other electromagnetic inverse techniques which show promise for 3D inversion are those based on inverse boundary conditions and the point by point reconstruction technique of Imbriale and Mittra and those using manipulations of dyadic Greens functions in a generalisation of the Porter

Bojarski integral equation.

REFERENCES

[1] H. P. Ahluwalia, W. M. Boerner, Application of a Set of Inverse Boundary Conditions to the Profile Characteristics Inversion of Conducting Circular Cylindrical Shapes, IEEE Trans. AP-21, Sept. 1973, 663-672.

[2] C. L. Bennet, Time Domain Inverse Scattering, IEEE Trans. AP-29, 1981, 213-219.

[3] W. M. Boerner, Polarisation Utilisation in Electromagnetic Inverse Scattering, Communications Laboratory Report 78-3, University of Illinois at Chicago Circle, October 1978.

[4] W. M. Boerner, M. B. El-Arani, C. Y. Chan, P. M. Mastoris, Polarisation Dependence in EM Inverse Problems, IEEE Trans. AP-29, 1981, 262-271.

[5] M. Born, E. Wolf, Principles of Optics, Pergamon Press 1965.

[6] S. K. Chaudhuri, W. M. Boerner, Polarisation Utilisation in Profile Inversion of Perfectly Conducting Prolate Spheroid, IEEE Trans. AP-25, July 1977, 505-511.

[7] S. K. Chaudhuri, Electromagnetic Imaging Based on Characteristic Properties of the Rayleigh Coefficient, Radio Science, Vol 19, 1984, 1367-1375.

[8] S. K. Chaudhuri, B. Y. Foo, W. M. Boerner, A Validation Analysis of Huynens Target Descriptor Interpretations of the Mueller Matrix Elements in Polarimetric Radar Returns Using Kennaughs Physical Optics Impulse Response Formulation, IEEE Trans. AP-34, Jan. 1986, 11-20.

[9] S. Cloude, Group Theory and Polarisation Algebra, OPTIK Vol 75 No 1 1986, 26-36.

[10] S. Cloude, Uniqueness of Target Decomposition Theorems in Radar Polarimetry" in Direct and Inverse Methods in Radar Polarimetry, eds. W. M. Boerner et al, D. Reidel Publ. Comp., Dordrecht 1989.

[11] S. Cloude, Measurement and Analysis in Radar Polarimetry, in Direct and Inverse Methods in Radar Polarimetry, eds. W.M. Boerner et al, D. Reidel Publ. Comp., Dordrecht 1989.

[12] D. C. Cox, Depolarisation of Radio Waves by Atmospheric Hydrometeors in Earth-Space Paths - a review, Radio Science Vol 16 Sept.-October 1981, 781-812.

[13] Y. Das, W. M. Boerner, On Radar Target Shape Estimation using Algorithms for Reconstructions from Projections, IEEE Trans. AP-26 March 1978, 274-279.

[14] B. Y. Foo, S. K. Chaudhuri, W. M. Boerner, A High Frequency Inverse Scattering Model to Recover the Specular Point Curvature from Polarimetric Scattering Matrix Data, IEEE Trans. AP-32 November 1984, 1174-1178.

[15] H. van de Hulst, Light Scattering by Small Particles, Dover 1981.

[16] W. A. Imbriale, R. Mittra, The Two-Dimensional Inverse Scattering Problem, IEEE Trans. AP-18 1970, 633-642.

[17] D. S. Jones, Acoustic and Electromagnetic Waves, Oxford Science Pub. 1986.

[18] E. M. Kennaugh, D. L. Moffat, Transient and Impulse Response Approximations, Proc. IEEE Vol 53 1965, 893-901.

[19] E. M. Kennaugh, Polarision Properties of Radar Reflections, MSc Thesis, Ohio State University 1952,

[20] K. J. Langennberg, T. Gurke, Towards a Unified Theory of Vector Diffraction Tomography, in Direct and Inverse Methods in Radar Polarimetry, eds. W. M. Boerner et al, D. Reidel Publ. Comp., Dordrecht 1989.

[21] R. M. Lewis, Physical Optics Inverse Diffraction, IEEE Trans. AP-17 May 1969, 308-314.

[22] E. Marx, Self Patch Integrals in Transient Electromagnetic Scattering, IEEE Trans. AP-33 July 1985, 763-767.

[23] H. Mieras. C. L. Bennet, Space Time Integral Equation Approach to Dielectric Targets, IEEE Trans. AP-30, 1982, 2-9.

[24] R. Mittra, Inverse Scattering and Remote Probing, Chapter 7 in Computer Techniques for Electromagnetics, ed. R. Mittra, Pergamon Press 1973.

[25] A. J. Poggio, E. K. Miller, Integral Equation Solutions of Three-Dimensional Scattering Problems, in Computer Techniques for Electromagnetics, ed. R. Mittra, Pergamon Press 1973.

[26] K. A. Shubert, J. D. Young, D. L. Moffat, Synthetic Radar Imaging, IEEE Trans. AP-25 July 1977, p477-483.

[27] V. H. Weston, W. M. Boerner, An Inverse Scattering Technique for Electromagnetic Bistatic Scattering, Can. J. Phys. Vol 47, 1969, 1177-1184.

[28] J. D. Young, Radar Imaging From Ramp Response Signatures, IEEE Trans. AP-24 1976, 277-287.

S. R. Cloude
University of Dundee
Department of Mathematics
Dundee DD1 4HN
Scotland, U.K.

G. DASSIOS AND K. KIRIAKI

Size orientation and thickness identification of an ellipsoidal shell

ABSTRACT

A confocal ellipsoidal shell of unknown semiaxes, orientation and thickness is excited into secondary radiation by a plane harmonic wave at low-frequency. The outer ellipsoid forms a penetrable surface while the inner one is a soft ellipsoid. We show that one measurement of the leading low-frequency approximation and six specific measurements of the next approximation for the scattering cross-section are enough to determine the size, the orientation and the shell thickness of the target.

1. INTRODUCTION

In a recent paper [4], the first author has developed a simple algorithm that provides the size and the orientation of an unknown soft ellipsoid out of seven measurements of the forward scattering amplitude. Actually, only six measurements are necessary but the seventh one simplifies the algorithm significantly. The method was based on the solution of the corresponding direct problem [2]. Furthermore, the problem of acoustic scattering by a soft ellipsoid coated by a penetrable confocal ellipsoidal shell has also been solved in the low-frequency realm in [3]. The present paper aims towards the development of an algorith that will provide the size, the orientation and the shell thickness of a soft scatterer which is coated by a penetrable shell with exterior boundary an ellipsoid possessing the same foci as the scatterer itself. It is shown that knowledge of the first two low-frequency approximations of the scattering cross-section for specific directions of

incidence are necessary to identify the scatterer completely. More accurately, we need to know the value of the leading term for any direction of excitation and the value of the second term for six specific directions of excitation. For scalar scattering by an ellipsoid one can look in [6,7,8]. Inverse scattering by an ellipsoid, from a completely different point of view, is developed in [1].

2. THE DIRECT PROBLEM

Consider the confocal ellipsoids

$$S_a: \frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \frac{x_3^2}{a_3^2} = 1 \tag{1}$$

and

$$S_b: \frac{x_1^2}{b_1^2} + \frac{x_2^2}{b_2^2} + \frac{x_3^2}{b_3^2} = 1 \tag{2}$$

where $a_1>a_2>a_3>0$, $b_1>b_2>b_3>0$,

$a_i>b_i$, $i = 1,2,3$ and the semi-interfocal distances h_1, h_2, h_3, where

$$h_1^2 = a_2^2 - a_3^2 = b_2^2 - b_3^2$$

$$h_2^2 = a_1^2 - a_3^2 = b_1^2 - b_3^2 \tag{3}$$

$$h_3^2 = a_1^2 - a_2^2 = b_1^2 - b_2^2$$

Suppressing the harmonic time dependence $e^{-i\omega t}$ and assuming the incident plane wave

$$\Phi(\underset{\sim}{r}) = e^{i\underset{\sim}{k}\cdot\underset{\sim}{r}}, \tag{4}$$

the direct scattering problem at hand asks for the evaluation of a function $\Psi^-(r)$ defined on the domain V^- between S_b and S_a and a function $\Psi^+(\underset{\sim}{r})$ defined

on the exterior to S_a domain V^+ which satisfy the following conditions. The total field

$$\Psi^+(\underset{\sim}{r}) = \Phi(\underset{\sim}{r}) + U(\underset{\sim}{r}), \tag{5}$$

where $U(r)$ is the scattered field, has to satisfy the Helmholtz equation with wave number k

$$(\Delta + k^2)\Psi^+(\underset{\sim}{r}) = 0, \quad \underset{\sim}{r} \in V^+ \tag{6}$$

while the inner field $\Psi^-(\underset{\sim}{r})$ has to satisfy the Helmholtz equation with a different wave number K

$$(\Delta + K^2)\Psi(\underset{\sim}{r}) = 0, \quad \underset{\sim}{r} \in V^-. \tag{7}$$

On the outer boundary S_a the boundary conditions

$$\Psi^+(\underset{\sim}{r}) = \Psi(\underset{\sim}{r}), \quad \underset{\sim}{r} \in S_a \tag{8}$$

$$\partial_n\Psi^+(\underset{\sim}{r}) = \bar{B}\partial_n(\Psi^-(\underset{\sim}{r}), \quad \underset{\sim}{r} \in S_a \tag{9}$$

that secure penetration are to be satisfied, while the inner boundary S_b consists of a soft boundary

$$\Psi^-(\underset{\sim}{r}) = 0, \quad \underset{\sim}{r} \in S_b \tag{10}$$

The positive constant

$$\bar{B} = \frac{\rho^+}{\rho^-} \tag{11}$$

determines the ratio of the outer to the inner mass density. The ratio of the bulk moduli of elasticity is given by

$$\frac{\gamma^+}{\gamma^-} = \bar{B}\frac{K^2}{k^2} = \bar{B}\eta^2 \tag{12}$$

where η stands for the relative index of refraction. For convenience we introduce the parameters

$$B = \bar{B} - 1 \tag{13}$$

$$C = \bar{B}\eta^2 - 1 \tag{14}$$

which imply

$$\eta^2 = \frac{C+1}{B+1}. \tag{15}$$

Finally, the scattered field $U(\underset{\sim}{r})$ should satisfy the Sommerfeld radiation condition at infinity.

The low-frequency analysis of this scattering problem can be found in [3]. In particular the leading two low-frequency coefficients of the scattering cross-section are provided by

$$\sigma = 4\pi\left[\frac{B+1}{H^o}\right]^2 + 4\pi k^2\left[T - \frac{1}{3}\left[\frac{B+1}{H^o}\right]^2 \hat{\underset{\sim}{k}}.\tilde{B}.\hat{\underset{\sim}{k}}\right] + O(k^4),\ k\to 0 \tag{16}$$

where

$$\tilde{B} = \sum_{n=1}^{3} b_n^2 \hat{\underset{\sim}{x}}_n \circledast \hat{\underset{\sim}{x}}_n \tag{17}$$

$$I_0^1(a_1) = \frac{1}{2}\int_0^{+\infty} \frac{dx}{\sqrt{x+a_1^2}\sqrt{x+a_2^2}\sqrt{x+a_3^2}} \tag{18}$$

$$I_0^1(b_1) = \frac{1}{2}\int_0^{+\infty} \frac{dx}{\sqrt{x+b_1^2}\sqrt{x+b_2^2}\sqrt{x+b_3^2}} \tag{19}$$

$$I_1^n(a_1) = \frac{1}{2}\int_0^{+\infty} \frac{dx}{(x+a_n^2)\sqrt{x+a_1^2}\sqrt{x+a_2^2}\sqrt{x+a_3^2}} \quad n = 1,2,3 \tag{20}$$

$$I_1^n(b_1) = \frac{1}{2}\int_0^{+\infty} \frac{dx}{(x+b_n^2)\sqrt{x+b_1^2}\sqrt{x+b_2^2}\sqrt{x+b_3^2}}, \quad n = 1,2,3 \tag{21}$$

$$I_0^1(\rho) = \frac{1}{2}\int_{\rho^2-a_1^2}^{+\infty} \frac{dx}{\sqrt{x+a_1^2}\,\sqrt{x+a_2^2}\,\sqrt{x+a_3^2}} \tag{22}$$

$$H^0 = B\, I_0^1(a_1) + I_0^1(b_1) \tag{23}$$

and

$$T = \frac{C(B+1)}{2\pi(H^0)^2}\int_V I_0^1(\rho)dv + \frac{B(B+1)^2}{(H^0)^3}\int_{b_1}^{a_1} \rho I_0^1(\rho)d\rho$$

$$- \frac{(a_1^2-b_1^2)(B+1)}{6(H^0)^3}\left[\left[-12(B+1)^2+3(C+1)(B+4)-B\,\frac{C+1}{B+1}\right]I_0^1(a_1)-(B+2)\frac{C+1}{B+1}\,I_0^1(b_1)\right]$$

$$- \left[\frac{B+1}{H^0}\right]^4 - \frac{2C(B+1)}{3(H^0)^2}\,(a_1a_2a_3-b_1b_2b_3)\,I_0^1(b_1) \tag{24}$$

$$- \frac{2C(B+1)^2a_1a_2a_3}{3(H^0)^3}\left[I_0^1(a_1)-I_0^1(b_1)\right]I_0^1(a_1)-\frac{a_1^2-b_1^2}{3(H^0)^2}\left[(2B+1)(B+1)-\frac{1}{2}\,B(C+1)\right]$$

$$+ \frac{2(b_1^2+b_2^2+b_3^2)(B+1)}{3(H^0)^3}\left[(B^2+2B-C)I_0^1(a_1)+(C+1)I_0^1(b_1)\right]-\left[\frac{B+1}{H^0}\right]^2\frac{b_1^2+b_2^2+b_3^2}{9}$$

$$- \frac{2(B+1)}{3(H^0)^3}\left[(C+1)\sum_{n=1}^{3} b_n^4 I_1^n(b_1)+(B^2+2B-C)\sum_{n=1}^{3} a_n^4 I_1^n(a_1)\right]$$

The expression for T involves only the two parameters B, C and the six semiaxes a_1, a_2, a_3, b_1, b_2, b_3. Furthermore, the dyadic $\tilde{B}$ corresponds to an ellipsoid which is reciprocal to (2).

3. THE INVERSE PROBLEM

The inverse problem we consider here asks for the determination of the semiaxes a_1, a_2, a_3, b_1, b_2, b_3 which specify the size and the thickness, and the three Euler angles ϕ, θ, y that provide the orientation of the scatterer. The form of (16) indicates that it is possible to use the method developed in [4] to solve the present problem as well. The only difference is due to the ellipsoidal shell that coats the soft scatterer, i.e. we need to determine b_1, b_2, b_3 and ϕ, θ, y as in [4], as well as the semiaxes a_1, a_2, a_3. Nevertheless, because of the confocality of S_a and S_b, relations (3) can be used to determine a_1, a_2, in terms of a_3 as

$$a_1^2 = b_1^2 + a_3^2 - b_3^2 \tag{25}$$

$$a_2^2 = b_2^2 + a_3^2 - b_3^2 \tag{26}$$

Consequently, we only need to determine b_1, b_2, b_3, a_3 and ϕ, θ, y.

An expression that related a_3 with b_1, b_2, b_3 can be obtained by a single measurement m_0 of the leading approximation of σ.

Indeed, if

$$m_0 = 4\pi \left[\frac{B+1}{H^0}\right]^2 \tag{27}$$

then, in view of (18), (19) and (23) we obtain

$$B \int_0^{+\infty} \frac{dx}{\sqrt{x+(b_1^2-b_3^2)+a_3^2}\sqrt{x+(b_2^2-b_3^2)+a_3^2}\sqrt{x+a_3^2}}$$

$$+ \int_0^{+\infty} \frac{dx}{\sqrt{x+b_1^2}\sqrt{x+b_2^2}\sqrt{x+b_3^2}} = 4(B+1)\sqrt{\frac{\pi}{m_0}} \tag{28}$$

which furnishes a_3 as a function of b_1, b_2, b_3

$$a_3 = a_3(b_1, b_2, b_3). \tag{29}$$

The existence of the function (29) is secured from the implicit function theorem with the help of the relation [2]

$$a_1 a_2 a_3 \left[I_1^1(a_1) + I_1^2(a_1) + I_1^3(a_1) \right] = 1 \tag{30}$$

which connects the corresponding elliptic integrals.

Following the procedure described in [4] we excite the scatterer from the six directions

$$\left.\begin{aligned} &\hat{k}_1', \ \hat{k}_2' \ \hat{k}_3' \\ &\hat{k}_4' = \frac{\hat{x}_1' + \hat{x}_2'}{\sqrt{2}} \\ &\hat{k}_5' = \frac{\hat{x}_2' + \hat{x}_3'}{\sqrt{2}} \\ &\hat{k}_6' = \frac{\hat{x}_3' + \hat{x}_1'}{\sqrt{2}} \end{aligned}\right\} \tag{31}$$

of an arbitrary coordinate system indicated with primed variables. If we denote by x_1, x_2, x_3 the coordinate system that fits the principal axes of the target ellipsoid then there exists an orthogonal dyadic $\tilde{P}$ such that

$$\underset{\sim}{r} = \tilde{P} . \underset{\sim}{r}' \tag{32}$$

where the components of $\tilde{P}$ are known trigonometric functions [4] of the three unknown Euler angles ϕ, θ, y.

Let

$$m_i = 4\pi\left[T - \frac{1}{3}\left[\frac{B+1}{H^o}\right]^2 \hat{\underset{\sim}{k}}_i . \tilde{B} . \hat{\underset{\sim}{k}}_i\right] \tag{33}$$

be the measured values of the second approximation of the scattering cross-section whenever the direction of excitation is given by $\hat{\underset{\sim}{k}}_i$, $i = 1,2,3,4,5,6$. A series of calculations similar to those in [4] leads to the dyadic equation

$$\tilde{B} - \frac{12\pi T}{m_o} I = \tilde{P} . \tilde{M} . \tilde{P}^* \tag{34}$$

where $\tilde{I}$ is the idemfactor, $\tilde{P}^*$ is the adjoint to $\tilde{P}$ dyadic and

$$\begin{aligned} M = \frac{3}{2m_o}\Big[& -2m_1 \hat{\underset{\sim}{x}}_1 \otimes \hat{\underset{\sim}{x}}_1 - 2m_2 \hat{\underset{\sim}{x}}_2 \otimes \hat{\underset{\sim}{x}}_2 - 2m_3 \hat{\underset{\sim}{x}}_3 \otimes \hat{\underset{\sim}{x}}_3 \\ & + (m_1+m_2-2m_4)(\hat{x}_1 \otimes \hat{x}_2 + \hat{x}_2 \otimes \hat{x}_1) \\ & + (m_2+m_3-2m_5)(\hat{\underset{\sim}{x}}_2 \otimes \hat{\underset{\sim}{x}}_3 + \hat{\underset{\sim}{x}}_3 \otimes \hat{\underset{\sim}{x}}_2) \\ & (m_3+m_1-2m_6)(\hat{\underset{\sim}{x}}_3 \otimes \hat{\underset{\sim}{x}}_1 + \hat{\underset{\sim}{x}}_1 \otimes \hat{\underset{\sim}{x}}_3)\Big] \end{aligned} \tag{35}$$

is the dyadic of measurements.

Since $\tilde{M}$ is a real symmetric dyadic its normal form involves three real eigenvalues and an orthogonal set of eigenvectors.

By virtue of (34), the eigenvectors of $\tilde{M}$ form the dyadic $\tilde{P}$ while its eigenvalues λ_1, λ_2, λ_3 are given by

$$\lambda_n = b_n^2 - \frac{12\pi T}{m_o}, \quad n = 1,2,3. \tag{36}$$

Hence the square of the three semiaxes of the core of the scatterer are provided by

$$b_n^2 = \lambda_n + \frac{12\pi T}{m_o}, \quad n = 1,2,3. \tag{37}$$

In view of (29) T is a function of b_1, b_2 and b_3. Therefore, (37) forms a highly non-linear system of three equations for the three unknowns b_1, b_2, b_3. Once a numerical scheme furnishes the value of b_1, b_2 and b_3, Formula (29)

provides a_3 and (25), (26) give the values of a_1 and a_2. This procedure identifies the size of the scatterer and the thickness of its core. The Euler angles that give the orientation as well are obtained from the knowledge of the dyadic $\tilde{P}$ exactly the same way as in [4].

We mention here that if the size and the orientation of the exterior ellipsoid are known, then a single measurement of the leading term of the scattering cross-section sufficies to identify the size of the core ellipsoid [5]. This procedure, which forms a nondistructive method of evaluating the size of the interior ellipsoid, is based on the Rayleigh approximation of a soft ellipsoid

$$\sigma' \sim 4\pi\left[\frac{1}{I_0^1(a_1)}\right]^2 \tag{38}$$

and the corresponding approximation

$$\sigma \sim 4\pi\left[\frac{B+1}{H^0}\right]^2 \tag{39}$$

given by (16).

Eliminating $I_0^{\,1}(a_1)$ between (38) and (39) we obtain the relation

$$I_0^1(b_1) = 2\sqrt{\pi}\left[\frac{B+1}{\sqrt{\sigma}} - \frac{B}{\sqrt{\sigma'}}\right], \tag{40}$$

which, for known σ and σ', can be solved numerically to obtain b_3 and then b_1 and b_2 from

$$b_1^2 = h_2^2 + b_3^2 \tag{41}$$

$$b_2^2 = h_1^2 + b_3^2 \tag{42}$$

as long as h_2 and h_3 are known from the knowledge of a_1, a_2, a_3.

ACKNOWLEDGEMENT:

The present work has been supported by the EEC SCIENCE program under grant number SCI-0079.

REFERENCES

[1] T. S. Angell and R. E. Kleinman, Polarizability Tensors in Low-Frequency Inverse Scattering, Radio Science, 22, 1120-1126 (1987).

[2] G. Dassios, Second Order Low-Frequency Scattering by the Soft Ellipsoid, SIAM J. Appl. Math. 38, 373-381 (1980).

[3] G. Dassios, Scattering of Acoustic Waves by a Coated Pressure-Release Ellipsoid, J. Acoust. Soc. Amer. 70, 176-185 (1981).

[4] G. Dassios, The Inverse Scattering Problem for the Soft Ellipsoid, J. Math. Phys. 28, 2858-2862 (1987).

[5] G. Dassios, S. A. Paipetis and A. Pournaras, An Inverse Scattering Problem for an Ellipsoidal Inclusion with an Interphase in an Acoustic Field, Proceeding of the Second International Symposium in Composite Materials, Patras (1988).

[6] B. D. Sleeman, The Scalar Scattering of a Plane Wave by an Ellipsoid, J. Inst. Math. Appl. 3, 4-15, (1967).

[7] B. D. Sleeman, The Low-Frequency Scalar Dirichlet Scattering by a General Ellipsoid, J. Inst. Math. Appl. 3, 291-312 (1967).

[8] W. E. Williams, Some Results for Low-Frequency Dirichlet Scattering by Arbitrary Obstacles and Their Application to the Particular Case of the Ellipsoid, J. Inst. Math. Appl. 7, 111-118 (1971).

G. Dassios
Division of Applied Mathematics
Department of Chemical Engineering
University of Patras and Institute of
Chemical Engineering and High Temperature
Chemical Processes
Patras GR 261 10, Greece

K. Kiriaki
Department of Mathematics
National Technical University of Athens
GR 157 73 Athens
Greece

G. DASSIOS, K. KIRIAKI AND V. KOSTOPOULOS

Inverse thermoelastic Rayleigh scattering by a rigid ellipsoid

ABSTRACT

In this paper the inverse scattering problem for the rigid ellipsoid in linear thermoelasticity is examined. We prove that six measurements of the far field pattern in the low-frequency region are necessary in order to evaluate the semiaxes of the ellipsoid as well as to fix the position of its principal axes.

1. INTRODUCTION

The inverse scattering problem, as it is well known, is concerned with the problem of determining the shape and/or the physical properties of the scattering object from the knowledge of the scattered far-field data.

The mathematical methods used to investigate the inverse, as well as the direct, scattering problems depend heavily on the frequency of the incident wave [11,12].

In the low-frequency region there is the problem that low frequency data does not provide enough information for a sharp resolution of the scattering surface and the optimisation procedure requires a direct scattering problem to be solved at each step of the iterative scheme for arriving at a solution. In low frequencies Angell and Kleinman [1] described a method for finding the dimensions and orientation of an ellipsoid from a constrained optimisation problem for a functional defined in terms of the polarisability tensor elements associated with the object. Dassios [5] solved the inverse scattering problem for the acoustical soft ellipsoid based on the informations

which the direct problem provides in the radiation region. The inverse scattering problem in linear elasticity was examined in [2].

In addressing oneself to the practical problem of reconstructing the shape of an obstacle from far-field data, one is faced with problems of numerical instability. The inverse scattering problem in general is ill-posed, in the sense of Hadamard, and intrinsically nonlinear. So, the aim of the provided methods for solving inverse scattering problems is to be relatively easy in numerical implementation. A survey of the research done in this area can be found in [3,4,13].

The inverse scattering problem which is considered in this paper is to determine the semiaxes and to fix the orientation of a rigid ellipsoidal scatterer embedded in an infinite, homogeneous, isotropic thermoelastic medium from the knowledge of the far-field data, in a finite number of directions.

In Section 2 we formulate the direct scattering problem and we give the necessary results obtained in [7,8].

In Section 3 we examine the inverse scattering problem for a rigid ellipsoid in linear thermoelasticity. We show that the necessary measurements in order to obtain all the information needed about the ellipsoidal scatterer are six. This is a consequence of the simplicity of the expressions for the leading term of the angular scattering amplitudes in low-frequency regions that allows for an exact solution of the corresponding inverse problem. The nonlinearity of the problem, expected for inverse problems, enters via the elliptic integrals, the second power of the values of the semiaxes and the quadratic expressions of the components of the directions of incidence.

Finally, in Section 4 remarks for the stability of the numerical method and a general discussion concerning our approach are presented.

2. FORMULATION OF THE DIRECT SCATTERING PROBLEM

Let us assume that the solid ellipsoid

$$\sum_{i=1}^{3} \frac{x_i^2}{a_i^2} \leq 1, \quad 0<a_3<a_2<a_1<+\infty \tag{1}$$

is embedded in an infinite, homogeneous and isotropic thermoelastic medium.

The isotropic thermoelastic material is characterised by the Lame constants λ,μ, the mass density ρ the coefficient of thermal diffusivity κ, and the linear expansion coefficient a. The unified four-dimensional field

$$\underline{U}(\underline{r}) = (u_1(\underline{r}),\ u_2(\underline{r}),\ u_3(\underline{r}),\ \theta(\underline{r})), \tag{2}$$

where $\underline{u}(\underline{r})$ denotes the displacement and $\theta(\underline{r})$ the temperature field, specifies the stationary thermoelastic state of the medium whenever it belongs to the kernel of the time-independent Biot operator

$$\tilde{L}(\partial_r) = \left[\begin{array}{c|c} (\mu\Delta + \rho\omega^2)\tilde{I}_3 + (\lambda+\mu)\nabla\nabla & -\gamma\nabla \\ \hline q\kappa\eta\nabla & \Delta+q \end{array}\right] \tag{3}$$

In the above expression, γ and η are coupling constants,

$$q = \frac{i\omega}{\kappa} \tag{4}$$

is the thermal analogue of the square of the wavenumber, ω stands for the angular frequency, which is suppressed through the harmonic dependence $e^{-i\omega t}$ and $\tilde{I}_n$ denotes the unit dyadic in n-dimensions. A convenient dimensionless coupling constant is provided by the parameter

$$\epsilon = \frac{\gamma\eta\kappa}{\lambda+2\mu} \tag{5}$$

In the limit as $\gamma \to 0+$, $\eta \to 0+$, $\epsilon \to 0+$ the thermoelastic problem decouples to the corresponding scattering problem in the classical theory of elasticity and an independent heat conduction problem.

A general incident plane wave propagating in the direction $\hat{\underline{k}}$ has the form [7]

$$\underline{\Phi}(\underline{r}) = \underline{\Phi}^1(\underline{r}) + \underline{\Phi}^2(\underline{r}) + \underline{\Phi}^s(\underline{r}) \tag{6}$$

where

$$\underline{\Phi}^1(\underline{r}) = A^1(\hat{\underline{k}},\beta_1)e^{ik_1\hat{\underline{k}}.\underline{r}} \tag{7}$$

is the elastothermal plane wave of amplitude A^1.

$$\underline{\Phi}^2(\underline{r}) = A^2(ik_2\beta_2\hat{\underline{k}},ik_2)e^{ik_2\hat{\underline{k}}.\underline{r}} \tag{8}$$

is the thermoelastic plane wave of amplitude A^2, and

$$\underline{\Phi}^s(\underline{r}) = A^s(\hat{\underline{b}},0)e^{ik_s\hat{\underline{k}}.\underline{r}} \tag{9}$$

is the transverse plane wave of amplitude A^s and polarisation along the direction $\hat{\underline{b}}$, orthogonal to the direction of propagation $\hat{\underline{k}}$. The factor ik_2 in (8) has been added in order to secure analyticity of $\underline{\Phi}^2$ with respect to the wave number. The amplitudes A^1 and A^s have dimensions of length, while A^2 has dimensions of length times temperature.

In consistency with physical reality, the wavenumbers k_1, k_2 are chosen to be those roots of the characteristic systems [7]

$$\left.\begin{aligned} k_1^2 + k_2^2 &= q(1+\epsilon) + k_p^2 \\ k_1^2k_2^2 &= qk_p^2 \end{aligned}\right\} \tag{10}$$

for which $\mathrm{Im}k_i > 0$, $i = 1,2$. These conditions reflect the dissipative character of the thermoelastic medium. The constants

$$\beta_1 = -\frac{k_1\omega\eta}{k_1^2-q}, \tag{11}$$

$$\beta_2 = \frac{ik_2\gamma}{(k_p^2-k_2^2)(\lambda+2\mu)} \tag{12}$$

furnish the appropriate factor in order for $\underline{\Phi}^1$ and $\underline{\Phi}^2$ to belong to the kernel of $\tilde{L}$. On the other hand, as $\epsilon \to 0+$, $\beta_1 \to 0$ and $\beta_2 \to 0$ and the system

decouples in a natural way.

In classical elasticity, we have the wave relations

$$\omega = k_p c_p = k_s c_s \tag{13}$$

where k_p and k_s denote the wavenumbers of the elastic longitudinal and transverse waves, and

$$c_p^2 = \frac{\lambda+2\mu}{\rho}, \qquad c_s^2 = \frac{\mu}{\rho} \tag{14}$$

specify the corresponding phase velocities.

In thermoelasticity, we denote the complex wavenumbers k_1 and k_2 by

$$\begin{aligned} k_1 &= \frac{\omega}{v_1} + id_1, \qquad v_1 > 0, \quad d_1 > 0 \\ k_2 &= \frac{\omega}{v_2} + id_2, \qquad v_2 > 0, \quad d_2 > 0 \end{aligned} \tag{15}$$

where v_1, v_2 are the phase velocities of the elastothermal and thermoelastic waves, respectively, and d_1, d_2 determine the corresponding dissipation coefficient. In the decoupled case, as

$$\left.\begin{aligned} &\epsilon \to 0+ \\ &k_1 \to k_p \\ &k_2 \to \sqrt{q} = (1+i)\sqrt{\frac{\omega}{2\kappa}} \end{aligned}\right\} \tag{16}$$

The general scattered field $\underline{U}$ has a corresponding decomposition into elastothermal, thermoelastic and transverse parts via the relations

$$\left.\begin{aligned} \underline{u}(\underline{r}) &= \underline{u}^1(\underline{r}) + \underline{u}^2(\underline{r}) + \underline{u}^s(\underline{r}) \\ \underline{\theta}(\underline{r}) &= \theta^1(\underline{r}) + \theta^2(\underline{r}) \end{aligned}\right\} \tag{17}$$

A set of ten asymptotic relations given by Kupradze express the radiation conditions which secure the well-posedness of our scattering problem.

The unified total field is expressed through

$$\underline{\Psi}(\underline{r}) = \underline{\Phi}(\underline{r}) + \underline{U}(\underline{r}) \tag{18}$$

The boundary conditions for the rigid ellipsoid in thermoelasticity are

described by

$$\tilde{B}_k(\partial_{\underline{r}}, \hat{\underline{n}})\underline{\psi}(\underline{r}) = \underline{0}, \ \underline{r} \in S, \quad k = 1,2, \tag{19}$$

where the boundary differential operators $\tilde{B}_k$ are expressed as follows:

(i) For the rigid ellipsoid at zero temperature

$$\tilde{B}_1(\partial_{\underline{r}}, \hat{\underline{n}}) = \tilde{I}_4 = \left[\begin{array}{c|c} \tilde{I}_3 & 0 \\ \hline 0 & 1 \end{array}\right] \tag{20}$$

(ii) For the thermally insulated rigid ellipsoid

$$\tilde{B}_2(\partial_{\underline{r}}, \hat{\underline{n}}) = \left[\begin{array}{c|c} \tilde{I}_3 & 0 \\ \hline 0 & \partial_n \end{array}\right] \tag{21}$$

$\hat{\underline{n}}$ being the outward unit normal. The integral postulation of the above problem is presented in [7]. The far-field behaviour of the scattered elastic and thermal fields have been studied in [7] where the following six normalised scattering amplitudes have been introduced.

As $r \to +\infty$, the following asymptotic forms hold true

$$\underline{u}^1(\underline{r}) = e^{-d_1 r}\left[g_r^1(\hat{\underline{r}}, \hat{\underline{k}})h(\frac{\omega}{v_1} r)\hat{\underline{r}} + 0(r^{-2})\right] \tag{22}$$

$$\underline{u}^2(\underline{r}) = e^{-d_2 r}\left[g_r^2(\hat{\underline{r}}, \hat{\underline{k}})h(\frac{\omega}{v_2} r)\hat{\underline{r}} + 0(r^{-2})\right] \tag{23}$$

$$\underline{u}^S(\underline{r}) = \left[g_\theta^S(\hat{\underline{r}}, \hat{\underline{k}})\hat{\underline{\theta}} + g_\phi^S(\hat{\underline{r}}, \hat{\underline{k}})\hat{\underline{\phi}}\right]h(\frac{\omega}{c_S} r)\hat{\underline{r}} + 0(r^{-2}) \tag{24}$$

$$\theta^1(\underline{r}) = e^{-d_1 r}\left[\ell^1(\hat{\underline{r}}, \hat{\underline{k}})h(\frac{\omega}{v_1} r)\hat{\underline{r}} + 0(\underline{r}^{-2})\right] \tag{25}$$

$$\theta^2(\underline{r}) = e^{-d_2 r}\left[\ell^2(\hat{\underline{r}},\hat{\underline{k}})h(\frac{\omega}{v_2} r)\hat{r} + 0(r^{-2})\right] \tag{26}$$

where g_r^m, $m = 1,2$ define the radial elastothermal (for $m = 1$) and the radial thermoelastic (for $m = 2$) normalised scattering amplitude, g_α^s, $a = \theta, \phi$ define the angular normalised scattering amplitudes, and ℓ^m, $m = 1,2$ define the corresponding thermal normalised scattering amplitudes for the elastothermal and thermoelastic waves respectively. The analytic expressions of the thermoelastic scattering amplitudes are given in [7].

In the theory of thermoelasticity, where the existence of dissipation manifests itself via the nonvanishing imaginary parts of the wavenumbers k_1 and k_2, the unit of energy, which enters the definition of the scattering cross-section, is not any longer invariant under translations in the direction of propagation.

This is the reason for introducing in [7] a local determination of the scattering cross-section which in general has the form

$$\sigma = \frac{e_{sc}(\hat{\underline{k}})}{e_{in}} = \frac{e_{sc}^{11}(\hat{\underline{k}})+e_{sc}^{22}(\hat{\underline{k}})+e_{sc}^{ss}(\hat{\underline{k}})+e_{sc}^{12}(\hat{\underline{k}})}{e_{in}^{11}+e_{in}^{22}+e_{in}^{ss}+e_{in}^{12}} \tag{27}$$

where e_{in}^{11} is the part of the unit of energy that corresponds to the elastothermal incident wave,

e_{in}^{22} is the part of the unit energy that corresponds to the thermoelastic incident wave,

e_{in}^{ss} is the part of the unit energy that corresponds to the transverse incident wave,

e_{in}^{12} is the part of the unit of energy that corresponds to the interaction between the elastothermal and thermoelastic incident waves,

and e_{sc}^{11}, e_{sc}^{22}, e_{sc}^{ss}, e_{sc}^{12} describe the corresponding parts of the energy of the scattered wave.

From the analytical expressions of the above quantities, given in [7], it is easy to observe that only the terms e_{sc}^{ss}, e_{in}^{ss} are independent of r, while all the other quantities die out as $r \to +\infty$. Therefore, the quantities e_{sc}^{11}, e_{in}^{11}, e_{sc}^{22}, e_{in}^{22}, e_{sc}^{12} and e_{in}^{12} play no leading role in the definition of the scattering cross-section, which, as $r \to +\infty$, has the asymptotic form

$$\sigma \sim \frac{e_{sc}^{ss}(\hat{\underline{k}})}{e_{in}^{ss}} = \frac{1}{k_s^2 (A^s)^2} \int_{|\hat{\underline{r}}|=1} \left[|g_\theta^s(\hat{\underline{r}},\hat{\underline{k}})|^2 + |g_\phi^s(\hat{\underline{r}},\hat{\underline{k}})|^2 \right] d\Omega(\hat{\underline{r}}) \tag{28}$$

In other words, the scattering cross-section is measured, to a first approximation, by the only nondissipative part of the incident and scattered wave which survives far away from the scattering region, i.e. the transverse part.

For the above equivalent thermoelastic scattering cross-section $\sigma \sim \sigma^{ss}$, an angular type scattering theorem as in classical elasticity [10] is obtained, from which we have the scattering relation

$$\sigma^{ss} = - \frac{4\pi}{k_s^2 (A^s)^2} \left[(\hat{\underline{b}}.\hat{\underline{\theta}}) R_e g_\theta^s(\hat{\underline{k}},\hat{\underline{k}}) + (\hat{\underline{b}}.\hat{\underline{\phi}}) R_e g_\phi^s(\hat{\underline{k}},\hat{\underline{k}}) \right] \tag{29}$$

In low-frequency region the leading terms of the angular thermoelastic scattering amplitudes for the two problems that will be examined in this paper are given by [8]

$$g_\alpha^s = - \frac{ik}{4\pi\mu} \hat{\underline{a}}.\int_S \left[\tilde{T}' \underline{u}_0(\underline{r}') \right] ds(\underline{r}') + 0(k^2), \quad k \to +0, \quad a = \theta, \phi \tag{30}$$

where for the case of a rigid ellipsoid $\underline{u}_0$ is the solution of the following boundary value problem

$$\tau_p^2 \Delta \underline{u}_0(r) + (1-\tau_p^2) \nabla\nabla.\underline{u}_0(\underline{r}) = \underline{0}, \qquad \rho > a_1 \tag{31}$$

$$\underline{u}_0(\underline{r}) = \underline{0}, \qquad \rho = a_1 \tag{32}$$

$$\underline{u}_0(\underline{r}) = (A^1 + A^2\beta_2)\hat{\underline{k}} + A^s\hat{\underline{b}} + 0(\frac{1}{\rho}), \qquad \rho \to +\infty \tag{33}$$

which can be solved by introducing a fictitious potential in the Papkovich representation and then use the freedom it provides to decompose the general boundary condition into two simpler ones. Following the method introduced in [6] we obtain, after very long calculations, the solution

$$u_0(r) = \left[(A^1 + A^2\beta_2)\hat{\underline{k}} + A^s\hat{\underline{b}}\right] \cdot$$

$$\cdot \sum_{m=1}^{3}\left[1 - \frac{(\tau_p^2-1)\alpha_m^2 I_1^m(\rho) - (\tau_p^2+1)I_0^1(\rho)}{(\tau_p^2-1)\alpha_m^2 I_1^m - (\tau_p^2+1)I_0^1}\right] \hat{\underline{x}}_m \otimes \hat{\underline{x}}_m -$$

$$- (\tau_p^2-1)\frac{(\rho^2-\alpha_1^2)}{\sqrt{\rho^2-\mu^2}\sqrt{\rho^2-\nu^2}}\hat{\rho} \otimes \left[(A^1+A^2\beta_2)\hat{\underline{k}}+\underline{A}^s\hat{\underline{b}}\right] \cdot$$

$$\cdot \sum_{m=1}^{3}\frac{\hat{\underline{x}}_m \otimes \hat{\underline{x}}_m \underset{\sim}{r}}{(\rho^2-\alpha_1^2+\alpha_m^2)[(\tau_p^2-1)\alpha^2 I_1^m - (\tau_p^2+1)I_0^1]} \tag{34}$$

where

$$I_0^1(\rho) = \int_\rho^{+\infty} \frac{du}{\sqrt{u^2-\alpha_1^2+\alpha_2^2}\sqrt{u^2-\alpha_1^2+\alpha_3^2}} \tag{35}$$

$$I_1^m(\rho) = \int_\rho^{+\infty} \frac{du}{(u^2-\alpha_1^2+\alpha_m^2)\sqrt{u^2-\alpha_1^2+\alpha_2^2}\sqrt{u^2-\alpha_1^2+\alpha_3^2}} \tag{36}$$

$$I_0^1 = I_0^1(\alpha_1), \quad I_1^m = I_1^m(\alpha_1), \quad m = 1,2,3 \tag{37}$$

and $\hat{\rho}$ is the outward unit normal on the ellipsoid.

The expressions of the scattering amplitudes involve the integral

$$\int_{\rho=\alpha_1} \tilde{T}\, \underline{u}_0(\underline{r})\, ds(\underline{r})$$

where

$$\tilde{T}\, \underline{u}_0(\underline{r}) = 2\mu\hat{\rho}.\nabla\underline{u}_0(\underline{r}) + \lambda\hat{\rho}\nabla.\underline{u}_0(\underline{r}) + \mu\hat{\rho}x(\nabla x\underline{u}_0(\underline{r})) \tag{38}$$

provides the traction field on the surface of the ellipsoid. In view of the potential integral

$$\int_{\rho=\alpha_1} \frac{ds}{\sqrt{\rho^2-\mu^2}\sqrt{\rho^2-\nu^2}} = 4\pi \tag{39}$$

and a series of vector and dyadic calculations we obtain

$$\int_{\rho=\alpha_1} \tilde{T}\, \underline{u}_0(\underline{r})ds(\underline{r}) = -8\pi\bar{\mu} \sum_{m=1}^{3} \frac{(A^1+A^2\bar{\beta}_2)k_m+A^s b_m}{(\tau_p^2-1)\alpha_m^2 I_1^m-(\tau_p^2+1)I_0^1}\, \hat{x}_m \tag{40}$$

Introducing (40) into Eq. (30) we conclude that the angular amplitudes assume the form

$$g_\theta^s(\hat{r},\hat{k}) = 2ik \sum_{m=1}^{3} \frac{[(A^1+A^2\bar{\beta}_2)k_m+A^s b_m]}{(\tau_p^2-1)\alpha_m^2 I_1^m-(\tau_p^2+1)I_0^1}\, (\hat{\theta}.\hat{x}_m) + 0(k^2) \tag{41}$$

$$g_\phi^s(\hat{r},\hat{k}) = 2ik \sum_{m=1}^{3} \frac{[(A^1+A^2\bar{\beta}_2)k_m+A^s b_m]}{(\tau_p^2-1)\alpha_m^2 I_1^m-(\tau_p^2+1)I_0^1}\, (\hat{\phi}.\hat{x}_m) + 0(k^2) \tag{42}$$

So from Eq. (28) by substitution of the angular scattering amplitudes and supposing only S-wave incidence we conclude that

$$\sigma^{ss} = \frac{32\pi}{3} \sum_{m=1}^{3} \left[\frac{b_n}{L_0^n}\right]^2 + 0(k^2) \tag{43}$$

where $\hat{\underline{b}} = (b_1, b_2, b_3)$, $\hat{\underline{b}}.\hat{\underline{k}} = 0$ and

$$L_o^n = (\tau_p^2 - 1)\alpha_n^2 I_1^n(\alpha_1) - (\tau_p^2 + 1) I_o^1(\alpha_1) \quad n = 1,2,3 \tag{44}$$

3. THE INVERSE SCATTERING PROBLEM

We assume that the directions of the principal axes of the ellipsoid described by Eq. (1) coincide with the unit vectors $\hat{\underline{x}}_i$ of the orthogonal cartesian system.

We choose six arbitrary directions of incidence of transverse waves. This choice of S-waves is dictated by the above discussion about the dissipation of all but e_{in}^{ss} and e_{st}^{ss} in the far field region.

In the sequel we will see why six measurements are enough to specify the three semiaxes as well as the three Euler angles that fix the position of the principal axes of the ellipsoid.

For S-incidence along the $\hat{\underline{k}}$-direction we can evaluate the angular thermoelastic scattering amplitudes, so that Eqs. (28, 29, 43) provide the thermoelastic scattering cross-section.

Let the six arbitrary directions of incidence be

$$\begin{aligned}
\hat{\underline{k}}'_j &= \hat{\underline{x}}'_j, \quad j = 1,2,3 \\
\hat{\underline{k}}'_4 &= \frac{1}{\sqrt{2}} (\hat{\underline{x}}'_1 + \hat{\underline{x}}'_2) \\
\hat{\underline{k}}'_5 &= \frac{1}{\sqrt{2}} (\hat{\underline{x}}'_2 + \hat{\underline{x}}'_3) \\
\hat{\underline{k}}'_6 &= \frac{1}{\sqrt{2}} (\hat{\underline{x}}'_3 + \hat{\underline{x}}'_1)
\end{aligned} \tag{45}$$

where $\{\hat{\underline{x}}'_1, \hat{\underline{x}}'_2, \hat{\underline{x}}'_3\}$ form an orthonormal set of an arbitrarily chosen cartesian

system whose origin coincides with the centroid of the ellipsoid. In the sequel with primes we will refer to the arbitrarily chosen system and without primes to the system determined by the principal axes of the ellipsoid.

Let $\tilde{P}$ be the orthogonal matrix that transforms (by rotation) the system $\hat{x}_1', \hat{x}_2', \hat{x}_3'$ to the system $\hat{x}_1, \hat{x}_2, \hat{x}_3$ which determines the principal directions of the ellipsoid, that is we have the relation

$$\underline{r} = \tilde{P}\underline{r}'. \tag{46}$$

The knowledge of the matrix $\tilde{P}$ will provide the orientation of the ellipsoid. The elements P_{ij} of $\tilde{P}$ are expressed via the three Euler angles ϕ, θ, y by the relations

$$\begin{aligned}
P_{11} &= \cos\phi \cos y - \cos\theta \sin\phi \sin y \\
P_{12} &= \sin\phi \cos y + \cos\theta \cos\phi \sin y \\
P_{13} &= \sin\theta \sin y \\
P_{21} &= -\cos\phi \sin y - \cos\theta \sin\phi \cos y \\
P_{22} &= -\sin\phi \sin y + \cos\theta \cos\phi \cos y \\
P_{23} &= \sin\theta \cos y \\
P_{31} &= \sin\theta \sin\phi \\
P_{32} &= -\sin\theta \cos\phi \\
P_{33} &= \cos\theta
\end{aligned} \tag{47}$$

Since the relation (43) is referred to the principal axes system it follows that (43) holds true after the transformation described by the matrix $\tilde{P}$, has been applied to the directions of polarisation $\hat{b}'$, that is

$$\hat{\underline{b}}_j = \tilde{P}\hat{\underline{b}}_j' \qquad j = 1,\ldots,6 \tag{48}$$

So, the six measurements, in view of Eq. (43), yield

$$m_j = \frac{32\pi}{3}(\hat{\underline{b}}_j)^T \tilde{A}\, \hat{\underline{b}}_j, \qquad j = 1,\ldots,6 \tag{49}$$

where

$$A_{ik} = 0, \qquad i \neq k$$

$$A_{ik} = \frac{1}{L_o^1} \qquad i = k = 1,2,3 \tag{50}$$

From Eqs. (48, 49) we conclude that

$$m_j = \frac{32\pi}{3}(\hat{\underline{b}}_j')^T \tilde{P}^T \tilde{A}\, \tilde{P}\, \hat{\underline{b}}_j' \tag{51}$$

Substituting Eq. (45) in Eq. (51) we obtain the following system of equations for the unknown quantities $\phi, \theta, y, a_1, a_2, a_3$

$$\sum_{n=1}^{3} \frac{P_{ni}^2}{L_o^n} = \frac{3m_i}{32\pi}, \qquad i = 1,2,3$$

$$\sum_{n=}^{3} \frac{(P_{ni}+P_{nj})^2}{L_o^n} = \frac{3m_k}{16\pi}, \qquad (i,j,k) = \{(1,2,4),(2,3,5),(3,1,6)\} \tag{52}$$

From Eqs. (52) we conclude that

$$\sum_{n=1}^{3} \frac{P_{ni}P_{nj}}{L_o^n} = \frac{3(2m_k - m_i - m_j)}{32\pi} \tag{53}$$

where (i,j,k) as in Eq. (52).

So, we have the matrix relation

$$\tilde{P}^T \tilde{A}\, \tilde{P} = \tilde{M} \tag{54}$$

where $\tilde{M}$ is a real symmetric matrix with known, from the six measurements, elements given by the relations

$$M_{ii} = \frac{3m}{16\pi} \qquad i = 1,2,3,$$

$$M_{ij} = M_{ji} = \frac{3(2m_k - m_i - m_j)}{32\pi} \tag{55}$$

(i,j,k) as in Eq. (52).

From Eq. (54) we conclude that

$$\tilde{A} = \tilde{P}\,\tilde{M}\,\tilde{P}^{T} \tag{56}$$

Since $\tilde{P}$ is an orthogonal matrix, $\tilde{M}$ is a real symmetric one, and $\tilde{A}$, from Eq. (50), is a diagonal matrix it follows that the eigenvalues of $\tilde{M}$ are given by

$$\lambda_n = \frac{1}{L_o^n}, \qquad n = 1,2,3 \tag{57}$$

while the columns of $\tilde{P}^{T}$ are the corresponding orthonormal eigenvectors. So, the eigenvectors of the known matrix $\tilde{M}$ specify the orientation of the ellipsoid. In order to evaluate the semiaxes of the ellipsoid we apply the following procedure.

From Eq. (57) we have that

$$L_o^n = \frac{1}{\lambda_n}, \qquad n = 1,2,3 \tag{58}$$

From Eq. (44) and the well-known formula

$$\sum_{n=1}^{3} a_n^2 I_1^n = I_o^1$$

which relates the elliptic integrals we derive that

$$I_o^1(a_1) = -\frac{1}{2(\tau_p^2+2)} \sum_{n=1}^{3} \frac{1}{\lambda_n} = M_o \tag{59}$$

From Eqs. (44, 57) we conclude that

$$a_n^2 I_1^n(a_1) = \frac{1}{(\tau_p^2-1)} \left\{ \frac{1}{\lambda_n} - \frac{\tau_p^2+1}{2(\tau_p^2+2)} \sum_{i=1}^{3} \frac{1}{\lambda_i} \right\} = M_n, \qquad n = 1,2,3 \tag{60}$$

where M_o, M_n are known quantities.

In order to bring the elliptic integral I_o^1 to its canonical form we

perform the transformation

$$x = \frac{a_1^2 - a_3^2}{t^2} - a_1^2 \tag{61}$$

and we conclude in standard notation of elliptic integrals that

$$I_0^1(a_1) = -(a_1^2 - a_3^2)^{-1/2} \int_0^{\sin\phi_0} \frac{dt}{\sqrt{1-t^2}\sqrt{1-t^2\sin^2 a_0}}$$

$$= -(a_1^2 - a_3^2)^{-1/2} F(\phi_0, a_0) \tag{62}$$

where

$$\phi_0 = \sin^{-1}\sqrt{\frac{a_1^2 - a_3^2}{a_1^2}}$$

$$\alpha_0 = \sin^{-1}\sqrt{\frac{a_1^2 - a_2^2}{a_1^2 - a_3^2}} \tag{63}$$

In order to bring the elliptic integral of the second kind I_n^1 to its canonical form $E(\phi_0, a_0)$, we perform the same as above transformation, while for the integral I_1^1 we conclude

$$\int_0^{\sin\phi_0} \frac{t^2 dt}{\sqrt{1-t^2}\sqrt{1-t^2\sin^2 a_0}} = -(a_1^2 - a_3^2)^{3/2} I_1^1(a_1) \tag{64}$$

Finally after some algebraic calculations we conclude that

$$I_1^1(a_1) = (a_1^2 - a_3^2)^{-3/2} \frac{1}{\sin^2 a_0}[E(\phi_0, \alpha_0) - F(\phi_0, a_0)] \tag{65}$$

If we introduce the notation

$$s_1 = \frac{a_2}{a_1}, \quad s_2 = \frac{a_3}{a_1} \tag{66}$$

use the relations

$$\sum_{n=1}^{3} I_1^n = \frac{1}{a_1 a_2 a_3}$$

$$\sum_{n=1}^{3} a_n^2 I_1^n = I_0^1$$

and the equations (62) and (64) we conclude the relations

$$-\frac{s_1^2 s_2^2 M_1 + s_2^2 M_2 + s_1^2 M_3}{s_1 s_2} (1-s_2^2)^{-1/2} F(\phi_0, a_0) = \sum_{i=1}^{3} M_i \tag{67}$$

$$\frac{s_1^2 s_2^2 M_1 + s_2^2 M_2 + s_1^2 M_3}{s_1 s_2} (1-s_2^2)^{-1/2} (1-s_1^2)^{-1} \left[E(\phi_0, a_0) - F(\phi_0, a_0) \right]$$
$$= \Lambda_1 \text{ (known quantity)} \tag{68}$$

where in view of (66)

$$\phi_0 = \sin^{-1}(1-s_2^2)^{1/2} \tag{69}$$

$$a_0 = \sin^{-1}\left(\frac{1-s_1^2}{1-s_2^2}\right)^{1/2}$$

In order to solve numerically the nonlinear system given by equations (67, 68) we use the same iterative scheme as in [2].

4. DISCUSSION

In this paper the inverse scattering problem of linear thermoelasticity for a rigid ellipsoid in the low frequency region is examined. We use information from the solution of the direct thermoelastic scattering problem and propose a method for the corresponding inverse.

In order to evaluate the dimensional characteristics and the orientation of the scatterer we need six measurements of the far-field data in the

low-frequency region under the condition that we have a knowledge of the shape and the boundary *a priori* conditions. This type of approach of the inverse problem has been proposed in [5]. With a similar technique the scattering problem in linear elasticity has been solved in [2].

If the boundary conditions on the surface of the scatterer alter the above approach can also be used. Obviously from the solution of the direct problem we obtain a highly nonlinear system which cannot be solved by a simple and rapidly convergent iterative scheme.

Looking closer at the results contained in this work, one can see no basic difference if we consider a thermally insulated rigid scatterer or a rigid scatterer at zero temperature. In other words, the results are independent of the temperature boundary condition on the surface of the rigid scatterer.

This is a reflection of the fact that thermal effects do not enter the Rayleigh approximation of the transverse field in the radiation zone.

The corresponding temperature field appeared in the low-frequency approximations of order higher than the leading one. Even in these higher order approximations the dependence on the temperature is implicit through its effects on the particular form of the displacement field.

ACKNOWLEDGEMENT

The present work has been supported by the EEC SCIENCE program under grant number SC1-0079.

REFERENCES

[1] Angell, T.S. and Kleinman, R.E., Polarizability Tensors in Low Frequency Inverse Scattering, Radio Science 22, 1120 (1987).

[2] Apostolopoulos, T., Kiriaki, K. and Polyzos, D, The Inverse Scattering Problem for a Rigid Ellipsoid in Linear Elasticity, Inverse Problems (in press).

[3] Colton, D., The Inverse Scattering Problem for Time-Harmonic Acoustic Waves, SIAM Rev. 26, 323 (1984).

[4] Colton, D. and Kress, R., Integral Equation Methods in Scattering Theory, J. Wiley, New York (1983).

[5] Dassios, G., The Inverse Scattering Problem for a Soft Eellipsoid, J. of Math. Physics 28 (12), 2858 (1987).

[6] Dassios, G. and Kiriaki, K., The Rigid Ellipsoid in the Presence of a Low-Frequency Elastic Wave, Q. Appl. Math. 43, 435 (1986).

[7] Dassios, G. and Kostopoulos, V., The Scattering Amplitudes and Cross-Sections in the Theory of Thermoelasticity, SIAM J. Appl. Math. 48, 79-98 (1988).

[8] Dassios, G. and Kostopoulos, V., On Rayleigh Expansions in Thermoelastic Scattering, SIAM J. Appl. Math. (in press).

[9] Dassios, G. and Kostopoulos, V., Thermoelastic Rayleigh Scattering by a Rigid Ellipsoid, (to appear).

[10] Dassios, G., Kiriaki, K. and Polyzos, D., On the Scattering Amplitudes for Elastic Waves, ZAMP, 856 (1987).

[11] Datta, S. K. and Sabina, F. J., Matched Asymptotic Expansions Applied to Diffraction of Elastic Waves, from Low and High Frequency Asymptotics Vol 2, edited by V. K. Varadan and V. V. Varadan, North Holland, 1986.

[12] King, R., Wu, T., Scattering and Diffraction of Waves, Harvard Univ. Press, 1959.

[13] Sleeman, B., The Inverse Problem of Acoustic Scattering, IMA J. Appl. Math. 29, 113 (1982).

G. Dassios
Division of Applied Mathematics
Department of Chemical Engineering
University of Patras, and
Institute of Chemical Engineering and
High Temperature Chemical Processes
GR 261 10 Patras, Greece

K. Kiriaki
Department of Mathematics
National Technical University of Athens
GR 157 73 Athens, Greece

V. Kostopoulos
Department of Applied Mechanics
University of Patras,
GR 261 10 Patras, Greece

A. FARIDANI

Reconstructing from efficiently sampled data in parallel-beam computed tomography

1. INTRODUCTION

In computed tomography (CT) an object is exposed to radiation which is measured after passing through this object. From these measurements a certain function f, characterizing the interaction of the object with the radiation is reconstructed. This function represents an 'image' of the interior of the object.

In diagnostic radiology, the classical application of CT, X-rays are used as radiation. The measurements are then line integrals of the X-ray absorption coefficient f. This leads to the mathematical problem of reconstructing a function from its line integrals. In two dimensions this means the inversion of the Radon transform.

In this paper we examine how many line integrals have to be measured in order to achieve a certain accuracy and resolution. We confine our investigation to the so-called parallel-beam sampling geometry. First we describe the application of multidimensional sampling theory to sampling the two-dimensional Radon transform. This approach was first taken by Lindgren and Rattey [11, 16] and further developed by Natterer [12]. It leads to sampling schemes which need a minimal amount of data to ensure that the Radon transform Rf is determined by the measured values up to a small error.

The question of how to achieve good reconstructions from such data has been studied by Kruse [9] who obtained error estimates for the filtered backprojection algorithm, the most popular reconstruction method. The main purpose of this paper is to derive estimates extending Kruse's results and to

use this theory to obtain improved reconstructions. The paper is organized as follows:

In the remainder of this section we give a brief description of the Radon transform and its inversion. In the next section we derive a version of Petersen and Middleton's sampling theorem [14] suitable for sampling functions defined on $[0,2\pi]^{n_1} \times \mathbb{R}^{n_2}$. Our proof differs from the approach taken in [12, p.64] and is based on a Poisson summation formula derived in [9]. Furthermore we characterize the class of sampling schemes to which the sampling theorem is applicable and which are suitable for sampling the two-dimensional Radon transform in practical applications. We call these schemes 'admissible sampling lattices'. Application of the sampling theorem leads to sampling conditions which have to be satisfied by a sampling scheme. The results of [11, 12, 16] for the usually employed standard lattice and the so-called 'interlaced lattice' are derived. The interlaced lattice, first suggested for sampling the Radon transform by Cormack [2], requires a minimal amount of data. Sampling conditions for the standard lattice were already derived in different ways by Bracewell [1], and by Crowther, De Rosier and Klug [3].

To clarify the question of accuracy of reconstructions, we give a detailed error analysis for a particular reconstruction method, the filtered backprojection algorithm. In section 3 we describe the implementation of the algorithm for data sampled on an admissible sampling lattice. In the following section we derive estimates for the reconstruction error. We extend the results of [9] by taking into account the influence of a certain interpolation step occuring in the algorithm. This interpolation was neglected in [9]. It turns out that it is critical when the interlaced

lattice is used, but that good reconstructions can be achieved with a sensible choice of the parameters of the algorithm.

The proof of the estimates already derived by Kruse is closely related to the proof of the sampling theorem given in section 2. This indicates that the filtered backprojection algorithm might be more suitable than other reconstruction methods when minimal sampling schemes like the interlaced lattice are used.

The last section is devoted to numerical experiments and their discussion in the light of the results of section 4. It turns out that the theoretical results explain the imminent numerical difficulties and show the way to obtain good reconstructions, thereby improving previous numerical results.

In the following we give a brief introduction to the Radon transform and its inversion and also introduce some notation.

Let $\mathbb{R}$, $\mathbb{N}$, $\mathbb{Z}$ denote the real numbers, natural numbers, and integers, respectively. Furthermore let Ω denote the unit disk in $\mathbb{R}^2$ and T the interval $[0,2\pi)$. For $\phi \in T$ the variable θ will always denote the unit vector $(\cos \phi, \sin \phi)^T$.

The Radon transform of a function $f \in C_0^\infty(\Omega)$ is given by

$$Rf(\phi,s) = \int_{-\infty}^{\infty} f(s\theta+t\theta^\perp)dt, \quad \phi \in T, \quad s \in \mathbb{R}. \tag{1}$$

For a survey on the mathematical properties of the Radon transform and its many applications see e.g. [4, 6, 7, 8, 12].

As we will see in Theorem 1.1 below, the Radon transform is closely related to the Fourier transform. The Fourier transform of a function $F \in L_1(\mathbb{R}^n)$ is given by

$$\hat{F}(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} F(x)e^{-ix.\xi}dx \tag{2}$$

where $x.\xi = \Sigma_{i=1}^n x_i\xi_i$ denotes the dot product. The Fourier transform can be

extended to the space of tempered distributions. Its inverse is given by the formula

$$F(x) = \hat{\hat{F}}(-x). \tag{3}$$

The convolution $f*g$ of two L_1-functions is given by $f*g(x) = \int_{\mathbb{R}^n} f(y)g(x-y)dy$.

We have

$$(f*g)\hat{}(\xi) = (2\pi)^{n/2}\hat{f}(\xi)\hat{g}(\xi). \tag{4}$$

In case of the two-dimensional Radon transform the Fourier transform with respect to the second variable alone is also of interest. It is defined by

$$(Rf)\hat{}^2(\phi,\sigma) = (2\pi)^{-1/2}\int_{\mathbb{R}} Rf(\phi,s)e^{-i\sigma s}ds \tag{5}$$

where $\hat{}^2$ indicates that the Fourier transform is taken with respect to the second variable only.

The Radon transform is closely related to the Fourier transform by means of the following theorem:

Theorem 1.1

Let $f \in C_0^\infty(\Omega)$. Then

$$(Rf)\hat{}^2(\phi,\sigma) = \sqrt{2\pi}\hat{f}(\sigma\theta). \tag{6}$$

Proof: Replace $Rf(\phi,s)$ in (5) by the righthand-side of (1) and use the definition (2) of the Fourier transform. □

Taking a two-dimensional inverse Fourier transform on both sides of (6) yields immediately an inversion formula for Rf:

$$f(x) = (2\pi)^{-1}\int_{\mathbb{R}^2}\hat{f}(\xi)e^{ix\cdot\xi}d\xi$$

$$= (2\pi)^{-1}\int_0^{2\pi}\int_0^\infty |\sigma|\hat{f}(\sigma\theta)e^{i\sigma x\cdot\theta}d\sigma d\phi$$

$$= \frac{1}{2}(2\pi)^{-1} \int_0^{2\pi} \int_{-\infty}^{\infty} |\sigma| \hat{f}(\sigma\theta) e^{i\sigma x.\theta} d\sigma d\phi$$

$$= \frac{1}{2}(2\pi)^{-3/2} \int_0^{2\pi} \int_{-\infty}^{\infty} |\sigma| (Rf)\hat{}^2(\phi,\sigma) e^{i\sigma x.\theta} d\sigma d\phi \tag{7}$$

$$= \frac{1}{4\pi} \int_0^{2\pi} q(\phi, x.\theta) d\phi$$

where $q(\phi,s) = H\frac{\partial}{\partial s} Rf(\phi,s)$, and H denotes the Hilbert transform acting on the second variable. For other types of inversion formulas see e.g. [12, Chapter II].

It is readily seen from (7) that an exact inversion of the Radon transform is unstable because of the amplification of high frequencies due to the filter $|\sigma|$ in the inner integral of (7). More suitable for numerical inversion are approximate inversion formulas of the following kind, where this instability is removed by means of a suitable low-pass filter:

Theorem 1.2

Let $W_b: R^2 \to R$ be a radially symmetric low-pass filter with cut-off frequency b, i.e.

$$\hat{W}_b(\xi) = (2\pi)^{-1} \psi(|\xi|/b)$$

with $0 \leq \psi(\sigma) \leq 1$ and $\psi(\sigma) = 0$ for $|\sigma| > 1$. Define w_b by

$$\hat{w}_b(\sigma) = \frac{1}{2}(2\pi)^{-3/2} |\sigma| \psi(|\sigma|/b)$$

and let $f \in C_0^\infty(\Omega)$. Then

$$W_b * f(x) = \int_0^{2\pi} \int_R w_b(x.\theta - s) Rf(\phi,s) ds d\phi. \tag{8}$$

Proof: Proceeding as above we obtain

$$W_b * f(x) = \int_{R^2} \hat{W}_b(\xi) \hat{f}(\xi) e^{i\xi.x} d\xi$$

$$= \frac{1}{4\pi} \int_0^{2\pi} \int_{-\infty}^{\infty} |\sigma| \psi(|\sigma|/b) \hat{f}(\sigma\theta) e^{i\sigma x.\theta} d\sigma d\phi$$

$$= \int_0^{2\pi} \int_{\mathbf{R}} w_b(x.\theta - s) Rf(\phi,s) ds d\phi$$

where we have used (4). □

This kind of approximate inversion formulas goes back to [19], and provides the basis for the filtered backprojection reconstruction algorithm [15], which we will describe in section 3. There are many other methods for the numerical inversion of the Radon transform. For a survey of such reconstruction methods see [12, Chapter V].

In tomography one has to compute approximations to f from measurements of Rf. The question arises, how many measurements are needed and at which points (ϕ,s) we should sample $Rf(\phi,s)$. Therefore we now discuss the relevant sampling theorem.

2. SAMPLING ON $[0,2\pi)^{n_1} \times \mathbf{R}^{n_2}$

Since $Rf(\phi,s)$ is 2π-periodic in the first variable, we need a sampling theorem for functions which are 2π-periodic in some of their variables. They can be regarded as functions on $T^{n_1} \times \mathbf{R}^{n_2}$, where T denotes the interval $[0,2\pi)$. One essential tool for deriving the sampling theorem is the Fourier transform on $T^{n_1} \times \mathbf{R}^{n_2}$ which is defined as follows: Let S^{n_1,n_2} denote the class of C^∞-functions of $n_1 + n_2$ variables which are 2π-periodic in their first n_1 variables and in $C_0^\infty(\mathbf{R}^{n_2})$ with respect to the last n_2 variables. The Fourier transform of a function $G \in S^{n_1,n_2}$ is a function $\hat{G}: Z^{n_1} \times \mathbf{R}^{n_2} \to C$ given by

$$\hat{G}(k,\sigma) = (2\pi)^{-(n_1+n_2)/2} \int_{T^{n_1}} \int_{R^{n_2}} G(\phi,s) e^{-i\sigma.s} e^{-ik.\phi} ds d\phi$$

with $(k,\sigma) \in Z^{n_1} \times R^{n_2}$ and $(\phi,s) \in T^{n_1} \times R^{n_2}$.

Let $L_1(Z^{n_1} \times R^{n_2})$ denote the space of all functions $g = Z^{n_1} \times R^{n_2} \to C$ for which

$$\sum_{k \in Z^{n_1}} \int_{R^{n_2}} |g(k,\sigma)| d\sigma < \infty.$$

The inverse Fourier transform of a function $g \in L_1(Z^{n_1} \times R^{n_2})$ is given by $\tilde{g}: T^{n_1} \times R^{n_2} \to C$,

$$\tilde{g}(\phi,s) = (2\pi)^{-(n_1+n_2)/2} \sum_{k \in Z^{n_1}} \int_{R^{n_2}} g(k,\sigma) e^{i\sigma.s} d\sigma \; e^{ik.\phi}.$$

For functions $G \in S^{n_1,n_2}$ we have $G = \tilde{\hat{G}}$.

For $(k,\sigma) \in Z^{n_1} \times R^{n_2}$ and $(\phi,s) \in T^{n_1} \times R^{n_2}$ we define the dot product in the natural way, i.e.

$$(k,\sigma).(\phi,s) = \sum_{i=1}^{n_1} k_i \phi_i + \sum_{j=1}^{n_2} s_j \sigma_j.$$

As to the set of points (ϕ,s), at which functions are sampled, we restrict our analysis to the so-called sampling lattices:

Definition 2.1

We call a non-singular (n_1+n_2, n_1+n_2)-matrix W feasible [9] for sampling on $T^{n_1} \times R^{n_2}$, if $2\pi e^i \in WZ^n$ for $i = 1,\ldots,n_1$, where the e^i are canonical unit vectors of $R^{n_1+n_2}$, i.e. $e^i_j = \delta_{ij}$. The set

$$L = (WZ^{n_1+n_2}) \cap (T^{n_1} \times R^{n_2})$$

$$= \{Wk \mid k \in Z^{n_1+n_2}, (Wk)_i \in T, \; i = 1,\ldots,n_1\}$$

is called the sampling lattice generated by W. The set

$$L' = 2\pi W^{-T} Z^{n_1+n_2} = \{2\pi W^{-T} k | k \in Z^{n_1+n_2}\}$$

is called the dual lattice with respect to L.

The simplest example for a sampling lattice is the so-called standard lattice on $T \times \mathbf{R}$. It is generated by a diagonal (2,2)-matrix with diagonal $(2\pi/P,h)$ where $P \in \mathbf{N}$, $h > 0$ and is equal to

$$L_S := \left\{ \begin{bmatrix} \phi \\ s \end{bmatrix} \in T \times \mathbf{R}: \ \phi = \frac{2\pi j}{P}, \ s = h\ell; \ j,\ell \in Z, \ 0 \leq j < P \right\}. \qquad (9)$$

The condition for feasibility means that $WZ^{n_1+n_2}$ is 2π-periodic in the first n_1 variables and implies that the dual lattice is a subset of $Z^{n_1} \times \mathbf{R}^{n_2}$ which is the domain of the Fourier transform of functions defined on $T^{n_1} \times \mathbf{R}^{n_2}$.

A given sampling lattice does not uniquely determine the generating matrix W. It does determine, however, |det W| as well as the dual lattice. The dual lattice L' can be characterized in terms of L by

$$L' = \{u \in Z^{n_1} \times \mathbf{R}^{n_2} | \forall v \in L: u.v/(2\pi) \in Z\}.$$

For a given lattice L we define the lattice constant c_L by

$$c_L = (2\pi)^{-(n_1+n_2)/2} |\det W|$$

where W is any feasible matrix generating L.

A lattice must satisfy a certain requirement for its density of sampling points in order to be a suitable lattice for sampling a given function g. The following theorem shows that this density requirement is determined by the size and shape of a set $K \subset Z^{n_1} \times \mathbf{R}^{n_2}$ in which the Fouier transform $\hat{g}$ is concentrated in the L_1-sense: The translates of K with respect to the dual lattice must be mutually disjoint. Note that this requirement of sparsity for the points of the dual lattice corresponds to a density requirement for

the points of the lattice itself.

Theorem 2.2

Let L be a sampling lattice for $T^{n_1} \times R^{n_2}$ and $K \subset Z^{n_1} \times R^{n_2}$ such that the sets $K + u$, $u \in L'$ are mutually disjoint. Let $g \in S^{n_1,n_2}$ and

$$S_L g(x) = c_L \sum_{v \in L} \tilde{\chi}_K(x-v) g(v)$$

where χ_K is the characteristic function of K. Then

$$\sup_{(\phi,s) \in T^{n_1} \times R^{n_2}} |(S_L g - g)(\phi,s)| \leq 2(2\pi)^{-(n_1+n_2)/2} \sum_{k \in Z^{n_1}} \int_{(k,\sigma) \notin K} |\hat{g}(k,\sigma)| d\sigma. \quad (10)$$

The theorem is a modified version of the sampling theorem of Petersen and Middleton [14]. It shows that the function g is essentially determined by its values on a sampling lattice L, if $|\hat{g}(k,\sigma)|$ is small outside a set K satisfying the conditions of the theorem. The key for the proof of the theorem as well as for proving some of the error estimates for the filtered backprojection algorithm to be presented later is the following lemma:

Lemma 2.3

Let L be a sampling lattice for $T^{n_1} \times R^{n_2}$, $G \in S^{n_1,n_2}$, $\hat{F} \in L_1(Z^{n_1} \times R^{n_2})$, and $F = (\hat{F})^{\sim}$ Then for every $y \in T^{n_1} \times R^{n_2}$

$$c_L \sum_{v \in L} F(y-v) G(v) =$$

$$= (2\pi)^{-(n_1+n_2)/2} \sum_{k \in Z^{n_1}} \int_{R^{n_2}} \hat{F}(k,\sigma) \sum_{u \in L'} \hat{G}((k,\sigma)-u) e^{i(k,\sigma).y} d\sigma. \quad (11)$$

Proof: Replace F(y-v) on the lefthand side of (11) by

$$(2\pi)^{-(n_1+n_2)/2} \sum_{k\in Z^{n_1}} \int_{R^{n_2}} \hat{F}(k,\sigma) e^{i(k,\sigma)\cdot(y-v)} d\sigma$$

and use the Poisson summation formula:

$$c_L \sum_{v\in L} G(v) e^{-i(k,\sigma)\cdot v} = \sum_{u\in L'} \hat{G}((k,\sigma)-u) \tag{12}$$

which is proved in [9].

□

Proof of Theorem 2.2: Using the lemma with $F = \tilde{\chi}_K$, we obtain

$$|(S_L g - g)(\phi,s)|$$

$$= \left| (2\pi)^{-(n_1+n_2)/2} \sum_{k\in Z^{n_1}} \int_{R^{n_2}} \left[\chi_K(k,\sigma) \sum_{u\in L'} \hat{g}((k,\sigma)-u) - \hat{g}(k,\sigma) \right] e^{ik\cdot\phi} e^{i\sigma\cdot s} d\sigma \right|$$

$$= \left| (2\pi)^{-(n_1+n_2)/2} \sum_{(k,\sigma)\in K} \sum_{u\in L', u\neq 0} \hat{g}((k,\sigma)-u) e^{ik\cdot\phi} e^{i\sigma\cdot s} d\sigma \right.$$

$$\left. - (2\pi)^{-(n_1+n_2)/2} \sum_{(k,\sigma)\in Z^{n_1}\times R^{n_2}\setminus K} \hat{g}(k,\sigma) e^{ik\cdot\phi} e^{i\sigma\cdot s} d\sigma \right|$$

$$\leq 2(2\pi)^{-(n_1+n_2)/2} \sum_{(k,\sigma)\in Z^{n_1}\times R^{n_2}\setminus K} |\hat{g}(k,\sigma)| d\sigma$$

where we have used the fact that the sets $K + u, u \in L'$ are actually disjoint, and therefore

$$\sum_{(k,\sigma)\in K} \sum_{u\in L', u\neq 0} |\hat{g}((k,\sigma)-u)| d\sigma \leq \sum_{Z^{n_1}\times R^{n_2}\setminus K} |\hat{g}(k,\sigma)| d\sigma.$$

□

In the following we apply Theorem 2.2 to the sampling of the Radon transform on $T \times \mathbf{R}$. In practice one wants to measure $Rf(\phi,s)$ for a collection of different 'views', where each view consists of a collection of measurements with ϕ fixed. Since this corresponds to measuring the line integrals for sets of parallel lines, this sampling geometry is called parallel-beam geometry. Suitable sampling lattices for this geometry should contain more than one element (ϕ,s) for each occurring value of ϕ. We call lattices with this property, i.e. with the property

$$(\phi,s) \in \mathbf{L} \Rightarrow \exists s' \neq s: (\phi,s') \in \mathbf{L}$$

admissible sampling lattices. The following lemma characterizes these lattices. It turns out that only a finite number LM of different values of ϕ occurs and that for fixed ϕ we get a set of equidistant values of s, where the distance d between two nearest neighbours does not depend on ϕ.

Lemma 2.4.

Let $\mathbf{L}$ be an admissible sampling lattice for $T \times \mathbf{R}$. Then there exists $d > 0$ and $L,M,N \in \mathbf{N}$ with $0 \leq N < M$ and $\gcd(M,N) = 1$ such that the matrix

$$W(d,L,M,N) = \begin{bmatrix} \frac{2\pi}{L} & \frac{2\pi N}{LM} \\ 0 & d/M \end{bmatrix}$$

generates $\mathbf{L}$. Hence $\mathbf{L}$ is equal to the set

$$\begin{aligned} &\mathbf{L}(d,L,M,N) \\ &:= \left\{ (\phi_j, s_{j\ell}) \in T \times \mathbf{R} \,\middle|\, \phi_j = \frac{2\pi j}{LM}, s_{j\ell} = d(\ell+\delta_j/M) \text{ with } \delta_j \in \mathbf{N} \right. \\ &\qquad \left. \text{such that } (N\delta_j - j)/M \in Z;\ j = 0,\ldots,LM-1;\ \ell \in Z \right\}. \end{aligned} \tag{13}$$

Proof: Consider an integer 2×2 matrix U with $|\det U| = 1$. Then $UZ^2 = Z^2$ and the lattices generated by a matrix W and by WU coincide. If W is feasible there exists an integer vector $(L_1,L_2)^T$ such that $W(L_1,L_2)^T = (2\pi,0)^T$. There exists an integer matrix U with $|\det U| = 1$ such

that $U(L,0)^T = (L_1,L_2)^T$ where $L = \gcd(L_1,L_2)$. Hence the first column of WU is equal to $(2\pi/L,0)^T$. Since L is admissible the ratio of the elements of the first row of WU is a rational number. Hence $(WU)_{12}$ can be written as $\frac{2\pi}{L}\frac{N}{M}$ with $\gcd(M,N) = 1$. Changing N by a multiple of M does not chance WUZ^2. Hence N can be chosen such that $0 \leq N \leq |M|-1$. Since replacing N by -N gives the same new lattice as replacement of N by M - N, it is readily seen that allowing for negative values of d,L,M does not yield further lattices. □

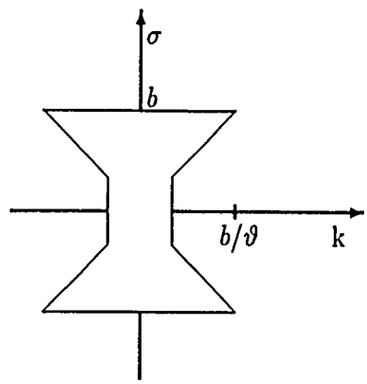


Figure 1: The set $K_0(\tau,b)$

In order to apply Theorem 2.2 to the Radon transform Rf of a function f we have to determine a suitable set K such that the righthandside of (10) is small. This has been done by Lindgren and Rattey [11, 16] who obtained the set

$$K_0(\tau,b) = \{(k,\sigma) \in \mathbb{Z} \times \mathbb{R} \mid |\sigma| < b, |k| < \frac{1}{\tau}\max(|\sigma|,(1-\tau)b)\} \qquad (14)$$

which is shown in Figure 1.

Here b plays the role of an *essential band-width* of f in the sense that

$$\epsilon_0(f,b) := \int_{|\xi|>b} |\hat{f}(\xi)|\,d\xi$$

is sufficiently small. τ is an auxiliary parameter between 0 and 1 and usually chosen very close to 1. The following rigorous estimate was given by Natterer [12, p. 73]:

Theorem 2.5

Let $f \in C_0^\infty(\Omega)$. Then

$$\pi^{-1} \sum_{(k,\sigma) \in \mathbf{Z} \times \mathbf{R} \setminus K_0(\tau,b)} \int |\widehat{Rf}(k,\sigma)|\,d\sigma \le \frac{8}{\pi\tau}\,\epsilon_0(f,b) + \|f\|_{L_1}\eta(\tau,b)$$

where $\eta(\tau,b)$ satisfies an estimate

$$0 \le \eta(\tau,b) \le C(\tau)e^{-\lambda(\tau)b}$$

with constants $C(\tau),\lambda(\tau) > 0$.

According to the last two theorems the procedure of determining appropriate sampling lattices is as follows: Use a-priori information to determine a value of b for which $\epsilon_0(f,b)$ is small. Choose τ close to one (e.g. 0.99) and look for lattices L such that the sets $K_0(\tau,b) + u$, $u \in L'$ are mutually disjoint.

If additional information about f is available it may be possible to replace $K_0(\tau,b)$ by a smaller set. For example if f is radially symmetric $\widehat{Rf}(k,\sigma) = 0$ for $k \neq 0$, and $K_0(\tau,b)$ can be replaced by the much smaller set $\tilde{K}(b) = \{(0,\sigma): |\sigma| \le b\}$. This of course greatly reduces the necessary amount of measurements. In the following we will assume that the only available a-priori information about f is the essential band-width b and therefore always work with $K_0(\tau,b)$.

For example, the standard lattice L_S given in (9) is identical with L(h,P,1,0) of Lemma 2.4. The dual lattice L_S' is therefore generated by the

matrix

$$2\pi W(h,P,1,0)^{-T} = \begin{bmatrix} P & 0 \\ 0 & 2\pi/h \end{bmatrix}.$$

The sets $K_0(\tau,b) + u$, $u \in L_S^*$ are mutually disjoint iff $P \geq 2b/\tau$ and $h \leq \pi/b$. Since the Radon transform satisfies the relation

$$Rf(\phi+\pi,-s) = Rf(\phi,s) \tag{15}$$

it is advantageous to choose P even, say $P = 2p$, $p \in \mathbb{N}$. Then $Rf(2\pi(j+p)/(2p),h\ell) = Rf(2\pi j/(2p),-h\ell)$, so that we only have to sample for $0 \leq j < p$.

Applying Theorems 2.2 and 2.5 yields the well-known sampling conditions $p \geq b/\tau$ and $h \leq \pi/b$ for the standard lattice:

Theorem 2.6

Let $f \in C_0^\infty(\Omega)$ and $L_S(h,p) := L(h,2p,1,0)$ be the standard lattice with $p \geq b/\tau$ and $h \leq \pi/b$. Then

$$\sup_{(\phi,s)\in I\times\mathbb{R}} |S_{L_S} Rf - Rf| \leq \frac{8}{\pi\tau}\,\epsilon_0(f,b) + \|f\|_{L_1}\eta(\tau,b).$$

The standard lattice is not the most efficient lattice suitable for sampling Rf. Lindgren and Rattey [11,16] found that a sampling lattice first suggested by Cormack [2] which needs only half as many samples satisfies the conditions of the theorems. This so-called 'interlaced lattice' contains the sampling points $(\phi_j,s_{j\ell}) = \left[\frac{2\pi j}{2p},\ 2h\ell+h\delta_j\right]$, $j = 0,\ldots,2p-1$, $\ell \in \mathbb{Z}$ where

$$\delta_j = j \bmod 2 = \begin{cases} 1 & j \text{ odd} \\ 0 & j \text{ even} \end{cases}$$

and is equal to $L(2h,p,2,1) =: L_I(h,p)$. We obtain the following conditions on p and h:

Theorem 2.7

Let $f \in C_0^\infty(\Omega)$ and $L_I(h,p)$ be the sampling lattice $L(2h,p,2,1)$ defined in

(13). If either

$$\pi/(2b) < h \le \pi/b \text{ and } p \ge (b+\max(\pi/h-b,(1-\tau)b))/\tau \tag{16}$$

or

$$h \le \pi/(2b) \text{ and } p \ge b/\tau, \tag{17}$$

then

$$\sup_{(\phi,s)\in I\times\mathbb{R}} |S_{L_I} Rf - Rf| \le \frac{S}{\pi\tau}\, \epsilon_0(\tau,b)\,\|f\|_{L_1}\,\eta(\tau,b).$$

Proof: The sets $K_0(\tau,b) + u$, $u \in L_I'$ are mutually disjoint if and only if the above conditions on p and h are satisfied The estimate follows again from Theorems 2.2 and 2.5.

□

We see that the interlaced lattice has a spacing of 2h between adjacent samples in the s-variable and needs only a slightly increased value of p. If p is even the values $Rf(\phi_j, s_{j\ell})$ for $j \ge p$ are redundant because of (15). In this case the interlaced lattice is nearly twice as efficient as the standard lattice. If p is odd the data for $j \ge p$ are not redundant and we obtain the same data as from the standard lattice.

While the theorem shows that Rf is determined by its values on the interlaced lattice, it is not immediately clear, how good reconstructions of f can be obtained from these data. The remainder of this paper is devoted to this question.

3 THE FILTERED BACKPROJECTION ALGORITHM

In this section we describe one of the standard reconstruction methods, the filtered backprojection algorithm [15] and its discrete implementation for data sampled on an admissible sampling lattice. The algorithm can be

regarded as a computer implementation of eq. (8) which read

$$W_b * f(x) = \int_0^{2\pi} \int_{\mathbb{R}} w_b(x.\theta - s) Rf(\phi,s) ds d\phi, \quad \theta = (\cos\phi, \sin\phi)^T.$$

So the goal is to obtain an accurate reconstruction of $W_b * f$ instead of f itself. The cut-off frequency b controls the accuracy with which $W_b * f$ approximates f. If it is greater or equal to the essential band-width of f, $W_b * f$ will be very close to f. Therefore we usually choose the cut-off frequency equal to the essential band-width. There is also a direct correspondence between the resolution of the image represented by $W_b * f$ and the cut-off frequency b: the smallest details discernible in the image have a size of roughly $2\pi/b$.

Sampling Rf on an admissible sampling lattice $L = L(d,L,M,N)$ permits to discretize the integrals by means of the trapezoidal rule:

$$W_b * f(x) \simeq \frac{2\pi d}{LM} \sum_{j=0}^{LM-1} \sum_{\ell \in Z} w_b(x.\theta_j - s_{j\ell}) Rf(\phi_j, s_{j\ell}) \tag{18}$$

$$s_{j\ell} = d(\ell + \delta_j/M), \quad \theta_j = (\cos\phi_j, \sin\phi_j)^T, \quad \phi_j = \frac{2\pi j}{LM}$$

where we have used (13). A computer implementation based on (18) demands the computation of the discretized convolution integral

$$Q_j(\phi_j, x.\theta_j) = d \sum_{\ell \in Z} w_b(x.\theta_j - s_{j\ell}) Rf(\phi_j, s_{j\ell})$$

for all points x where the image is computed and all j, $0 \le j \le LM - 1$. Since this would be computationally too demanding, we compute only $Q_j(\phi_j, \Pi k)$, with $\Pi > 0$, $k \in Z$, $-\frac{1}{\Pi} \le k \le \frac{1}{\Pi}$ and obtain an approximation for $Q_j(\phi_j, x.\theta_j)$ by linear interpolation with stepsize Π. This means that we approximate $Q_j(\phi_j, x.\theta_j)$ by

$$I_\Pi Q_j(\phi_j, x.\theta_j) := \sum_k B_\Pi(x.\theta_j - \Pi k) Q_j(\phi_j, \Pi k)$$

where B_Π is given by

$$B_\Pi(s) = \begin{cases} 1 - |s/\Pi| & \text{if } |s| \le \Pi \\ 0 & \text{otherwise} \end{cases}.$$

In the following we will always assume that H is chosen such that

$$\pi/b \geq H = \frac{d}{mM} \text{ with } m \in \mathbb{N}. \tag{19}$$

A short computation shows that because of $d/(HM) \in \mathbb{N}$

$$I_H Q_j(\phi_j, x.\theta_j) = d \sum_{\ell \in Z} (I_H w_b)(x.\theta_j - s_{j\ell}) Rf(\phi_j, s_{j\ell}).$$

This means, that the effect of the interpolation can be expressed by replacing w_b with the interpolated kernel

$$I_H w_b(s) = \sum_{k \in Z} B_H(s - Hk) w_b(Hk).$$

Its Fourier transform is given by

$$\begin{aligned}(I_H w_b)\hat{}(\sigma) &= (2\pi)^{1/2} H^{-1} \hat{B}_H(\sigma) \sum_{\ell \in Z} \hat{w}_b\left[\sigma - \frac{2\pi\ell}{H}\right] \\ &= \operatorname{sinc}^2(H\sigma/2) \sum_{\ell \in Z} \hat{w}_b\left[\sigma - \frac{2\pi\ell}{H}\right]\end{aligned}$$

where $\operatorname{sinc}(s) = \frac{\sin s}{s}$. Hence $I_H w_b$ is not bandlimited. For $|\sigma| < \frac{\pi}{H}$ we have

$$(I_H w_b)\hat{}(\sigma) = \operatorname{sinc}^2(H\sigma/2)\hat{w}_b(\sigma)$$

since we assumed that $b \leq \pi/H$. Therefore we can split $I_H w_b$ in a bandlimited part w_{LO} with band-width b and Fourier transform $\hat{w}_{LO}(\sigma) = \operatorname{sinc}^2(H\sigma/2)\hat{w}_b(\sigma)$, and a high-frequency part w_{HI} with

$$\hat{w}_{HI}(\sigma) = \operatorname{sinc}^2(H\sigma/2) - \sum_{\ell \neq 0} \hat{w}_b\left[\sigma - \frac{2\pi\ell}{H}\right] = \begin{cases} 0 & |\sigma| < \pi/H \\ (I_H w_b)\hat{}(\sigma) & |\sigma| \geq \pi/H. \end{cases}$$

This means that the algorithm computes the function

$$\begin{aligned} f_R(x) &:= \frac{2\pi d}{LM} \sum_{j=0}^{LM-1} \sum_{\ell \in Z} (I_H w_b)(x.\theta_j - s_{j\ell}) Rf(\phi_j, s_{j\ell}) \\ &= \frac{2\pi d}{LM} \sum_{j=0}^{LM-1} \sum_{\ell \in Z} (w_{LO} + w_{HI})(x.\theta_j - s_{j\ell}) Rf(\phi_j, s_{j\ell}). \end{aligned} \tag{20}$$

4 ERROR ESTIMATES

From equation (20) we see that the filtered backprojection algorithm computes

a function $f_R(x)$ which approximates $W_b * f(x)$ and can be decomposed as $f_R(x) = f_{LO}(x) + f_{HI}(x)$ where

$$f_{LO}(x) = \frac{2\pi d}{LM} \sum_{j=0}^{LM-1} \sum_{\ell \in Z} w_{LO}(x.\theta_j - s_{j\ell}) Rf(\phi_j, s_{j\ell}) \tag{21}$$

and $f_{HI}(x)$ is obtained by replacing w_{LO} by w_{HI}. We will show that f_{LO} coincides with $W_b * f$ up to a small error and that $|f_{HI}(x)|$ can be kept small by a proper implementation of the algorithm. From Theorem 1.2 we see that

$$\int_0^{2\pi} \int_R w_{LO}(x.\theta - s) Rf(\phi, s) ds \, d\phi = W_{LO} * f(x)$$

where $\hat{W}_{LO}(\xi) = 2\sqrt{2\pi}|\xi|^{-1} \hat{w}_{LO}(|\xi|)$. Hence

$$W_{LO}(x) = (G_H * W_b)(x) \tag{22}$$

with the additional low-pass filter G_H given by

$$\hat{G}_H(\xi) = \begin{cases} (2\pi)^{-1} \operatorname{sinc}^2(H|\xi|/2) & |\xi| \le b \\ 0 & |\xi| > b \end{cases}$$

compare [12, Theorem V.1.2]. Hence $f_{LO}(x)$ can be regarded as approximattion for $W_{LO} * f(x) = G_H * W_b * f(x)$. The discrepancy $G_H * W_b * f(x) - W_b * f(x)$ can be made arbitrarily small by choosing a sufficiently small H.

We can use the following slightly modified result of Kruse [9, Theorem 6.1] to estimate the error $|f_{LO}(.) - G_H * W_b * f(x)|$:

Theorem 4.1

Let $f \in C_0^\infty(\Omega)$, $g(\phi, s) := Rf(\phi, s)$, W_b, w_b as in Theorem 1.2, $L = L(d, L, M, N)$ an admissible sampling lattice for $T \times R$ as given in (13), and $K \subset Z \times R$ such that the sets $K + u, u \in L'$ are mutually disjoint. Then

$$\frac{2\pi d}{LM} \sum_{j=0}^{LM-1} \sum_{\ell \in Z} w_b(x.\theta_j - s_{j\ell}) g(\phi_j, s_{j\ell}) = W_b * f(x) + e_1(x) + e_2(x) \tag{23}$$

where

$$|e_1(x)| \leq \frac{b}{2\pi} \sum_{k\in Z} \int_{(k,\sigma)\notin K} |\hat{g}(k,\sigma)|\,d\sigma$$

$$|e_2(x)| \leq \frac{bd}{4\pi LM} \sum_{v\in L} |g(v)| \sum_{k\in Z} \int_{(k,\sigma)\notin K,\,|\sigma|\leq b} |J_k(\sigma|x|)|\,d\sigma$$

with J_k the Bessel function of the first kind of order k.

Applying Theorem 4.1 with W_b and w_b replaced by $G_H * W_b$ and w_{L0}, respectively, yields the desired estimate for f_{L0}:

Corollary 4.2

Let f, L,K as in Theorem 4.1, f_{L0} as in (21) and W_{L0} as in (22). Then

$$f_{L0}(x) = W_{L0} * f(x) + e_1 + e_2$$

where e_1 and e_2 satisfy the estimates of Theorem 4.1.

The sampling conditions of Theorem 4.1 and the estimate for e_1 correspond directly to the hypothesis and estimate of the sampling Theorem 2.2. This indicates that the filtered backprojection algorithm is ideally suited to exploit the advantages of minimal sampling lattices. The proof of Theorem 4.1 given below employs Lemma 2.3 in a crucial way, as was done in the proof of the sampling theorem.

Proof of Theorem 4.1: For $x \in \Omega$ define q_x: $T \times R \to R$,

$q_x(\phi,s) = w_b\left[x.\begin{bmatrix}\cos\phi\\ -\sin\phi\end{bmatrix} + s\right]$. We have

$$W_b * f(x) = \int_{T\times R} q_x(-y)g(y)\,dy$$

$$= \sum_{k\in Z} \int_R \hat{q}_x(k,\sigma)\hat{g}(k,\sigma)\,d\sigma \qquad (24)$$

and (see [9, Lemma 6.1])

$$|\hat{q}_x(k,\sigma)| = (4\pi)^{-1}|\sigma\psi(\sigma/b)J_k(\sigma|x|)| \qquad (25)$$

$$\leq (4\pi)^{-1}b|J_k(\sigma|x|)| \qquad (26)$$

$$\leq (4\pi)^{-1}b. \qquad (27)$$

Observing that $c_L = d/(LM)$ the lefthand-side of (23) can be written as

$$2\pi c_L \sum_{v\in L} q_x(-v)g(v).$$

Applying Lemma 2.3 gives together with (24):

$$|e(x)| := \left|2\pi c_L \sum_{v\in L} q_x(-v)g(v) - W_b * f(x)\right|$$

$$= \left|\sum_{k\in Z} \int_R \hat{q}_x(k,\sigma) \sum_{u\in L', u\neq 0} \hat{g}((k,\sigma)-u)d\sigma\right|. \qquad (28)$$

Let χ_K denote the characteristic function of K. Since the sets $K+u$, $u \in L'$ are mutually disjoint, we obtain

$|e(x)|$

$$\leq \sup_{(k,\sigma)\in K}(|\hat{q}_x(k,\sigma)|) \sum_{k\in Z} \int_{(k,\sigma)\notin K} |\hat{g}(k,\sigma)|d\sigma$$

$$+ \sum_{k\in Z} \int_R (1-\chi_K(k,\sigma))|\hat{q}_x(k,\sigma)| \left|c_L \sum_{v\in L} g(v)e^{-i(k,\sigma)\cdot v} - \hat{g}(k,\sigma)\right| d\sigma$$

where we have used the Poisson summation formula (12) for the second term. Hence

$$|e(x)| \leq 2 \sup_{(k,\sigma)\in Z\times R} (|\hat{q}_x(k,\sigma)|) \sum_{k\in Z} \int_{(k,\sigma)\notin K} |\hat{g}(k,\sigma)|d\sigma$$

$$+ c_L \sum_{v\in L} |g(v)| \sum_{k\in Z} \int_{(k,\sigma)\notin K} |\hat{q}_x(k,\sigma)|d\sigma.$$

Now the estimate for e_1 is obtained by applying (27) to the first term on the righthandside. The estimate for e_2 results from observing that $\hat{q}_x(k,\sigma) = 0$ for $|\sigma| > b$ and using (26) for the second term.

□

The results derived so far provide the following procedure for choosing an appropriate sampling lattice and obtain accurate reconstructions: First use

a-priori information about f to determine the essential bandwidth of f which should also be chosen as cut-off frequency for the filter W_b. If the cut-off frequency is too small, $W_b * f$ will not be a good approximation for f and if it is bigger than the essential bandwidth, the reconstruction error e_2 will be large. Otherwise e_2 is not very critical, as we will see below.

Then choose the sampling lattice L so that the translated sets K+u, $u \in L'$ with $K \supseteq K_0(\tau,b)$ are mutually disjoint. According to the results of section 2 this guarantees that Rf is properly sampled and leads according to Corollary 4.2 to a small reconstruction error e_1.

The application of Corollary 4.2 to the standard and interlaced lattices is as follows: As we have seen in Theorems 2.6 and 2.7 the required disjointness of the sets $K_0(\tau,b) + u$, $u \in L'$ translates into the sampling conditions $p \geq b/\tau$, $h \leq \pi/b$ for the standard lattice and into the conditions (16), (17) for the interlaced lattice. Hence the parameters p and h have to be chosen accordingly. For the standard lattice the disjointness of the sets $K_0(\tau,b) + u$, $u \in L_s'$ implies that the sets $K_1(\tau,b) + u$, $u \in L_s'$ with K_1 the rectangle

$$K_1(\tau,b) = \{(k,\sigma) \in Z \times R: |k| \leq b/\tau, |\sigma| \leq b\}$$

are also mutually disjoint. Therefore we can take $K_1(\tau,b)$ for the set K in Corollary 4.2. In case of the interlaced lattice we have to choose $K = K_0(\tau,b)$ to achieve optimally sparse sampling. This is of little consequence for e_1 but increases e_2. To see this we reformulate the estimate for e_2 and bring out its dependence on $|x|$: Let K_0, K_1 be as above, $\kappa_0 := (1-\tau)b/\tau$, $\kappa_1 := b/\tau$ and $\beta := \frac{1}{3}(1-(\tau|x|)^2)^{3/2}$. Using the estimates of [12, p. 65] we obtain the estimate (i = 0,1):

$$\sum_{k\in Z}\int_{(k,\sigma)\notin K_i(\theta,b),\,|\sigma|<b}|J_k(\sigma|x|)|\,d\sigma \le \sqrt{\frac{2}{\pi}}\,\frac{2\tau}{(1-\tau^2|x|^2)^{1/4}}\sum_{k>\kappa_i}\sqrt{k}\;e^{-\beta k}. \quad (29)$$

According to this estimate we expect for τ close to 1 a strong growth of the error e_2 near the boundary of Ω. Since $\kappa_1 > \kappa_0$ the error should be more pronounced in the case of interlaced sampling. A comparison of the reconstructions shown in Figures 3 and 6 of the next section shows that these expectations are justified. With increasing b the error decays exponentially.

The performance of the filtered backprojection algorithm using the standard lattice has already been studied in [10, 12, 13, 20] There the errors due to the discrete convolution and the numerical integration with respect to ϕ have been treated separately. It turns out that for the standard lattice both numerical integrations yield accurate results when the sampling conditions are satisfied. This is no longer the case when the interlaced lattice is used. Then the discrete convolutions are highly inaccurate because of the stepsize 2h being too large. The results derived in [5] show that these errors cancel out during the subsequent numerical integration over T. But these cancellations may be disturbed by the interpolation step inserted in between the two integrations. This provides an intuitive explanation for the fact that, as we will see below, the interpolation is harmless for standard sampling but critical for interlaced sampling.

Having seen that f_{L0} represents the picture we want to have it remains to clarify under which conditions $|f_{III}(x)|$ will be small. We obtain

Theorem 4.3

Let $f \in C_0^\infty(\Omega)$, ψ as in Theorem 1.2, and $L(d,L,M,N)$ be an admissible

sampling lattice. Then

$$|f_{HI}(x)| \le \frac{1}{2} \sup_{\phi\in T} \int_{-b}^{b} \sin^2(H\sigma/2)\,|\sigma\psi(\sigma/b)| \sum_{\ell\in Z} \left|\hat{f}\left[\left[\sigma - \frac{2\pi\ell}{d}\right]\theta\right]\right| d\sigma. \qquad (30)$$

Proof: With $s_{j\ell} = d(\ell+\delta_j/M)$ (see (13)) we obtain

$$f_{HI}(x) = \frac{2\pi d}{LM} \sum_{j=0}^{LM-1} \sum_{\ell\in Z} w_{HI}(x.\theta_j - s_{j\ell}) Rf(\phi_j, s_{j\ell})\, d\sigma$$

$$= \frac{\sqrt{2\pi}\, d}{LM} \sum_{j=0}^{LM-1} \int_R \hat{w}_{HI}(\sigma) e^{i\sigma x.\theta_j} \sum_{\ell\in Z} e^{i\sigma d(\ell+\delta_j/M)} Rf(\phi_j, d(\ell+\delta_j/M))\, d\sigma$$

$$= \frac{2\pi}{LM} \sum_{j=0}^{LM-1} \int_R \hat{w}_{HI}(\sigma) e^{i\sigma x.\theta_j} \sum_{\ell\in Z} (Rf)\hat{}^2\left[\phi_j, \sigma - \frac{2\pi\ell}{d}\right] e^{-i2\pi\ell\delta_j/M} d\sigma$$

where we have used Poisson's summation formula (12) with $n_1 = 0, n_2 = 1$ and $L = \{d\ell, \ell \in Z\}$. With Theorem 1.1 and the assumptions (19) we obtain the estimate

$$|f_{HI}(x)|$$

$$\le (2\pi)^{3/2} \sup_{\phi\in T} \int_R |\hat{w}_{HI}(\sigma)| \sum_{\ell\in Z} \left|\hat{f}\left[\left[\sigma - \frac{2\pi\ell}{d}\right]\theta\right]\right| d\sigma$$

$$= (2\pi)^{3/2} \sup_{\phi\in T} \int_R \mathrm{sinc}^2(H\sigma/2) \sum_{k\neq 0} \hat{w}_b\left[\sigma - \frac{2\pi k}{H}\right] \sum_{\ell\in Z} \left|\hat{f}\left[\left[\sigma - \frac{2\pi\ell}{d}\right]\theta\right]\right| d\sigma$$

$$= (2\pi)^{3/2} \sup_{\phi\in T} \sum_{k\neq 0} \int_{-b}^{b} \frac{\sin^2(H\sigma/2)}{\left[\frac{H\sigma}{2} + \pi k\right]^2} \hat{w}_b(\sigma) \sum_{\ell\in Z} \left|\hat{f}\left[\left[\sigma + \frac{2\pi(kMm-\ell)}{d}\right]\theta\right]\right| d\sigma$$

$$\le \frac{1}{2} \sup_{\phi\in T} \int_{-b}^{b} \sin^2(H\sigma/2)\,|\sigma\psi(\sigma/b)| \sum_{\ell\in Z} \left|\hat{f}\left[\left[\sigma - \frac{2\pi\ell}{d}\right]\theta\right]\right| d\sigma.$$

□

The interpretation of this estimate is more involved than the one of Corollary 4.2. As an example we discuss the case of standard sampling with

$d = h$, and f essentially bandlimited with bandwidth $b = \pi/h$, which is also the cut-off frequency for the filter w_b. Hence $2\pi/d = 2b$ and therefore the terms with $\ell \neq 0$ in $\Sigma_{\ell \in Z} \left| \hat{f}\left[\left[\sigma - \frac{2\pi\ell}{d}\right]\theta\right] \right|$ are negligibly small. The term with $\ell = 0$ is equal to $|\hat{f}(\sigma\theta)|$. In most applications we have $f(x) \geq 0$. This means that $|\hat{f}(\xi)|$ has a sharply peaked maximum at the origin. For $|\xi|$ close to the cut-off frequency b, $|\hat{f}(\xi)|$ is usually very small. In such a case the integral in (30) will be small since $\sin^2(H\sigma/2)$ is small for $|\sigma| << 1/H$. So with standard sampling it is usually safe to choose $H = h$. An example for this is the Shepp-Logan phantom which we use for our numerical simulations in the next section.

If on the other hand $|\hat{f}(\xi)|$ is not small for $|\sigma|$ close to b, the interpolation stepsize H has to be chosen considerably smaller than π/b. Then $\sin^2(H\sigma/2)$ is small for all values $-b \leq \sigma \leq b$. As an example for this case we will perform numerical tests with $f(x) = J_1(b|x-x_0|)/|x-x_0|$. The Fourier transform of this function is constant for $|\xi| \leq b$ and vanishes for $|\xi| > b$.

5 NUMERICAL RESULTS

In this section we present numerical tests for the theory derived so far. We will concentrate on the standard and the interlaced lattices and will see that the numerical results can be understood in detail using the theorems of the last section. The theoretical results will in particular enable us to remove the numerical difficulties with reconstructions from the interlaced lattice reported in [9].

Figure 2 shows the first object we used for our tests. It is a mathematical phantom due to Shepp and Logan [18] and simulates a cross-section of a human head. Here f is given by a linear combination of characteristic

functions of ellipses. The density values of the ellipses are the values of Rowland [17] multiplied by 900 and were also used by Kruse [9]. The displayed values are those between 1 (black) and 75 (white). The biggest value occuring in the picture is 900. All pictures are displayed on a 256 by 256 grid. As essential bandwidth we choose $b = 128\pi \simeq 402$.

Figure 3 shows a reconstruction made using the standard sampling lattice $L_S(h,p)$ with parameters

$$h = \pi/b = 1/128 \quad p = 404 \quad H = h$$
$$\psi(\sigma) = \mathrm{sinc}(\sigma\pi/2)\chi_{[-1,1]}(\sigma) \tag{31}$$

where $\chi_{[-1,1]}$ denotes the characteristic function of the interval $[-1,1]$. The sampling conditions of Theorem 2.6 are satisfied with $\tau = b/p \simeq 0.995$. From the discussions following Corollary 4.2 and Theorem 4.3 we expect a good reconstruction. This is obviously the case.

In Figure 4 we set $H = h/2$ which introduces a high-frequency pattern in the interior of the object. According to Theorem 4.3 a smaller value of H should lead to a smaller error f_{HI}, hence to a more accurate reconstruction of $W_b * f$. This is indeed the case. The high-frequency pattern stems from the jump discontinuity of the filter function ψ which causes a jump discontinuity of $\hat{W}_b(\xi)$ at $|\xi| = b$. Therefore for this choice of ψ the high-frequency pattern, though undesired, is a true feature of $W_b * f$. It can be shown that if $H = h = \pi/b$ as in the previous picture, the error term f_{HI} removes the discontinuity in the Fourier transform of the reconstructed function and thus causes the high-frequency pattern to disappear.

If we reconstruct using the interlaced lattice $L_I(h,p) = L(2h,p,2,1)$ with the same parameters (31) as in Figure 3, the sampling condition (16) is satisfied with $\tau = 2b/(p+b) \simeq 0.998$. But now the error f_{HI} caused by the interpolation becomes critical. The estimate of Theorem 4.3 reads

$$|f_{HI}(x)| \le \frac{1}{2} \sup_{\theta \in T} \int_{-b}^{b} \sin^2(H\sigma/2)\,|\sigma\psi(\sigma/b)| \sum_{\ell \in Z} \left|\hat{f}\left[\left[\sigma - \frac{\pi\ell}{h}\right]\theta\right]\right| d\sigma. \qquad (32)$$

Since $|\hat{f}(\xi)|$ is small for $|\xi| > b$ and because of $h \le \pi/b$, only the terms with $|\ell| \le 1$ in the sum on the righthandside are not negligibly small. As we have already discussed the term with $\ell = 0$ will not be critical for an object like this one. For $|\ell| = 1$ however, $|\hat{f}((\sigma - \pi\ell/h)\theta)| = |\hat{f}((\sigma \pm b)\theta)|$. This means that the integrand on the righthandside of (32) might assume large values for $|\sigma|$ close to b, since $\sin^2(H\sigma/2)$ is not small for these values of σ if $H \simeq \pi/b$. So we have to expect a considerable reconstruction error in this case. The picture of Figure 5 shows this error which makes the reconstruction totally useless. The picture is essentially the same as the one of [9, Fig. 8.2.(c)]. But our estimate (32) also suggests the following three ways to remedy the problem:

1. Choose $H \ll \pi/b$ so that $\sin^2(H\sigma/2) \ll 1$ for $|\sigma| \le |b|$. While this method will always work, the next two possibilities are only suitable if $|\hat{f}(\xi)|$ is peaked around the origin.
2. Choose h smaller than π/b so that $|\sigma \pm \pi/h|$ stays away from 0 if $|\sigma| \le b$. If $|\hat{f}|$ is peaked around the origin, even a small decrease in h should give considerable improvement, see Figure 7 below. On the other hand decreasing h might result in a violation of the condition that the translated sets $K_0(\tau,b) + u$, $u \in L'$ are mutually disjoint and thus introduce new artifacts.
3. Choose the filter ψ such that $|\psi(\sigma/b)| \ll 1$ for $|\sigma|$ close to b. Then the integrand in (32) remains small also for these critical values of $|\sigma|$.

These three options are tested in the following pictures.

In Figure 6 we put $H = \pi/(16b)$ so that $\sin^2(H\sigma/2) < 0.01$ for $|\sigma| \le b$. In

the interior the picture is virtually identical to the reconstruction with the standard lattice and small N shown in Figure 4. In particular the high-frequency patterns due to the discontinuity of ψ coincide. Aside from the white ring-type artifact at the boundary of the picture we have a good reconstruction of $W_b * f$. This corroborates the assertion that the large error in the previous reconstruction is caused by the interpolation. The artifact at the boundary can be traced to the error e_2 of Corollary 4.2. The estimate (29) shows that this error term grows strongly for $|x|$ approaching 1. This explains that the artifact appears at the boundary of the unit circle. Furthermore according to (29) the artifact should vanish if p is slightly increased, corresponding to a smaller value of τ. Figure 8 shows that this is indeed the case.

The second possibility of improvement, namely to decrease h, was tried in Figure 7. We used again the parameters (31) but set h = 1/150. Hence $|\sigma \pm \pi/h| \geq \pi/h - b = 22\pi$ if $|\sigma| \leq b = 128\pi$. Hence the peak of $|\hat{f}(\xi)|$ near the origin is avoided, and as expected the reconstruction error is strongly reduced. In the interior of the picture there are still some disturbances. A comparison with the next picture suggests that these artifacts result from the overlapping of the sets $K_0(\tau,b) + u$, $u \in L'$ caused by decreasing h without increasing p. In Figure 8 we used again h = 1/150 but increased p to 470, so that this overlap is greatly reduced. In addition we can assume a smaller value of τ in the error estimate (29). As expected from our reasoning above the disturbances in the interior are removed as well as is the white ring at the boundary. The high-frequency structures in the interior of the phantom are very similar in Figures 7 and 8 but clearly different from the patterns visible in Figures 4 and 6. The latter ones are, as we have seen no reconstruction errors but belong to $W_b * f$. The difference stems from the

interpolation and can be removed by reducing H, e.g. setting $H = \pi/(3b)$. This was essentially done by Kruse in [9, Fig. 8.2.(d)]. In his picture, however, the high-frequency pattern appears as a low-frequency disturbance which was not completely removed even by strong smoothing, see [9, Fig. 8.2.(e)]. The reason for this is that the picture is displayed on a 160 by 160 instead of a 256 by 256 grid, which causes aliasing.

The third method mentioned above, i.e. choosing the filter ψ such that $|\hat{\psi}(\sigma/b)| \ll 1$ for $|\sigma|$ close to b yielded the best result for this phantom. In Figure 9 we used the filter $\hat{\psi}(\sigma) = \cos(\sigma\pi/2)\chi_{[-1,1]}(\sigma)$ and the other parameters as in (31). We obtain a good reconstruction without artifacts and of comparable quality as the reconstruction of Figure 3 where the standard lattice was used. The small values of $\hat{\psi}(\sigma/b)$ for $|\sigma|$ close to b seem to remove the ring-like artifact at the boundary. The continuity of ψ removes the undesired high-frequency patterns in the interior. The price for this is a slight loss in resolution. So this method will only work if $|\hat{f}(\xi)|$ is already small for $|\xi|$ close to b.

In summary, the numerical tests show that for this object good reconstructions using the interlaced lattice are possible if the algorithm is implemented in the right way.

Our second test object is the strictly bandlimited function

$$f(x) = \frac{J_1(b_0|x-x_0|)}{|x-x_0|}$$

with $b_0 = 20\pi$ and $x_0 = (-0.4,0.7)^T$. The main difference between the two phantoms can be seen from the behaviour of their Fourier transforms. The Fourier transform of the Shepp-Logan phantom is peaked around the origin and relatively small for $|\xi|$ close to b, but has no compact support. The Fourier transform of the function above is constant for $|\xi| \leq 20\pi$ and vanishes for

$|\xi| > 20\pi$. Since $\epsilon_0(f,b) = 0$ for $b \geq 20\pi$ we should be able to obtain very accurate reconstructions from both lattices. We will see that this is indeed possible. To achieve high accuracy we have to take into account that f has no compact support and that the values of

$$Rf(\phi,s) = 2\ \mathrm{sinc}(20\pi(s-x_0\cdot\theta))$$

for $|s| > 1$ are not small enough to be completely neglected. For our purposes it will be sufficient to use the data for $|s| \leq 6$ for reconstructing f inside the unit circle. As measure for the accuracy of the reconstructions we use a discrete relative L_2-error

$$L_2E := \left[\sum_{i,j}(f(x_{ij}) - f_R(x_{ij}))^2\right]^{1/2}\left[\sum_{i,j}(f(x_{ij}))^2\right]^{-1/2}$$

where the x_{ij} are the points at which the reconstruction is computed.

Phantoms of this kind have been used by Rowland [17], Natterer [12] and Kruse [9] for reconstructions using the standard lattice and by Kruse also for reconstructing with the interlaced lattice. All three authors choose $x_0 = 0$. This choice is not appropriate for our purpose since it makes the object radially symmetric. This means that $(Rf)\hat{}(k,\sigma) = 0$ for $k \neq 0$ which is too great a simplification. Rowland demonstrated that it is necessary to choose $H \ll \pi/b_0$ even when using the standard lattice. The reason for this can be seen from Corollary 4.2. The additional filter G_H is now critical because $|\hat{f}(\xi)|$ is not small for $|\xi|$ close to b. While a choice of H close to π/b_0 leads to discrete relative L_2-errors around 0.4, much more accurate results are possible with H small. For our tests we used the following parameters:

$$b = 20\pi \quad p = 70 \quad H = \pi/(16b)$$

$$\psi(\sigma) = \chi_{[-1,1]}(\sigma).$$

The reconstruction using the standard lattice with $h = \pi/b = 1/20$ yields an excellent result with $L_2E = 0.0085$. Using the interlaced lattice with the

same parameters we obtain an equally good result with $L_2E = 0.0093$. But in order to achieve this accuracy it is crucial to satisfy the sampling condition, i.e. to make sure that the sets $K_0(\tau,b) + u$, $u \in L_I^1$, $\tau = 2b/(p+b)$ are mutually disjoint. Increasing h to 1/18 increases L_2E to 0.56. Even decreasing h increases the error as the following table shows:

1/h	L_2E
16	0.74
18	0.56
20	0.0093
22	0.042
24	0.22
30	0.57
36	0.45
38	0.33
40	0.0087
42	0.0072
50	0.0070

The explanation for this behaviour is that the sets $K_0(\tau,b) + u$, $u \in L_I^1$ are mutually disjoint only for $h = 1/20$ and $h \leq 1/40$. Hence satisfying one of the sampling conditions (16), (17) is absolutely crucial in this case, where the Fourier transform of the object is not small for frequencies close to b.

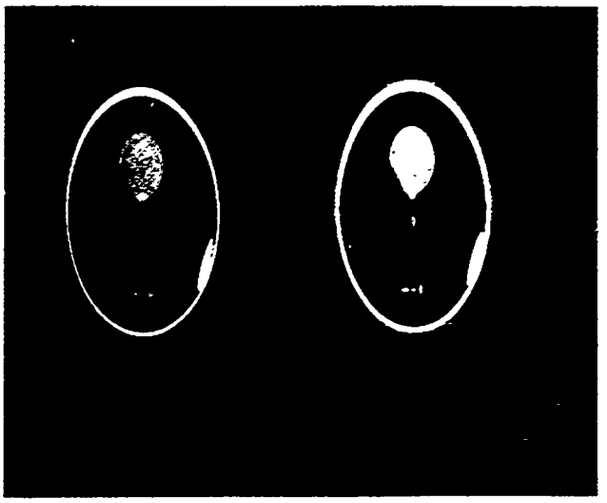


Figure 2: left Figure 3: right

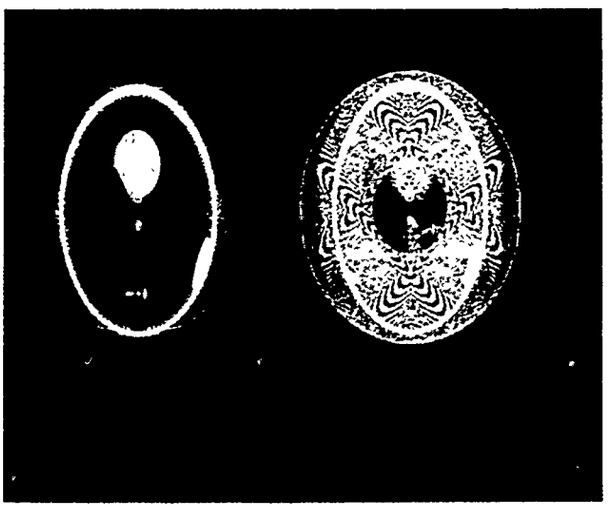


Figure 4: left Figure 5: right

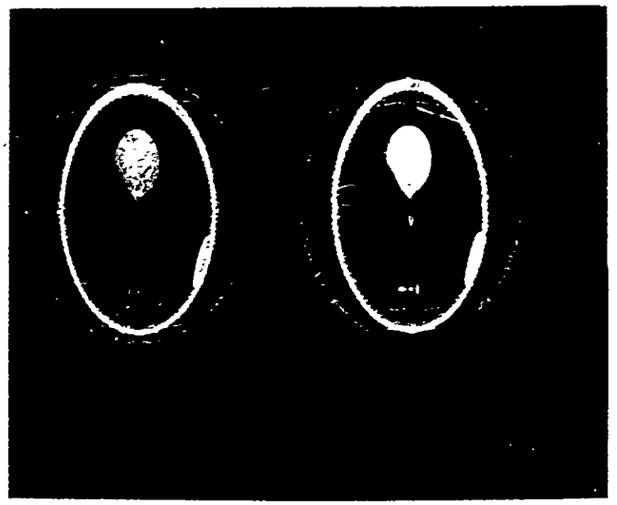


Figure 6: left Figure 7: right

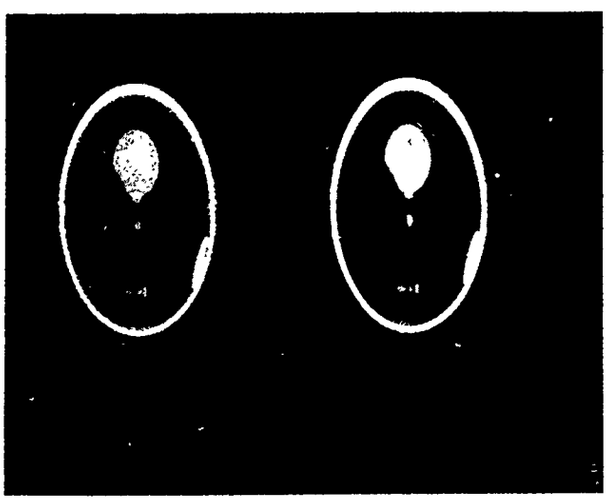


Figure 8: left Figure 9: right

ACKNOWLEDGEMENT

Part of this work was supported by a Feodor-Lynen-Fellowship of the Alexander von Humboldt Foundation while the author was on leave at Oregon State University, Department of Mathematics, Corvallis, OR 97331, U.S.A.

REFERENCES

[1] R. N. Bracewell, Strip integration in radio astronomy, Aus. J Phys., 9, 198-217, 1956.

[2] A M. Cormack, Sampling the Radon transform with beams of finite width, Phys. Med. Biol., 23, 1141-1148, 1978.

[3] R. A. Crowther, D. J. De Rosier and A. Klug, The reconstruction of a three-dimensional structure from projections and its application to electron microscopy, Proc. R. Soc. London Ser. A, 317, 319-340, 1970.

[4] S. R. Deans, The Radon Transform and some of its Applications, Wiley, 1983.

[5] A. Faridani, Sampling and resolution in diffraction tomography II: An error analysis of the filtered backpropagation algorithm, accepted for publication in SIAM J. Appl. Math.

[6] S. Helgason, The Radon Transform, Birkhäuser, 1980.

[7] G. T. Herman (ed.), Image Reconstruction from Projections, Springer, 1979.

[8] G. T. Herman, Image Reconstruction from Projections. The Fundamentals of Computerized Tomography, Academic Press, 1980.

[9] H. Kruse, Resolution of reconstruction methods in computerized tomography, SIAM J. Sci. Stat. Comput., 10, 447-474, 1989.

[10] R. M. Lewitt, R. H. T. Bates and T. M. Peters, Image reconstruction from projections II: Modified back-projection methods, Optik, 50, 180-205, 1978.

[11] A. G. Lindgren and P. A. Rattey, The inverse discrete Radon transform with applications to tomographic imaging using projection data, Advances in Electronics and Electron Physics, 56, 359-410, 1981.

[12] F. Natterer, The Mathematics of Computerized Tomography, Wiley, New York, 1986.

[13] F. Natterer and A. Faridani, Basic algorithms in tomography, to appear in Signal Processing, IMA Vols. in Math. and Appl., Springer, New York.

[14] D. P. Petersen and D. Middleton, Sampling and reconstruction of wave-number-limited functions in N-dimensional euclidean space, Inf. Control, 5, 279-323, 1962.

[15] G. N. Ramachandran and A. V. Lakshminarayanan, Three-dimensional reconstruction from radiographs and electron micrographs: application of convolutions instead of Fourier transforms, Proc. Nat. Acad. Sci., 68, 2236-2240, 1971.

[16] P. A. Rattey and A. G. Lindgren, Sampling the 2-D Radon transform, IEEE Trans. Acoust. Speech Signal Processing, ASSP-29, 994-1002, 1981.

[17] S. W. Rowland, Computer implementation of image reconstruction formulas, in G. T. Herman (ed.), Image Reconstruction from Projections, Springer, 1979.

[18] L. A. Shepp and B. F. Logan, The Fourier reconstruction of a head section, IEEE Trans. Nucl. Sci., NS-21, 21-43, 1974.

[19] K. T. Smith, Reconstruction formulas in computed tomography, Proc. of Symposia in Applied Mathematics, 27, 7-23, 1983.

[20] O. J. Tretiak, The point spread function for the convolution algorithm, in R. Gordon (ed.), Image processing for 2-D and 3-D reconstructions from projections: Theory and practice in medicine and the physical sciences. Digest of technical papers, Stanford, California, August 4-7, 1975. Sponsored by the Optical Society of America.

A. Faridani,
Westfaelische Wilhelms-Universitaet
Institut fuer Numerische und instrumentelle Mathematik
Einsteinstr. 62, D-4400 Muenster
FRG.

R.F. MILLAR

An inverse moving boundary problem for Laplace's equation

I INTRODUCTION

Suppose that one or more incompressible, viscous fluids flow slowly in the narrow gap separating two plane parallel plates. Such a configuration is known as a Hele-Shaw cell. For steady flow, the problem is essentially two-dimensional and, in the absence of gravitational effects, the average of the fluid velocity across the stratum is proportional to the pressure gradient ([11, §330]). Since Darcy's law for flow through porous media is of the same form, two-dimensional flow through a porous medium may be modelled by flow in a Hele-Shaw cell. Because Hele-Shaw flow is more easily accessible to observation and experiment than flow in a porous medium, this connection has been utilised frequently to examine phenomena such as the interfacial instabilities known to occur in both; see, for example, [19], and the recent review articles [1], [5] and [18].

In a typical Hele-Shaw problem, one viscous fluid displaces a second that has different viscosity or density, and interest centres on the evolution of the interface between them. Instabilities in the interface may arise when the displaced fluid is more viscous than the other liquid. The easiest situation to analyse occurs when one of the fluids has negligible viscosity and density. In these circumstances, one may neglect the motion of this fluid, the pressure of which is taken to be zero. We shall confine attention to this most simple case of single-phase flow and, although the relationship between fluid velocity and pressure gradient is based on the assumption of steady flow, we shall assume its validity in the time-dependent case as well.

We shall also employ generally-accepted boundary conditions on the interface, even though the question of which conditions are physically correct can hardly be considered as settled ([1], [5], [18], [20]).

One may distinguish between direct problems, and inverse (synthesis or control) problems. In the former, the real physical sources, or singularities, of the flow are prescribed in space and time and, starting from some given initial state, one wishes to predict the state for subsequent times.

For an inverse problem, the evolution of the flow is prescribed, and the object is to determine what sources - if, indeed, any - will sustain the motion. Because of the theme of this meeting (but also since the method to be described is better suited to that task than to the direct problem), we shall consider only the inverse problem and, for certain examples, we shall discuss whether or not the prescribed motion of the interface can be realised in practice.

Conformal mapping techniques have provided a powerful means for studying Hele-Shaw problems. For time-dependent (as opposed to steady-state) problems, the usual procedure consists in obtaining a nonlinear differential equation for the time-evolution of the unknown analytic function that maps the unit disc onto the flow region ([1], [6], [7], [8], [9], [15]). It is believed that these methods have not been used to study inverse problems.

In the present work, a different approach is adopted. For the single-phase Hele-Shaw problem, it will be seen that the pressure is the solution to a Cauchy problem for the Laplace equation, with analytic data prescribed on the unknown interface. An explicit representation exists for this solution, analytic in a neighbourhood of the interface ([12]). Moreover, the Cauchy data can be expressed completely in terms of what is

called the Schwarz function S of the interface curve ([2]). Consequently, the pressure can be written explicitly in terms of S and its derivatives alone. Prescription of S as a function of time is thus equivalent to prescribing the evolution of the interface, and the singularities required to maintain the process are readily found. Some of these will lie on one side of the interface in the fluid region; others may lie outside the fluid where the solution corresponds to the analytic continuation of the pressure. (The Schwarz function has arisen in previous analyses of Hele-Shaw problems; see [6], [9], [10], [15], [16] and [17].)

In the following section, the problem is formulated. Then the solution is expressed in terms of the Schwarz function of the interface curve. The effect of surface tension (T) on the interface is included, and some immediate consequences are briefly described. The physical realisability of solutions is discussed. Time-dependent problems are examined under the assumption that $T = 0$. Consideration is given to interior problems, in which the fluid occupies a bounded, simply-connected domain, and to exterior problems in which the complement of the fluid domain is of this form. Specific attention is given to interfaces that are circles, ellipses, and limacons. The stability problem is not addressed. For the interior problem, an interface that is elliptical for all time can be generated by a system of simple sources on the interfocal segment. In the exterior case, an interface that is elliptical for all time can be generated only if the ellipses have constant eccentricity and the pressure becomes unbounded logarithmically at infinity. In both interior and exterior problems for limaçons there are singularities in the finite plane inside and outside the fluid region. For the interior problem, it is possible to generate the solution. In the case of the exterior problem, however, the relevant singularities extend to infinity and it seems

unlikely that this solution could be realised in practice.

In connection with these examples, some consequences of nonvanishing surface tension are investigated briefly. The interior solution for an ellipse still seems to be realisable, but the exterior solution is not. For the limaçon, the mathematical solution does not exist when $T \neq 0$.

Finally, we mention some other problems that can be formulated in the same manner as those described here.

2. FORMULATION OF THE PROBLEM

Let D_t denote a domain in the complex z-plane. Its boundary C_t evolves with time t. An incompressible, viscous fluid fills a region between two closely-spaced, parallel plates, and projects orthogonally onto D_t. Suppose that this Hele-Shaw cell is unbounded and consider an inverse problem in which the evolution of C_t is prescribed. One wishes to determine the sources or sinks, if any, that will produce this behaviour.

It will be assumed that C_t is a simple, closed, analytic curve, oriented positively. The unit normal $\hat{n}$ to C_t will always be drawn out of the bounded domain enclosed by C_t.

If $z \in C_t$, then the Schwarz function S of C_t is defined by $\bar{z} = S(z,t)$, ([2]), so knowledge of S determines C_t. It may be obtained by rewriting the equation $f(x,y;t) = 0$ for C_t in terms of z and $\bar{z}$, and solving for $\bar{z}$. The Schwarz function is analytic near C_t.

Let the angle ψ be defined in terms of the unit positive tangent vector $\hat{t}$ to C_t by $\hat{t} = (\cos\psi, \sin\psi)$. Then $\hat{n} = (\sin\psi, -\cos\psi)$, and the derivative $S_z(z,t) = e^{-i2\psi}$, $z \in C_t$ [2, (7.8)]. We shall define $(S_z(z,t))^{1/2}$ by

$$S_z^{1/2} := e^{-i\psi}, \quad z \in C_t,$$

from which it follows that $dz/ds = S_z^{-1/2}$ on C_t. Thus with $S_z^{1/2}$ is associated

a particular orientation of C_t.

A discussion of the Hele-Shaw equations is given in [19]. If v denotes the fluid velocity in D_t then, if gravitational effects are negligible,

$$v = -\frac{1}{K}\nabla u, \tag{2.1}$$

in which $K := 6\mu/d^2$ where d is the distance between the plates, μ is the coefficient of viscosity of the fluid, and u is the fluid pressure. Incompressibility implies that

$$\Delta u = 0, \qquad z \in D_t, \tag{2.2}$$

except at singularities. At a point on C_t,

$$u = \pm T\kappa, \tag{2.3}$$

and $v.\hat{n} = V_n$, so

$$\frac{\partial u}{\partial n} = -2KV_n; \tag{2.4}$$

here T is the surface tension coefficient, assumed to be non-negative for immiscible fluids, $\kappa := d\psi/ds$ denotes the curvature of C_t, and V_n is the velocity of C_t at z in the direction $\hat{n}$. The sign ambiguity in (2.3) is resolved in the following manner: for interior problems in which C_t contains D_t the upper sign is chosen; for exterior problems we choose the lower sign. (This sign ambiguity is a result of the unique definition of $S_z^{1/2}$ for both interior and exterior problems.)

3. SOLUTION REPRESENTATION

Equations (2.2) to (2.4) define a Cauchy problem for Laplace's equation. Since C_t and the data are analytic, a unique solution exists in a neighbourhood of C_t. It is our intention to determine the singularities of this solution when C_t is prescribed. The goal may be attained by using a representation for the solution.

Suppose that D is a real neighbourhood of an analytic arc C. Let S denote the Schwarz function of C, and suppose that u is harmonic on D. Define $z := x + iy$, $z^* := x - iy$, where x and y may be complex. Assume that C is parametrised by arclength s as $z = z(s)$, and let $s = s(z)$ be the inverse relationship. If $u = v(s)$, $\partial u/\partial n = w(s)$ on C, then ([12]) $u(x,y) = U(z,\bar{z})$, where

$$U(z,z^*) = \frac{1}{2}[v(s(z)) + v(\bar{s}(z^*))] + \frac{1}{2} i \int_{\bar{s}(z^*)}^{s(z)} w(s)ds; \qquad (3.1)$$

here an overbar denotes the conjugate function and the integral is in the complex plane.

If $V(z) := v(s(z))$ and $W(z) := w(s(z))ds/dz$. then (3.1) becomes

$$U(z,z^*) = \frac{1}{2}[V(z) + \bar{V}(S(z^*))] + \frac{1}{2} i \int_{S(z^*)}^{z} W(\zeta)d\zeta, \qquad (3.2)$$

and $z^* = \bar{z}$ gives

$$u(x,y) = \frac{1}{2}[V(z) + \overline{\bar{V}(S(z))}] + \frac{1}{2} i \int_{\overline{S(z)}}^{z} W(\zeta)d\zeta. \qquad (3.3)$$

The quantities κ and V_n in (2.3) and (2.4) may be expressed in terms of S and its derivatives; for, by [2,(7.17)],

$$\kappa = \frac{1}{2} iS_{zz}S_z^{-3/2}$$

$$= -i\frac{\partial}{\partial z}(S_z^{-1/2}),$$

for $z \in C_t$. Also, if now $z(t) := x(t) + iy(t)$, where $(x(t),y(t))$ always lies on C_t, then differentiation of $\overline{z(t)} = S(z(t),t)$ immediately gives $v.\hat{n}$, and

$$V_n(z) = -\frac{1}{2} iS_tS_z^{-1/2}, \qquad z \in C_t. \qquad (3.4)$$

Insertion of the Cauchy data into (3.3) yields

$$u(x,y;t) = \pm \frac{1}{2} T \operatorname{Re}\left\{\frac{iS_{zz}(z,t)}{[S_z(z,t)]^{3/2}}\right\} - \frac{1}{2} K \int_{\overline{S(z,t)}}^{z} S_t(\zeta,t)d\zeta, \qquad (3.5)$$

and the solution is expressed completely in terms of S and its derivatives.

It may be shown that the integral in (3.5) is real. The contour of integration is obtained by continuously deforming the path for a point near C_t into that for the given point z, and will depend on the mapping properties of S. If S is multivalued, this integral will be on the corresponding Riemann surface. It is possible, however, to show that

$$\int_{\overline{S(z,t)}}^{z} S_t(\zeta,t)d\zeta = \int_{z_o}^{z} S_t(\zeta,t)d\zeta + \int_{\bar{z}_o}^{\bar{z}} \bar{S}_t(\zeta,t)d\zeta,$$

in which z_o is any point on C_t. (A similar observation has been made in [4].) Consequently (3.5) may be rewritten as

$$u(x,y;t) = \operatorname{Re}\left\{\pm \frac{1}{2} iT \frac{S_{zz}(z,t)}{[S_z(z,t)]^{3/2}} - K \int_{z_o}^{z} S_t(\zeta,t)d\zeta\right\}, \qquad (3.6)$$

and if a complex potential w is defined by

$$w(z,t) := \pm \frac{1}{2} iT \frac{S_{zz}(z,t)}{[S_z(z,t)]^{3/2}} - K \int_{z_o}^{z} S_t(\zeta,t)d\zeta \qquad (3.7)$$

(to which any purely imaginary function of t may be added), then

$$u(x,y;t) = \operatorname{Re}\, w(z,t).$$

Now, $w_z = u_x - iu_y$, so

$$u_x - iu_y = \pm \frac{1}{2} iTS_z^{1/2}\{S,z\} - KS_t, \qquad (3.8)$$

in which $\{S,z\}$ is the Schwarzian derivative of S. From (3.8) for $z \in C_t$, and the result (7.23') of [2]: $d\kappa/ds = \frac{1}{2}i\{S,z\}$, together with (2.1) and 3.4), we deduce that

$$v.t = \mp \frac{T}{2K} \frac{d\kappa}{ds}, \qquad (3.9)$$

and verify that

$$\mathbf{v}.\hat{\mathbf{n}} = V_n, \tag{3.10}$$

on C_t. Thus, if $T \neq 0$ there is a tangential component of fluid velocity on C_t unless $d\kappa/ds = 0$. This component vanishes if C_t is a circle; but generally there is a discontinuity in tangential velocity across C_t, which corresponds to a layer of vorticity on C_t.

The results (3.9) and (3.10) refer to behaviour on C_t. Suppose, however, that we are interested in fluid velocity at some point P interior to the fluid. Then (3.8) is valid, in which S refers to C_t. Let C' be any simple, closed, analytic curve through P, lying in the domain bounded externally by C_t and having positive orientation. Suppose that $\hat{t}'$ and $\hat{n}'$ are the unit tangent and normal vectors to C' at P. Thus, if $\hat{t}' = (\cos\psi', \sin\psi')$, then $e^{i\psi'} = (S_z')^{-1/2}$, where S' is the Schwarz function of C'; this result depends only on the direction of C' at P. Consequently, at P,

$$\mathbf{v}.(\hat{t}'+i\hat{n}') = \pm \frac{iT}{2K}(S_z')^{-1/2} \frac{\partial^2}{\partial z^2}(S_z^{-1/2}) + \frac{1}{2}(S_z')^{-1/2}S_t, \tag{3.11}$$

in which the result $S_z^{1/2}\{S,z\} = -2\dfrac{\partial^2}{\partial z^2}(S_z^{-1/2})$ has been used.

If $T = 0$, the complex potential and velocity are determined by S_t alone; in particular, time independent singularities of S do not play a role. To be more precise, let $z \in D_{t'}$ for $0 \leq t' \leq t$, and suppose that there are no singularities of pressure in $D_{t'}$. Then, from (3.8) with $T = 0$,

$$\int_0^t w_z(z,t')dt' = K(S(z,0) - S(z,t)),$$

from which it follows that the singularities in D_t of $S(z,t)$ coincide with those of $S(z,0)$, are constant in time, and do not affect $w_z(z,t)$. These points have been noted previously ([9], [10]), and used by Howison to simplify the integration of a system of differential equations that determine the

evolution of C_t for an exterior problem. On the other hand, if $T \neq 0$, then zeros of S_z become important. For, if $S_z(z,t) = (z-z_1)^n H(z,t)$, where H is analytic near z_1, $H(z_1,t) \neq 0$, and n is a positive integer, then, near z_1, $S_{zz}.S_z^{-3/2} \sim n(z-z_1)^{-1-1/2n} H(z_1,t)^{-1/2}$; thus w has a branch point at z_1 of order at least -3/2 if n is odd, and a pole of order at least two if n is even. Although mathematically significant, such singularities are likely to be difficult, if not impossible, to synthesise physically. In general, then, such an S will not be admissible, and the corresponding C_t will not be realisable. Some illustrative examples will be given later.

The elimination of possible solutions by surface tension is consistent with an observation in [6]. There, when $T = 0$, it was shown that a cusp develops on the fluid boundary when a zero of the derivative of a function that maps the fluid region conformally and 1-1 onto the unit disc reaches the unit circle from its exterior. Such a zero may be shown to correspond to a zero of S_z in the fluid region.

4. PHYSICAL REALISABILITY OF SOLUTIONS

In studying inverse problems, one objective is to determine all singularities of the solution. Some of these may be internal singularities that lie in the region occupied by fluid and others may be external singularities that exist in the analytic continuation of the solution beyond the fluid boundary C_t. For a given interface C_t, into which category a particular singularity falls depends on whether we are considering an interior problem or an exterior problem.

The external singularities may be regarded as image singularities that are induced in the solution and depend on the form of C_t. This interpretation is common in electrostatics, where the effect in the region of interest of the

image of a source in a conductor is produced by an appropriate induced charge distribution on the conductor. In the Hele-Shaw probem, the fluid external to D_t reacts on C_t against the internal fluid to create an effect equivalent to that of the external singularities, thus resulting in the required pressure on, and motion of, C_t. Consequently, we need not account explicitly for the effect of external singularities, except to note that analyticity of C_t will be destroyed at time t at a point where such a singularity meets C_t ([6], [10], [15]).

A further aim is to decide whether the solution is uniquely determined by its internal singularities and, if so, whether they can be realised in practice. For an interior problem, it was shown originally by Richardson ([15], [16]) that if the internal singularities are simple point sources, then all the complex moments of D_t are determined, given the initial configuration. From these is found the part of S that is analytic outside C_t and zero at infinity. (Richardson's analysis would appear to generalize to other types of singularity.) It was also pointed out in [16] that knowledge of the moments of D_t alone does not necessarily determine D_t uniquely, as shown by the example of Sakai ([21]). Nevertheless, from the fact that functions and domains in the Hele-Shaw problem depend continuously on time, Richardson was led to conjecture that some such domains would be uniquely determined by their moments, and thus by their internal singularities. Specifically, he has determined the evolution of C_t when there are as many as four internal simple sources. His method involves the conformal mapping procedure; it becomes more and more complicated as the number of singularities increases, and impractical if the singularities of S inside C_t are not poles ([10]). (Note that S is meromorphic inside C_t if and only if the mapping function occurring in this procedure is rational ([2, p. 158]).)

In the examples that follow, we shall assume that internal singularities in the form of simple point sources, doublets, or continuous distributions of simple sources on finite curves can be synthesised physically and, even when $T \neq 0$, like Richardson, we shall conjecture that the internal singularities and initial configuration uniquely determine the evolution of C_t.

5. TIME-DEPENDENT INVERSE PROBLEMS WITH T = 0

Here we consider time-dependent problems in which surface tension is neglected. Many such solutions have been obtained: see, for example, [6], [7], [8], [9], [10], [15], [16], [17]. In Richardson's work, solutions are obtained that can be generated by realistic sources. It seems less likely, however, that many of the other solutions can be realized in practice, even with the neglect of surface tension effects.

Let us begin with a simple example. Suppose that C_t is the circle of radius $r(t)$, with centre at 0. Then $S(z,t) = r^2(t)/z$, and (3.5) gives

$$u(x,y;t) = -K\dot{r}r[\log z - \log(r^2/\bar{z})],$$

in which the principal branch of the logarithm is selected. Then $\log z - \log(r^2/\bar{z}) = 2\log(|z|/r)$, and

$$u(x,y;t) = -2K\dot{r}r \log(|z|/r), \tag{5.1}$$

a result that follows also from (3.6).

This solution is valid whether D_t is bounded internally or externally by C_t. For the interior problem, $0 < |z| < r$, and one finds $u \gtrless 0$ accordingly as $\dot{r} \gtrless 0$. When $\dot{r} > 0$, fluid is injected at 0 and extraction takes place if $\dot{r} < 0$. The fluid velocity is radial, the outward component being $r\dot{r}/|z|$. The source is at 0, and the flux across any curve enclosing 0 is equal to the time rate of change of the area of D_t.

When D_t lies outside C_t, $|z| > r$ and $u \lessgtr 0$ accordingly as $\dot{r} \gtrless 0$. If $\dot{r} > 0$, fluid is removed at infinity; fluid is injected if $\dot{r} < 0$. Again, the velocity is radial and equal to $r\dot{r}/|z|$, and there is a logarithmic source of pressure at infinity. In practice, this flow could be generated in the region between a circle centred on 0 in $|z| > r$ and the circle $|z| = r$ by setting up the appropriate pressure difference between these circles, the pressure in $|z| \leq r$ being zero.

5.1. Interior problem for ellipses

We consider a rather more difficult example, in which the curves $C_t (t \geq 0)$ evolve in the form of a family of ellipses.

Let C_t be the ellipse $b^2x^2 + a^2y^2 = a^2b^2$, where a and b are prescribed functions of time, and $a > b > 0$. The Schwarz function for C_t is

$$S(z,t) = \frac{a^2+b^2}{c^2} z - \frac{2ab}{c^2} (z^2-c^2)^{1/2} \tag{5.2}$$

([2, (5.13)]), in which $c^2 := a^2 - b^2$, and the complex plane is cut from $-c$ to c with $(z^2-c^2)^{1/2} > 0$ for $z > c$.

We shall consider the velocity field v of the fluid in the domain D_t contained within C_t. It is given by (2.1) where, by (3.8),

$$u_x - iu_y = -K\left[z \frac{d}{dt}\left[\frac{a^2+b^2}{c^2}\right] - (z^2-c^2)^{1/2} \frac{d}{dt}\left[\frac{2ab}{c^2}\right] + \right.$$

$$\left. + \frac{1}{(z^2-c^2)^{1/2}} \cdot \frac{2ab\dot{c}}{c}\right]. \tag{5.3}$$

Integration of (5.3) on a curve C enclosing the segment $(-c,c)$ shows that the flux across C is the rate of change of the area of D_t.

The analytic function $u_x - iu_y$ may be represented by its Cauchy integral and so related to its singularities. Deformation of the path of integration around singularities and into a circle at infinity, followed by some simplification, leads to the following result:

$$u_x - iu_y = -Kz \frac{d}{dt}\left[\frac{a-b}{a+b}\right] + \frac{K}{\pi}\int_{-c}^{c}\left[\sqrt{(c^2-\xi^2)}\,\frac{d}{dt}\left[\frac{2ab}{c^2}\right] + \frac{1}{\sqrt{(c^2-\xi^2)}}\,\frac{2\dot{abc}}{c}\right]\frac{d\xi}{\xi - z}$$

and, since $u_x - iu_y = w_z$, we find that

$$w(z,t) = -\frac{1}{2}Kz^2\frac{d}{dt}\left[\frac{a-b}{a+b}\right] - \frac{2K}{\pi}\int_{-c}^{c}\left[\sqrt{(c^2-\xi^2)}\frac{d}{dt}\left[\frac{ab}{c^2}\right] + \frac{1}{\sqrt{(c^2-\xi^2)}}\,\frac{\dot{abc}}{c}\right]\log(z-\xi)d\xi + g(t). \tag{5.4}$$

Here any branch of the logarithm may be selected, since only u(= Re w) is of interest.

From (5.4) it is seen that w is generated by internal simple sources along the interfocal segment, while the term in z^2 arises from external sources at infinity. In general, c is dependent on time and the internal sources are not stationary.

We consider two special cases, but only one in detail. In the first, it is assumed that the ellipses have constant eccentricity. Then $\frac{d}{dt}\left[\frac{a-b}{a+b}\right] = 0$, $\frac{d}{dt}\left[\frac{ab}{c^2}\right] = 0$, and (5.4) simplifies to

$$w(z,t) = -\frac{2K\dot{ab}}{\pi}\int_{-c}^{c}\frac{\log(z-\xi)}{\sqrt{(c^2-\xi^2)}}\,d\xi + g(t).$$

The function g may be chosen so that $u = 0$ on C_t, and

$$u(x,y;t) = 2K\dot{ab}\left[\log\left[\frac{a+b}{2}\right] - \frac{1}{\pi}\int_{-c}^{c}\frac{\log|z-\xi|}{\sqrt{(c^2-\xi^2)}}\,d\xi\right]. \tag{5.5}$$

From this, one sees that the motion is generated by a continuous distribution of simple sources on (-c,c) with density $-2K(\dot{ab}/\pi)/\sqrt{(c^2-\xi^2)}$, $-c < \xi < c$.

When z lies on the segment (-c,c), it is found that the integral in (5.5) is equal to $\pi \log(\frac{1}{2}c)$, so

$$u(x,0;t) = 2K\dot{a}b \log\left[\frac{a+b}{c}\right], \quad -c < x < c; \tag{5.6}$$

here the logarithm is constant in time, and u is constant for $-c < x < c$, $y = 0$, for a given time.

Let C' be an ellipse in D_t that is confocal with C_t. Then it may be shown that u is constant on C'. More precisely, if a' and b' are its semi-major and -minor axes, respectively, then

$$u(x,y;t) = 2K\dot{a}b \log\left[\frac{a+b}{a'+b'}\right], \quad z \in C', \tag{5.7}$$

which reduces to (5.6) if C' collapses onto (-c,c).

Since u is constant on C', v is normal to C'. A calculation based on (3.11) with T = 0 verifies this and gives

$$v = \frac{a'\dot{a}b}{\sqrt{(a'^4 - c^2x^2)}}\,\hat{n}', \tag{5.8}$$

in which $\hat{n}'$ is the unit normal to C'. This determines the analytic function $u_x - iu_y$ on C' and, hence, everywhere. In principal, then, if this velocity is impressed across C', the correct motion will be established inside C_t; but because analytic continuation is unstable, small errors in impressed velocity on C' will lead to large errors elsewhere.

For the second case, we assume that the ellipses are confocal. Then, on omitting details, it is found that

$$u(x,y;t) = \frac{1}{2}(y^2-x^2)\frac{d}{dt}\left[\frac{a}{a+b}\right] - \frac{1}{\pi c^2}\frac{d}{dt}(ab)\int_{-c}^{c}\sqrt{(c^2-\xi^2)}\log|z-\xi|\,d\xi$$

$$+ \frac{a^3b}{2c^2}\frac{d}{dt}\left[\frac{b}{a}\right] + \frac{1}{2}\left[\log\left[\frac{a+b}{2}\right] - \frac{1}{2}\right]\frac{d}{dt}(ab).$$

It may be seen that u(x,0;t) depends on x when $-c < x < c$, in contrast to the

result in the previous example. In fact, the curves on which u is constant are confocal ellipses only when the ellipses C_t are members of a family with constant eccentricity.

5.2 Exterior problem for ellipses

Suppose that the domain occupied by fluid is outside C_t, with extraction from, or injection into, D_t at infinity. It is assumed that S (and, hence, S_t) has no singularities in the finite plane outside C_t; thus the only sources in the fluid region are at infinity. We shall also suppose that $u_x - iu_y$ behaves like z^{-1} as $|z| \to \infty$, so that the time rate of change of the area inside C_t is bounded; any greater growth at infinity seems to be unreasonable on physical grounds. Consequently S_t is of order z^{-1} at infinity. Finally, it will be assumed that S is at most of order z as $|z| \to \infty$.

Subject to the assumptions on S alone, it may be shown that $C_t(t \geq 0)$ determines a family of ellipses ([22], [23], [14]), for which, in general, S_t is of order z at infinity and the behaviour of $u_x - iu_y$ is inadmissible. (It is conjectured that $u_x - iu_y$ would have inadmissible behaviour at infinity if S grew more rapidly than z, but we can offer no proof.) Only if the family of ellipses has constant eccentricity will S_t be of order z^{-1}, as required, and clearly C_0 must be a member of this family.

Then, subject to the assumptions above, one concludes that C_t can sweep over every point outside C_0 if and only if C_0 is an ellipse, and $C_t(t > 0)$ is an ellipse with the same eccentricity as C_0. This result has been given previously in [8], and generalised to $\mathbb{R}^n$ in [3].

To discuss the possibility of synthesising this solution, the form of u at infinity is needed. With notation unchanged from the previous section, (3.7)

and $S_t(z,t) = 2\dot{a}b(z^2-c^2)^{-1/2}$ give

$$w(z,t) = -2K\dot{a}b \int_{z_0}^{z} \frac{d\zeta}{(\zeta^2-c^2)^{1/2}}, \qquad z_0 \in C_t.$$

Since u is independent of z_0 we may take $z_0 = a$. Then the integral is expanded for $|z| > c$ to yield

$$w(z,t) = -2K\dot{a}b[\log z - \frac{1}{4} c^2 z^{-2} - \frac{3}{32} c^4 z^{-4} + \ldots] + h(t),$$

in which log denotes the principal branch, and h is real. Thus

$$u(x,y;t) = -2K\dot{a}b \log|z| + h(t) + O(|z|^{-2}), \qquad |z| \to \infty.$$

If C_t is a circle, so $c = 0$, this result reduces to (5.1). There are no internal singularities of pressure, but there is a logarithmic singularity at infinity. The external singularities lie along the interfocal segment of the ellipse C_t.

In practice, instead of a singularity at infinity, a source or sink of pressure would be impressed along some closed curve C'; D_t would be bounded internally by C_t and externally by C'. If C' were chosen to be a large circle centred on 0 then, at least initially, the desired flow would be generated. Alternatively, if C' were an ellipse, instantaneously confocal with C_t and with semi-axes a' and b' ($a' > a$, $b' > b$), then the pressure on C' is given by (5.7). The maintenance of this pressure on C' would produce the flow in the region between C' and C_t but, because C' changes with time, this would be difficult to realise physically.

5.3. Interior and exterior problems for limacons

As a further example, it is assumed that C_t is a limacon. This has been a popular choice in earlier work that uses the conformal mapping technique to generate solutions, because the mapping function is a quadratic polynomial;

but no attempt seems to have been made to relate the solutions to their singularities.

In polar coordinates, with the pole of the system at $z = -a$, C_t has the form

$$r = 1 + 2a \cos \theta;$$

here a is a function of t and it is assumed that $0 \leq a < \frac{1}{2}$, so that C_t is simple. Then ([2, p.54])

$$S(z,t) = \frac{(R+1)[2z+a(R+1)]}{4z^2}, \tag{5.9}$$

in which $R := (1+4az)^{1/2}$, and $R > 0$ for $z > -1/(4a)$. Since $a < \frac{1}{2}$, the branch point lies outside C_t, and the complex plane is cut along the negative real axis from $-1/(4a)$ to $-\infty$. If $a = 0$ for $t = 0$, then C_0 is a circle of unit radius, centred on 0. The branch point moves to the right as a increases, and meets C_t when $a = \frac{1}{2}$.

The family of limaçons considered here is a one-parameter subset of the original two-parameter family considered by others ([6], [7]). In the earlier work, the two parameters are related in such a manner that the corresponding S_t has only a simple pole at $z = 0$. Here, in contrast, S_t will have a pole of order two at $z = 0$. More precisely, for $a > 0$, S has a pole of order two at $z = 0$ and no other singularities inside C_t. Near $z = 0$,

$$S(z,t) = \frac{a}{z^2} + \frac{2a^2+1}{z} + (a-a^3) + \ldots, \tag{5.10}$$

and

$$S_t(z,t) = \frac{\dot{a}}{z^2} + \frac{4a\dot{a}}{z} + (1-3a^2)\dot{a} + \ldots . \tag{5.11}$$

By employing (3.8) and the Cauchy integral representation for $u_x - iu_y$, we obtain

$$u_x - iu_y = -K\left[\frac{\dot{a}}{z^2} + \frac{4a\dot{a}}{z} - \frac{\dot{a}}{2\pi}\int_{-\infty}^{1/(4a)} \frac{2\xi^2+6a\xi+1}{\xi^2|1+4a\xi|^{1/2}} \frac{d\xi}{\xi-z}\right]. \qquad (5.12)$$

Thus, to within an additive function of t alone,

$$w(z,t) = K\left[\frac{\dot{a}}{z} - 4a\dot{a}\log z - \frac{\dot{a}}{2\pi}\int_{-\infty}^{1/(4a)} \frac{2\xi^2+6a\xi+1}{\xi^2|1+4a\xi|^{1/2}} \log(\xi-z)d\xi\right],$$

and

$$u(x,y;t) = K\left[\frac{\dot{a}x}{x^2+y^2} - 4a\dot{a}\log|z| - \frac{\dot{a}}{2\pi}\int_{-\infty}^{1/(4a)} \frac{2\xi^2+6a\xi+1}{\xi^2|1+4a\xi|^{1/2}} \log|\xi-z|d\xi\right].$$

Suppose now that the fluid region D_t is interior to C_t. The internal singularities at $z = 0$ correspond to a simple source and a doublet. The external singularities are a distribution of simple sources on the negative real axis from $-\infty$ to $-1/(4a)$. If $0 \leq a < \frac{1}{2}$, the motion will be generated by the internal sources; when $a = \frac{1}{2}$, an external singularity meets C_t, which loses analyticity at this point and the solution breaks down.

If the fluid is external to C_t, then the internal singularities are simple sources on $(-\infty,-1/(4a))$, and the external singularities correspond to a simple source and a doublet at the origin. Starting from an admissible initial configuration (for example, $a(0) = 0$ so that C_0 is the unit circle), this flow could be generated by the source distribution on $(-\infty,-1/(4a))$; in practice, this source configuration would be difficult to set up.

6. TIME-DEPENDENT PROBLEMS WITH T ≠ 0

We shall now briefly re-examine the examples discussed earlier, but with surface-tension effects included. It has already been noted that zeros of S_z in the flow region are not permissible, so any such cases will be omitted.

It goes almost without saying that in general one effect of surface

tension is to complicate the analysis. But if C_t is a circle, the curvature κ is constant, and the condition (2.3) merely raises or lowers the pressure u by a constant amount. Thus surface tension does not affect the motion in this case, an observation made previously in [24].

The complex potential is given by (3.7), from which it follows that

$$u_x - iu_y = \mp iT \frac{\partial^2}{\partial z^2}(S_z^{-1/2}) - \kappa S_t. \qquad (6.1)$$

Once again $u_x - iu_y$ may be related to its singularities by means of the Cauchy integral representation. In addition to singularities that arise when $T = 0$, the zeros of S_z now play a role.

For the ellipse, from (5.2) we find that S_z has two simple zeros at $z = \pm(a^2+b^2)/c$. Since $a(a-c) + b^2 > 0$, it follows that these lie outside the ellipse. Thus the term in (3.7) or (3.8) that depends on T has two branch points inside C_t and two outside C_t. Since S_z has zeros outside the ellipse, for reasons mentioned earlier we shall disregard the exterior problem. For the interior problem, additional cuts are now made in the complex plane outside C_t from $\pm(a^2+b^2)/c$ to $\pm\infty$. In addition to the integral arising when $T = 0$, the representation for w when $T \neq 0$ will contain an integral on $(-c,c)$ and integrals on these new cuts. Their integrands will behave like $[z\pm(a^2+b^2)/c]^{-3/2}$ near these points, and the integrals will involve second derivatives of the potentials of simple source distributions. Since these are external singularities, we conjecture that the flow is realisable, in accordance with the discussion in section 4.

In the case of a limaçon, from (5.9) it is found that S_z has only one zero: $z = -2a(1-2a^2)$; because $0 \leq a < \frac{1}{2}$, this lies inside C_t, and the corresponding term in w will have a branch point of order $-3/2$ at this point. Since it is impossible to define $u_x - iu_y$ by (6.1) as a single-valued analytic

function in a neighbourhood of C_t, we conclude that the limaçon is not a solution for either interior or exterior problems if $T \neq 0$. This, then, is a specific example that illustrates how surface tension limits the class of possible solutions.

7. CONCLUDING REMARKS

The one-phase Hele-Shaw flow problem on an unbounded domain can be formulated as a Cauchy problem for the Laplace equation. By adopting this point of view, and using an explicit representation for the solution, the required harmonic function is expressed completely in terms of the Schwarz function of the interfacial curve. This form of the solution is well-suited to the study of inverse problems, in which the Schwarz function is prescribed. The realisability in practice of the flow when the interface is circular, elliptical, or in the form of a limaçon, has been examined by relating the Schwarz function to its singularities, with or without the consideration of surface tension effects.

Other problems can be studied in this way. These include steady-state problems, in which the curve C_t is merely the translation of C_0 with uniform velocity V_0 along the x-axis. Then $S(z,t) = V_0 t + S(z-V_0 t,0)$ ([2,(8.8)]), so $S_t(z,t) = V_0[1-S_z(z-V_0 t,0)]$ and the integration in (3.5) and (3.6) can be performed explicitly. Such problems can be reduced to examination of $S(z;0)$; they can also be formulated directly in a more elementary way ([13]). Examples include the motion and shape of bubbles and fingers in Hele-Shaw cells of finite width.

Problems of two-phase flow can be formulated in the same manner, although the Cauchy data on the interface are not given explicitly in terms of S, since there is coupling between the solutions in the two fluids.

Moving boundary problems arise also in many other contexts; see, for example, the introduction to [6]. Some of these involve synthesis or control, and the present approach should be useful there.

Direct problems are more difficult to examine in the present formulation, for they require the determination of S from prescribed sources. Some consideration has been given to this general problem ([6], [9], [10], [15], [16]). Much remains to be done, and it is hoped that others will be encouraged to study the question.

ACKNOWLEDGEMENT

I am grateful to S. D. Howison, J. R. Ockendon, S. Richardson, and R. J. Tait for helpful discussions of this work. This research was supported by the Natural Sciences and Engineering Research Council of Canada under Grant A 5388.

REFERENCES

[1] Bensimon, D., Kadanoff, L. P., Liang, S., Shraiman, B. I. and Tang, C., Viscous flows in two dimensions, Rev. Mod. Phys. 58 (1986), 977-999.

[2] Davis, P. J., The Schwarz Function and its Applications, The Carus Mathematical Monographs, No. 17, Mathematical Association of America, 1974.

[3] Friedman, A., and Sakai, M., A characterization of null quadrature domains in $\mathbb{R}^N$, Indiana Univ. Math. Journal 35 (1986), 607-610.

[4] Garabedian, P. R., An example of axially symmetric flow with a free surface, in "Studies in Mathematics and Mechanics Presented to Richard von Mises", 149-159, Academic Press, New York, 1954.

[5] Homsy, G. M., Viscous fingering in porous media, Ann. Rev. Fluid Mech. 19 (1987), 271-311.

[6] Howison, S. D., Ockendon, J. R. and Lacey, A. A., Singularity development in moving-boundary problems, Quart. J. Mech. Appl. Math. 38 (1985), 343-360.

[7] Howison, S. D., Cusp development in Hele-Shaw flow with a free surface, SIAM J. Appl. Math. 46 (1986), 20-26.

[8] Howison, S. D., Bubble growth in porous media and Hele-Shaw cells, Proc. Roy. Soc. Edinburgh 102A (1986), 141-148.

[9] Howison, S. D., Fingering in Hele-Shaw cells, J. Fluid Mech. 167 (1986), 439-453.

[10] Lacey, A. A., Moving boundary problems in the flow of liquid through porous media, J. Austral. Math. Soc. (Series B) 24 (1982), 171-193.

[11] Lamb, H., Hydrodynamics, 6th ed., Dover, New York, 1945.

[12] Millar, R. F., The analytic continuation of solutions to elliptic boundary value problems in two independent variables, J. Math. Anal. Appl. 76 (1980), 498-515.

[13] Millar, R. F., The steady motion and shape of a bubble in a Hele-Shaw cell, Proceedings of the Canadian Applied Mathematics Society Conference on Continuum Mechanics and its Applications, Simon Fraser University, Burnaby, Canada, 30 May - 3 June, 1988, G. A. C. Graham and S. K. Malik, editors, Hemisphere Publishing Corporation, New York, 1989, 603-610.

[14] Millar, R. F., Inverse problems for a class of Schwarz functions, Complex Variables Theory Appl. (to appear).

[15] Richardson, S., Hele Shaw flows with a free boundary produced by the injection of fluid into a narrow channel, J. Fluid Mech. 56 (1972), 609-618.

[16] Richardson, S., Some Hele Shaw flows with time-dependent free boundaries, J. Fluid Mech. **102** (1981), 263-278.

[17] Richardson, S., Hele Shaw flows with time-dependent free boundaries in infinite and semi-infinite strips, Quart. J. Mech. Appl. Math. **35** (1982), 531-548.

[18] Saffman, P. G., Viscous fingering in Hele-Shaw cells, J. Fluid Mech. **173** (1986), 73-94.

[19] Saffman, P. G. and Taylor, G. I., The penetration of a fluid into a porous medium or Hele-Shaw cell containing a more viscous liquid, Proc. Roy. Soc. London **A 245** (1958), 312-329.

[20] Saffman, P. G. and Tanveer, S., Prediction of bubble velocity in a Hele-Shaw cell: thin film and contact angle effects, Phys. Fluids **A1** (1989), 219-223.

[21] Sakai, M., A moment problem on Jordan domains, Proc. Amer. Math. Soc. **70** (1978), 35-38.

[22] Sakai, M., Null quadrature domains, J. d'Analyse Math. **40** (1981), 144-154.

[23] Shapiro, H. S., Domains allowing exact quadrature identities for harmonic functions - an approach based on P.D.E., in Anniversary Volume on Approximation Theory and Functional Analysis, P. L. Butzer et al (ed.), International Series of Numerical Mathematics, Vol. 65, Birkauser Verlag, Basel, 1984, 335-354.

[24] Taylor, G. I. and Saffman, P. G., A note on the motion of bubbles in a Hele-Shaw cell and porous medium, Quart. J. Mech. Appl. Math. **12** (1959), 265-279.

R. F. Millar
Department of Mathematics
University of Alberta
Edmonton, Alberta
Canada T6G 2G1

A.B. PARRY

The scattering of velocity fields by an airfoil in compressible flow

1. INTRODUCTION

The interaction of an airfoil with a fluctuating velocity field, and the calculation of the resultant scattered field, is generally of great importance in both aerodynamics and hydrodynamics. Most of the work in these ares has been concentrated on low frequency interaction problems in incompressible or weakly compressible flow (von Karman & Sears 1938; Sears 1940; Kemp 1952, 1973; Osborne 1973; Goldstein 1976, chap. 3). The latter can, of course, be reduced to an equivalent incompressible flow problem by a suitable transformation (see, for example, Ward 1955; Landau & Lifshitz 1959). Here, however, we consider such scattering problems in as much as they apply to noise generation by blade row interactions on the new generation of advanced propellers. In this application the blades operate in the high subsonic regime - even at 'take off' and 'approach' conditions.[1] Accordingly, the usual methods of dealing with these interaction problems are inappropriate. Moreover - as we will see below - the disturbance velocity field is not always convected with the mean flow, as is usually the case.

The way in which the fluctuating velocity field is modelled, and the far-field sound obtained from the unsteady pressure distribution across the airfoil chord (and, indeed, across the whole span of the blade) on advanced propellers, is described in detail in Parry (1988) and Parry & Crighton

[1] In virtually all of the present designs, the blade tips only operate supersonically at the 'cruise' condition or design point.

(1989). Here, therefore, it is only necessary to consider the interaction of a single Fourier component of the velocity field with a flat plate. The problem is assumed to be two-dimensional.[2] Our aim is to calculate the scattered field and thereby the distribution of unsteady pressure across the plate.

In §2 we consider the interaction of an airfoil with a convected gust from upstream infinity. Applications of this to the wake interactions on a counter-rotation propeller show considerable differences between the measured and predicted acoustic field. In §3 the work is extended to include the interaction of an airfoil with potential (i.e. non-convected) velocity fields from both upstream and downstream infinity. In addition, an iterative technique is presented which can be used to put the solution in the form of an asymptotic series. Comparisons between predictions and measurements of the far-field noise of a counter-rotation propeller show that the predictions are extremely accurate, in terms of both absolute levels and noise directivity.

2. CONVECTED GUST INTERACTIONS

The major source of aerodynamic interference between blade rows is taken, usually (and naturally enough), to be the unsteady velocity field associated with the viscous wakes generated by the upstream blade row.[3] We suppose that

---

[2] Some justification for the two-dimensional approximation comes from the asymptotic analysis of Parry & Crighton (1989) who showed that noise generation is highly localised at discrete radii.

[3] In addition the tip-vortex, of current interest with regard to advanced propellers, can be represented by convected gusts. Tip vortex interactions can, therefore, also be described by the methods of this section - providing, of course, that the velocity field is known.

the gusts - representing the velocity field - are of relatively high frequency; this is appropriate for advanced propellers which have large numbers of blades (in addition, we must remember that the ear is most sensitive to the higher frequency interactions which, therefore, assume greater importance).

A transformation

We start with a single Fourier component of the convected velocity field given by

$$v = u \exp[ik(Ut-x)] \tag{1}$$

on $y = 0$. The axes x and y are oriented as shown in figure 1 and centred on the airfoil leading edge with U the velocity in the x direction and k the wavenumber. The disturbance velocity potential ϕ will satisfy the convected wave equation which we write as

$$(1-M^2)\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} - 2ikM^2\frac{\partial\phi}{\partial x} + k^2M^2\phi = 0 \tag{2}$$

where $M = U/c_0$ is the Mach number and the time dependence e^{ikUt} is implied. The boundary condition on the airfoil surface is simply

$$\frac{\partial\phi}{\partial y} + v = 0 \quad \text{on} \quad y = 0,\ 0 \le x \le c \tag{3}$$

where c is the chord length of the airfoil.

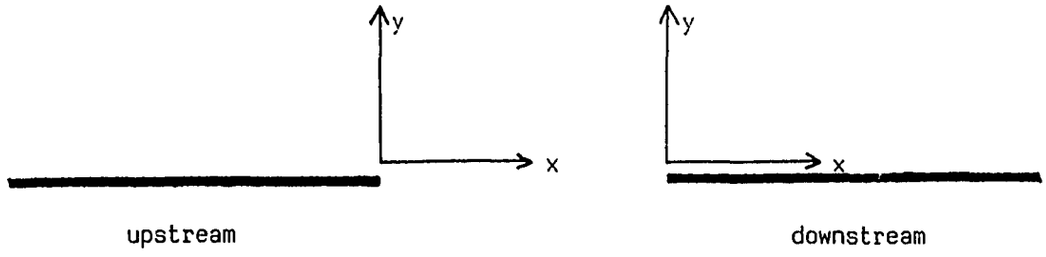


Figure 1. The coordinate axes for downstream (leading edge) interactions or upstream (trailing edge) interactions.

We now introduce the Prandtl-Glauert coordinates, scaled on the semi-chord c/2, given by

$$X = 2x/c, \quad Y = 2\beta y/c, \tag{4}$$

where $\beta = \sqrt{(1-M^2)}$ is the usual compressibility factor. We also introduce a new potential ψ such that

$$\phi(x,y) = \frac{cu}{2\beta}\,\psi(X,Y)\exp\left[i\,\frac{\sigma M^2}{1-M^2}\,X\right]. \tag{5}$$

The result of (4) and (5) is that (2) then reduces to the standard Helmholtz equation

$$\nabla^2\psi + K^2\psi = 0 \tag{6}$$

where the new nondimensional wavenumber K is given by

$$K = \frac{\sigma M}{1-M^2} \tag{7}$$

and $\sigma = kc/2$ is the reduced frequency. In the new coordinate system the boundary condition on the plate is given by

$$\frac{\partial \psi}{\partial Y} = -\exp(-i\kappa X) \quad \text{on} \quad Y = 0,\ 0 \le X \le 2, \tag{8}$$

and the pressure $p = -\rho D\phi/Dt = -\rho U(ik+\partial/\partial x)\phi$ is related to the transformed potential ψ by

$$p = -\frac{\rho U u}{\sqrt{(1-M^2)}} \exp\left[i\,\frac{\sigma M^2}{1-M^2}\,X\right]\left[i\kappa + \frac{\partial}{\partial X}\right]\psi . \tag{9}$$

where the wavenumber κ in (8) and (9) is

$$\kappa = \frac{\sigma}{1-M^2}. \tag{10}$$

Airfoil response

The solution to (6), subject to the boundary condition (8), at high frequencies produces a pressure distribution which oscillates rapidly away from the leading edge of the airfoil where, indeed, it will be (integrably) singular. The pressure will be, therefore, to a large part self-cancelling. The trailing-edge region, where a Kutta condition is applied, should then be relatively unimportant so that the pressure distribution is much the same as that on a flat plate extending to downstream infinity. This leading-edge problem, or two-part boundary problem, can be solved by the Wiener-Hopf technique (see, for example, Noble 1958; Crighton 1977). The solution has, however, been given previously by Landahl (1961) and Goldstein (1976), who

used an alternative approach, and we simply quote the result as[4]

$$\Delta p(X) = \frac{2\rho U u}{[\pi\sigma(1+M)X]^{1/2}} \exp\left[-i\frac{\pi}{4} - i\frac{\sigma M}{1+M}X\right]. \qquad (11)$$

Measurement vs. prediction

It is difficult to validate this result directly as we have no suitable data on blade unsteady response. However, by combining it with the noise radiation formulae for counter-rotation propellers (Hanson 1985), and an appropriate model for the (viscous wake) unsteady velocity field (Parry 1988), we can obtain predictions of far-field noise. (The way in which these different stages can be combined into a robust prediction scheme has been described by Parry & Crighton 1989). These predictions will be compared with measurements taken from the flyover tests carried out by Rolls-Royce on the Fairey Gannet counter-rotation propeller: details of these tests have been discussed by Bradley (1986). This comparison has been given before by Parry & Crighton (1989) but, for the sake of completeness, and since the comparison serves as a check on the blade response calculation, we will give it again here along with a brief discussion. Since the front and rear 4-blade rows on the Gannet were run at slightly different speeds, the interaction tone components could all be separated out in terms of frequency, thus allowing us to examine each tone individually.

The first interaction tone generated by the Gannet is the (1,1)

[4] Our solution is the complex conjugate of Goldstein's since he chose a time dependence $e^{-i\omega t}$ and we have used $e^{i\omega t}$. In addition, a phase term $e^{i\sigma}$ is missing from our result since here the reference is the airfoil leading edge and not the mid-chord as in Goldstein's work.

interaction.[5] The directivity plot (sound pressure level versus radiation angle where 0^{o} represents the direction of flight or the propeller axis) for this tone is shown in figure 2. We can see that there is a null at the 90^{o} radiation angle in both the measured and predicted data. This is to be expected since the (1,1) interaction tone generates a plane-wave mode (on the Gannet which has equal blade numbers, B_1 and B_2 on the front and rear rows), i.e. $n_1B_1 - n_2B_2 = 0$. This tone peaks on the propeller axis and is zero in the plane of the rotor. However, predictions are typically 25dB below the measurements.

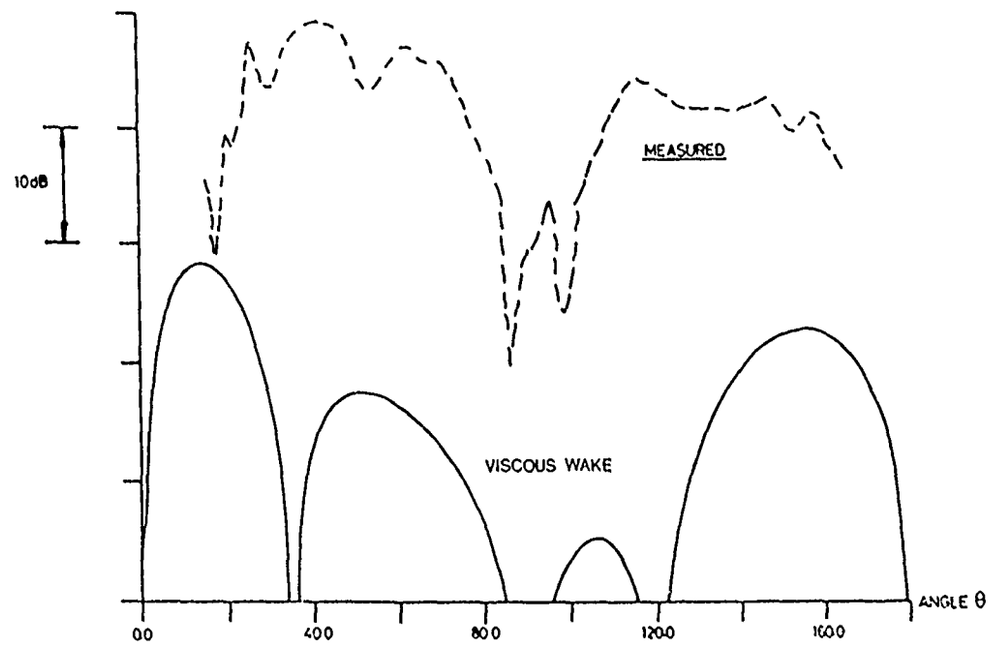


Figure 2. Gannet measurements vs predicted wake interaction noise for the (1,1) interaction tone.

[5] We will use the notation (n_1,n_1) to indicate an aerodynamic interaction tone generated at the combination frequency $n_1f_1 + n_2f_2$ where f_1 and f_2 are the blade passing frequencies of the front and rear blade rows respectively.

The next two interaction tones generated by the Gannet are the (2,1) and (1,2) interactions, for which directivity plots are shown in figures 3 and 4 respectively. For these interaction tones the predictions are at least 10dB below the measurements. In fact, it is not only the levels which are incorrectly predicted but also the directivities: predictions are 20dB below the measurements for the (1,2) interaction in the forward arc, and 40dB below the measurements for the (2,1) interaction in the rear arc.

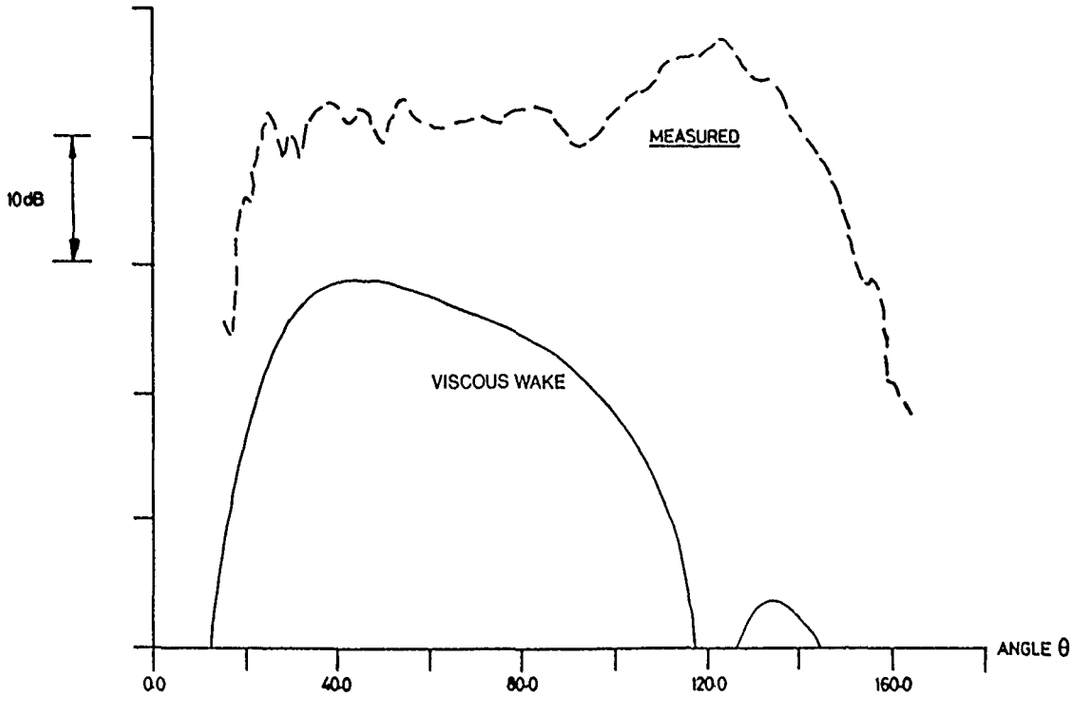


Figure 3. Gannet measurements vs. predicted wake interaction noise for the (2,1) interaction tone.

We now refer to the discussion on interaction tone directivities given by Bradley (1986) and Parry & Crighton (1989). There it was shown that, although the measured directivities are similar in level in forward and rear arcs, any noise source on the rear blade row will produce only asymmetrical directivities (except for the plane wave case where $n_1 = n_2$). There must, therefore, be a significant noise source on the forward blade row. In order to generate tones at the (n_1,n_2) interaction tone frequencies, the noise source can only be that due to the interaction of the forward blade row with the potential field generated by the rear row.

3. POTENTIAL FIELD INTERACTIONS

Accordingly, we proceed to consider the scattering of a potential velocity field by an airfoil. In addition to the interaction of the forward blade row with the potential field generated by the rear row, it seems appropriate to include the interaction of the downstream blades with the bound potential field generated by the forward blade row, because that is likely to generate at least as large a field as that from the interaction of the forward row with the potential field of the rear row. The latter involves a trailing edge (weakly loaded if a Kutta condition is satisfied), the former a leading edge (highly loaded). As before we assume that the detailed velocity field is known (the modelling of the bound potential flow field of a counter-rotation propeller, in compressible flow, is described in full by Parry 1988) and consider a single harmonic component of the upwash

$$v = u \exp(i\omega t - i\gamma x) \qquad (12)$$

on $y = 0$. The axes x and y are centred on the airfoil trailing edge, for upstream interactions, or on the leading edge, for downstream interactions, as shown in figure 1. The velocity in the x-direction is U and γ is the complex

wavenumber.

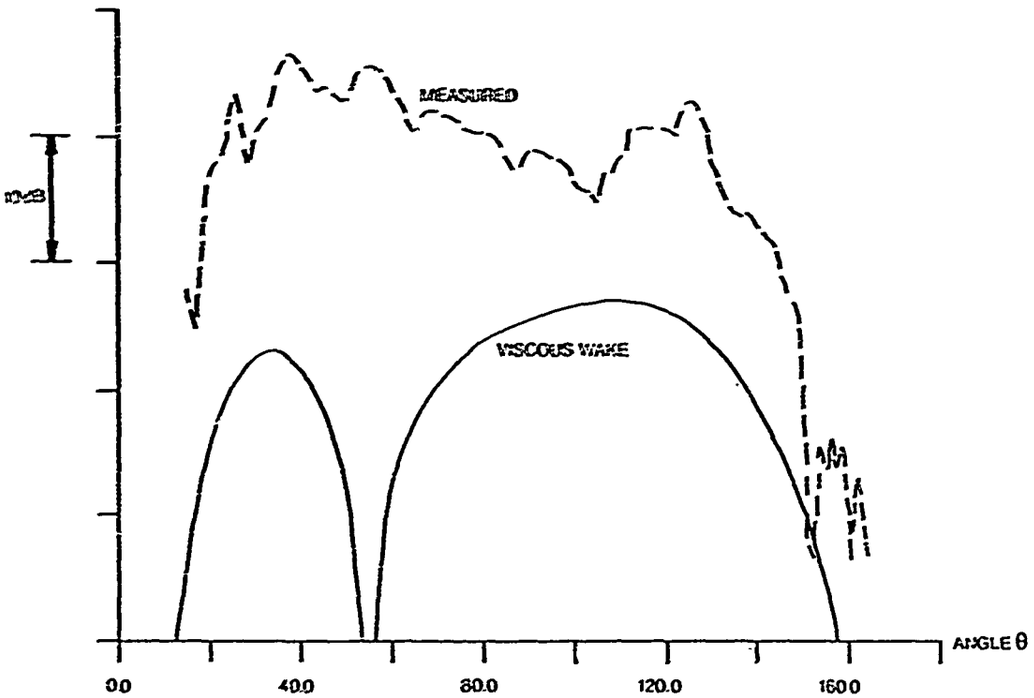


Figure 4. Gannet measurements vs. predicted wake interaction noise for the (1,2) interaction tone.

Airfoil response - upstream row

An analysis similar to that given in §2 shosws that the transformed velocity potential ψ, related to ϕ by (5), again satisfies the Helmholtz equation $\nabla^2\psi + K^2\psi = 0$. The boundary condition, however, is now given by

$$\frac{\partial\psi}{\partial Y} = -\exp(-i\kappa^* X) \quad \text{on} \quad Y = 0, \qquad -\infty \leq X \; 0 \tag{13}$$

where

$$\kappa^* = \mu + \frac{\sigma M^2}{1-M^2} ,$$

$$\mu = \gamma c/2. \tag{14}$$

For simplicity, since the pressure p is related to the transformed potential ϕ by (9), we define a transformed pressure

$$p^* = \left[i\kappa + \frac{\partial}{\partial X}\right]\phi. \tag{15}$$

The solution is obtained by the Wiener-Hopf technique and will be described here in detail. We define + and - Fourier transforms by

$$\Phi_{\pm}(s,Y) = \int_{-\infty}^{\infty} \phi(X,Y)H(\pm X)e^{isX}dX \tag{16}$$

where H(X) is the Heaviside unit function. The inverse transform is given by

$$\phi(X,Y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Phi(s,Y)e^{isX}\,ds. \tag{17}$$

On Fourier transforming the Helmholtz equation (6), we obtain

$$\Phi''(s,Y) + \chi^2\Phi(s,Y) = 0, \tag{18}$$

where we use primes to denote differentiation with respect to Y and

$$\chi = \sqrt{(K^2-s^2)}. \tag{19}$$

Here we choose the branch of the square root so that $\chi \to -i|s|$ as $s \to \pm\infty$. The branch cuts in the complex plane are shown in figure 5. The wavenumber K is taken to have a small imaginary part.

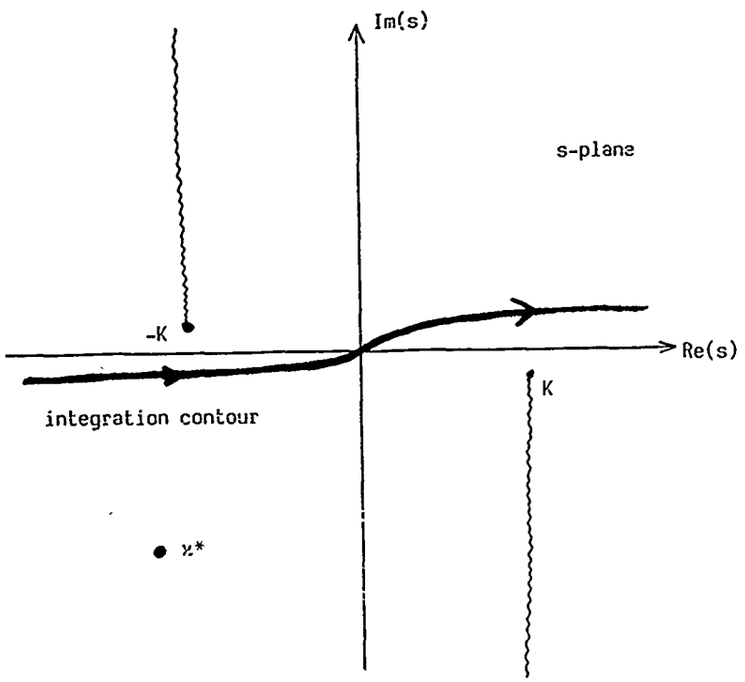


Figure 5. The integration contour and branch cuts in the complex plane.

We now consider the region X > 0 where we assume continuity of pressure across the wake. From (9) and (15) p^* is also continuous across the wake so that, on Fourier transforming (15), we obtain

$$i(\kappa - s)[\Psi_+(s,0-) - \Psi_+(s,0+)] = 0. \tag{20}$$

Since ψ must be an odd functin of Y this leads to

$$\Psi_+(s,0+) = 0. \tag{21}$$

If we now take the Fourier transform of p^* in the region X < 0 we obtain

$$-2i(\kappa - s)\Psi_-(s,0+) = \Delta P_-^*(s) \tag{22}$$

where we have again used the fact that ψ is odd in Y and ΔP_-^* represents the transform of the jump in p^* across the airfoil.

The boundary condition (13) can be Fourier transformed to give

$$\Psi_-'(s,0) = -\int_{-\infty}^{0} e^{-i\kappa^* X + isX} dX$$

$$= \frac{i}{s-\kappa^*} \tag{23}$$

where the integral converges provided that $\Im(s) < \Im(\kappa^*)$ and $\Im(z)$ denotes the imaginary part of z.

From (6) and the requirement that the fields decay at infinity we obtain

$$\Psi(s,Y) = C(s)e^{-i\chi Y} \tag{24}$$

where we have taken Y to be positive. (For Y negative we can again use the fact that ψ is odd in Y.) On differentiating (24) and setting Y = 0+ we obtain

$$\Psi'(s,0) = -i\chi\Psi(s,0+). \tag{25}$$

Substituting (21) and (23) into (25) we find that

$$\Psi_+'(s,0) + \frac{i}{s-\kappa^*} = i\chi\Psi_-(s,0+). \tag{26}$$

We now put $\chi = \sqrt{(K+s)(K-s)}$ so that, on dividing through by $\sqrt{(K-s)}$, (26) becomes

$$\frac{\Psi_+'(s,0)}{\sqrt{(K-s)}} + \frac{i}{(s-\kappa^*)\sqrt{(K-s)}} = i\sqrt{(K+s)}\Psi_-(s,0+). \tag{27}$$

Here the first term on the left-hand side is a + function and the right-hand side is a - function. The second term on the left-hand side can be split, in the usual way, into the sum of a + function and a - function. We thus rewrite (27) as

$$\frac{\Psi_+'(s,0)}{\sqrt{(K-s)}} + \frac{i}{(s-\kappa^*)}\left[\frac{1}{\sqrt{(K-s)}} - \frac{1}{\sqrt{(k-\kappa^*)}}\right]$$

$$= -i\sqrt{(K+s)}\Psi_-(s,0+) - \frac{i}{(s-\kappa^*)\sqrt{(K-\kappa^*)}} \tag{28}$$

where the left-hand side is a + function and the right-hand side is a - function. By the usual arguments resulting from Liouville's theorem, both sides of (28) are equal to an entire function E(s) which must be identically equal to zero (otherwise $\Psi_+'(s,0)$ would diverge as $s \to \infty$ implying that ψ has a singularity at the origin). From (22) and the right-hand side of (28) we then obtain

$$\Delta P_-^*(s) = \frac{2i(\kappa - s)}{(s-\kappa^*)\sqrt{(K-\kappa^*)(K+s)}}. \tag{29}$$

We now apply the inverse Fourier transform, as defined by (17), whence

$$\Delta p^*(X) = \frac{i}{\pi\sqrt{(k-\kappa^*)}} \int_{-\infty}^{\infty} \frac{(\kappa - s)}{(s-\kappa^*)\sqrt{(K+s)}} e^{-isX}\, ds. \tag{30}$$

On wrapping the integration contour around the branch cut in the upper half plane we find that

$$\Delta p^*(X) = \frac{2e^{iKX+i\pi/4}}{\pi\sqrt{(K-\kappa^*)}} \int_0^{\infty} \frac{(\kappa+K-is')}{\sqrt{s'}(s'+iK+i\kappa^*)} e^{Xs'}\, ds' - \frac{2(\kappa-\kappa^*)}{\sqrt{(K^2-\kappa^{*2})}} e^{-i\kappa^* X} \tag{31}$$

where the final term represents the contribution from the pole at $s = \kappa^*$. On evaluating the integral we obtain

$$\Delta p^*(X) = \frac{2(\kappa-\kappa^*)}{\sqrt{(K^2-\kappa^{*2})}}\left\{e^{iKX} w\left[i\sqrt{|X|(iK+i\kappa^*)}\right] - e^{-i\kappa^* X}\right\} + \frac{2e^{iKX-i\pi/4}}{\sqrt{[\pi(K+\kappa^*)|X|]}} \tag{32}$$

where $w(x) = e^{-x^2}\,\mathrm{erfc}\,(-ix)$ is the complex error functin (see Abramowitz & Stegun 1965).

We now impose a Kutta condition at the trailing edge of the airfoil. In the appendix we show that this is equivalent, in this case, to removing the inverse square-root singularity at the trailing edge. Then, from (9), the pressure jump across the airfoil is given by

$$\Delta p(X) = \frac{-2i\rho U(\sigma-\mu)}{\sqrt{(1-M^2)(i\overset{*}{\sigma}+i\mu)(i\tilde{\sigma}-i\mu)}}\left\{e^{i\overset{*}{\sigma}X}w\left[i\sqrt{-X(i\overset{*}{\sigma}+i\mu)}\right]-e^{-i\mu X}\right. \qquad (33)$$

where we have defined the normalised frequencies

$$\overset{*}{\sigma} = \frac{\sigma M}{(1-M)}, \qquad \tilde{\sigma} = \frac{\sigma M}{(1+M)}. \qquad (34)$$

Leading edge correction

In the previous section we dealt with just the trailing edge problem and consequently the leading edge effects have been neglected. Since leading edge effects can be important, as they are usually heavily loaded, we will discuss a correction to the present results to account for such effects.

We will use a technique developed by Landahl (1961) and Adamczyk (1974), and discussed by Amiet (1975), for downstream convected gust interactions. This involves an iterative technique for the solution of a 3 part boundary problem. The situation is shown in figure 6 with the upwash specified on $-2 \leq X \leq 0$.

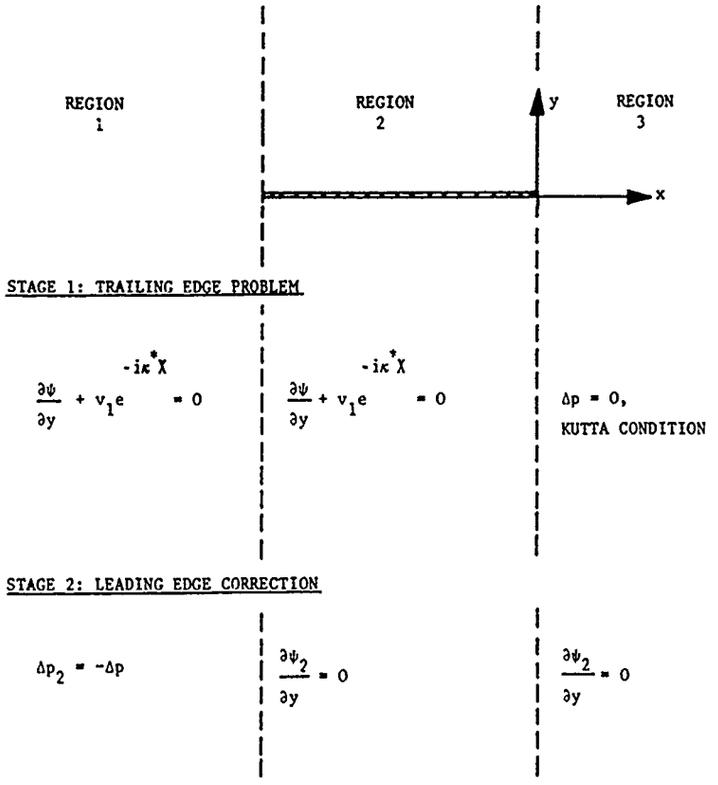


Figure 6. Three-part boundary value problem.

The first iteration involves solving the trailing edge problem with upwash specified on $-\infty < X \leq 0$; this is the case discussed in the previous section, i.e. we set $\Delta p_1 = \Delta p, \partial\psi/\partial Y = -\exp(-i\kappa^* X)$, on $Y = 0$. The second iteration involves correcting the upstream boundary condition on $-\infty < X \leq -2$ without affecting the boundary condition on the airfoil, i.e. on $-2 \leq X \leq 0$. We therefore require the 'new' pressure difference across $-\infty < X \leq -2$ to be minus that obtained on the first iteration, i.e. $\Delta p_2 = -\Delta p_1$ on $-\infty < X \leq -2$, and the upwash on $-2 \leq X < \infty$ to be zero.[6] This second iteration produces an error in

[6] We will use the suffix 2 throughout this section to denote 'second iteration' values.

the boundary condition on $0 \leq X < \infty$ which could be corrected by a third iteration and so on.

We define a new coordinate $\bar{X}$ centred on the airfoil leading edge so that $\bar{X} = X + 2$. The pressure jump Δp_2 across $-\infty < X \leq 0$ is minus that obtained in the previous section so that

$$\Delta p_2(\bar{X}) = -\Delta p(X) = -\Delta p(\bar{X}-2). \tag{35}$$

Then, from (9) and (15), the normalised pressures are related by a phase shift: $\Delta p_2^*(\bar{X}) = -\exp[-2i\sigma M^2/(1-M^2)]\Delta p^*(\bar{X}-2)$. From (32) this leads to[7]

$$\Delta p_2^*(\bar{X}) = -\frac{2(\kappa-\kappa^*)}{(K^2-\kappa^{*2})}\exp\left[-2i\frac{\sigma M^2}{1-M^2}\right]$$

$$\times\left\{e^{iK(\bar{X}-2)}w[i\sqrt{|\bar{X}-2|(iK+i\kappa^*)}]-e^{-i\kappa^*(\bar{X}-2)}\right\}. \tag{36}$$

Since we are considering a high frequency problem, σ and μ are both large. Then, from (7) and (14), the argument of the complex error function in (36) is also large on $\bar{X} < 0$ so that, from Abromowitz & Stegun (1965), we can use the approximation

$$w[i\sqrt{|\bar{X}-2|(iK+i\kappa^*)}] \sim \frac{1}{\sqrt{\pi|\bar{X}-2|(iK+i\kappa^*)}}. \tag{37}$$

(In order to use this approximation we have used the fact that $|\arg\sqrt{(iK+i\kappa^*)}| < \pi/2$.) In addition, since μ is large and $\Im(\mu) \neq 0$, the last term in braces in (36) is from (14), exponentially small on $\bar{X} < 0$; consequently, this term will be neglected. The jump in p_2^* across $\bar{X} < 0$ can

[7] Recall that the inverse square root singularity term has been removed in order to satisfy the Kutta condition.

therefore be approximated by

$$\Delta p_2^*(\bar{X}) \sim -\frac{2(\kappa-\kappa^*)}{\sqrt{\pi(iK+i\kappa^*)(K^2-\kappa^{*2})|X-2|}}\exp\left[-2i\frac{\sigma M^2}{1-M^2}+iK(\bar{X}-2)\right]. \qquad (38)$$

The boundary condition is that of no upwash so that

$$\frac{\partial \psi_2}{\partial \bar{Y}} = 0, \qquad \bar{X} > 0, \qquad (39)$$

where, for consistency, we have replaced Y with $\bar{Y}$.

The problem is therefore defined by the Helmholtz equation with the boundary condition (39) and the jump in p_2^* across $\bar{X} < 0$ given by (38). The solution is again obtained by the Wiener-Hopf technique with + and - transforms defined by (16) and the inverse transform defined by (17).

The Fourier transform $\Delta p_2^*(\bar{X})$ is given by

$$\Delta P_{2-}^*(s) = -\frac{2(\kappa-\kappa^*)}{\sqrt{\pi(iK+i\kappa^*)(K^2-\kappa^{*2})}}\exp\left[-2i\frac{\sigma M^2}{1-M^2}-2iK\right]\int_{-\infty}^{0}\frac{e^{i(s+K)\bar{X}}}{\sqrt{(2-\bar{X})}}\,dX. \qquad (40)$$

We now use the fact that K is large so that the integrand in (40) oscillates rapidly. Then the integral is dominated by contributions from close to X = 0 and, by following Murray (1974), we can approximate (40) to leading order in 1/(K+s), by

$$\Delta P_{2-}^*(s) = -\frac{2(\kappa-\kappa^*)}{\sqrt{\pi(iK+i\kappa^*)(K^2-\kappa^{*2})(s+K)}}\exp\left[-2i\left[\frac{\sigma M^2}{1-M^2}+K\right]\right]. \qquad (41)$$

From this point we proceed with the Wiener-Hopf technique in the usual manner and the pressure jump $\Delta p_2(X)$ is finally obtained as

$$\Delta p_2(X) = \frac{\sqrt{2}i\rho uU(\sigma-\mu)e^{i\tilde{\sigma}(1-X)-4i\bar{\sigma}}}{(i\mu+i\sigma^*)\sqrt{\pi(1-M^2)(i\tilde{\sigma}-i\mu)}}$$

$$\times \left\{ w[-e^{-i\pi/4}\sqrt{2\bar{\sigma}(1+X)}] - \frac{\sqrt{2}e^{-i\pi/4}}{\sqrt{\pi\hat{\sigma}}} \frac{1}{\sqrt{1+X}} \right\} \quad (42)$$

where we have reintroduced the normalised frequencies σ, μ, σ^* and $\tilde{\sigma}$ and, in addition, we have defined

$$\bar{\sigma} = \frac{M}{1-M^2}\sigma, \qquad \hat{\sigma} = \frac{(1+M)}{M(1-M)}\sigma. \quad (43)$$

We note, from (42), that we have restored the inverse square root singularity at the airfoil leading edge. The pressure distribution on the airfoil is given by the sum of (33) and (42). Comparing these equations we see that, apart from the exponential decay term $e^{-i\mu X}$ in p and the inverse square root singularity $1/\sqrt{(1+X)}$ in p_2, the first and second stage solutions are basically of the same form apart from the factor $1/\sqrt{(i\mu+i\sigma^*)}$ in p_2. Since we are considering a high frequency problem both μ and σ (and hence σ^*) are large. In the high frequency limit therefore the first correction to the trailing edge problem is $O(1/\sqrt{\sigma})$ smaller than the leading order term - even though the correction term includes the inverse square root singularity at the airfoil leading edge. We conclude that, in the high frequency limit, the semi-intinite airfoil model is a valid approximation and provides accurate results to leading order in σ.

Airfoil response - downstream row

The airfoil response calculation for the downstream blades is calculated in much the same way as for the upstream blades. The only difference is that now we have to solve a leading-edge problem instead of a trailing-edge

problem. The least singular solution is chosen, and has an inverse square root pressure singularity at the leading edge.

The Wiener-Hopf technique can be used, as before, to obtain the pressure difference across the airfoil (which now occupies $y = 0,\ 0 < x < \infty$) as

$$\Delta p(X) = \frac{-2\rho u U}{\sqrt{(1-M^2)(i\mu+i\sigma^*)}}$$

$$\times \left\{ \frac{i(\sigma-\mu)}{\sqrt{(-i\mu+i\sigma)}} \left[e^{-i\tilde{\sigma}X} w\left[i\sqrt{X(-i\mu+i\tilde{\sigma})} \right] - e^{-i\mu X} \right] - \frac{e^{-i\tilde{\sigma}X}}{\sqrt{\pi X}} \right\}. \qquad (44)$$

Measurement vs. prediction

As in §2 we combine the results given here for airfoil response with the appropriate noise radiation formulae and a model for the unsteady potential velocity field (Hanson 1985; Parry 1988) in order to obtain predictions of far-field noise which can be compared with the Gannet data. Note that here, unlike the case considered in §2, we have noise sources on both front and rear blade rows. The relative phasing of the two sources should, therefore, be corrected to account for spatial separation (as discussed by Hanson 1985). However we will, for the present, consider the two fields separately.

The first interaction tone generated by the Gannet is the (1,1) interaction. The far-field directivity of this tone is shown in figure 7.

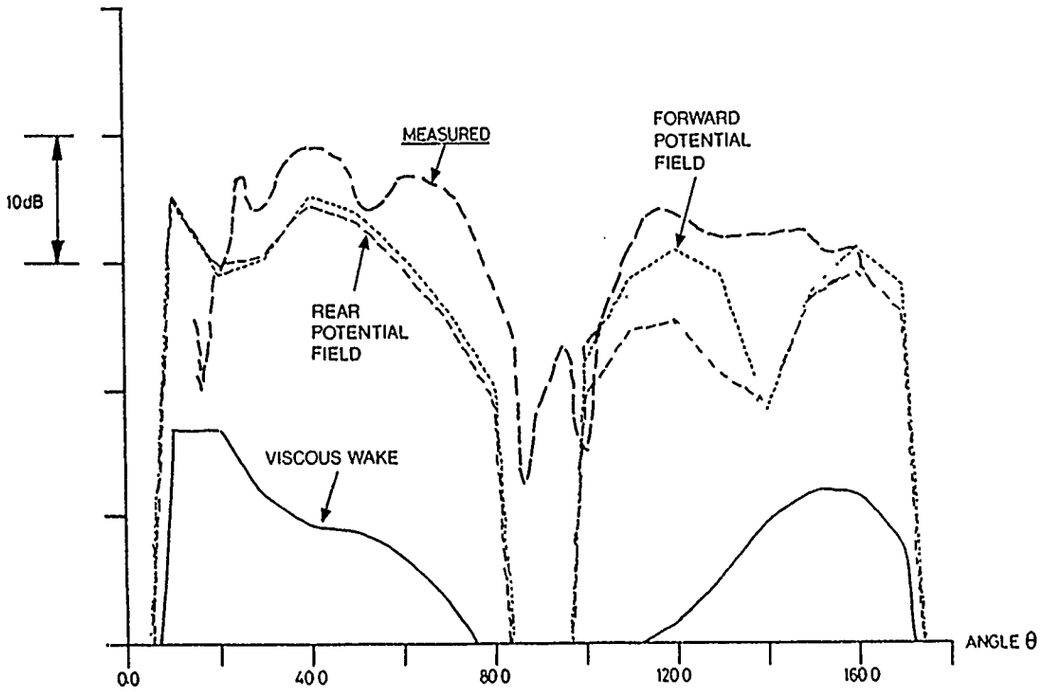


Figure 7. Gannet measurements vs. predicted wake and potential field interaction noise for the (1,1) interaction tone.

The first thing to note is that the predicted potential field interaction noise is significantly greater than the predicted wake interaction noise: the predicted forward and rearward potential field interaction tones are, typically, 20dB greater than the predicted wake interaction tone. The predicted potential field interaction noise agrees extremely well with the measured data in both forward and rear arcs - except, perhaps, for a discrepancy in the range 140 - 160 degrees,

The next two interaction tones generated by the Gannet are the (2,1) and (1,2) interaction tones, for which directivity plots are shown in figures 8 and 9. Here we see, again, that the predicted potential field interaction noise levels are significantly greater than the predicted wake interaction

noise levels - typically by 20dB (and more at some angles). Note that, as we remarked in §2, sources on the forward blade row (due to the upstream potential field interaction) generate far-field directivities different from those of sources on the rear blade row (due to the downstream row and potential field interactions).

We emphasise that the predictions, shown in figures 7 - 9, are **absolute level predictions** governed solely by the theoretical prediction scheme outlined above and dependent on the high-frequency approximation. The inputs for the calculation of the scattered fields are the incoming velocity fields: models for these are described by Parry (1988).

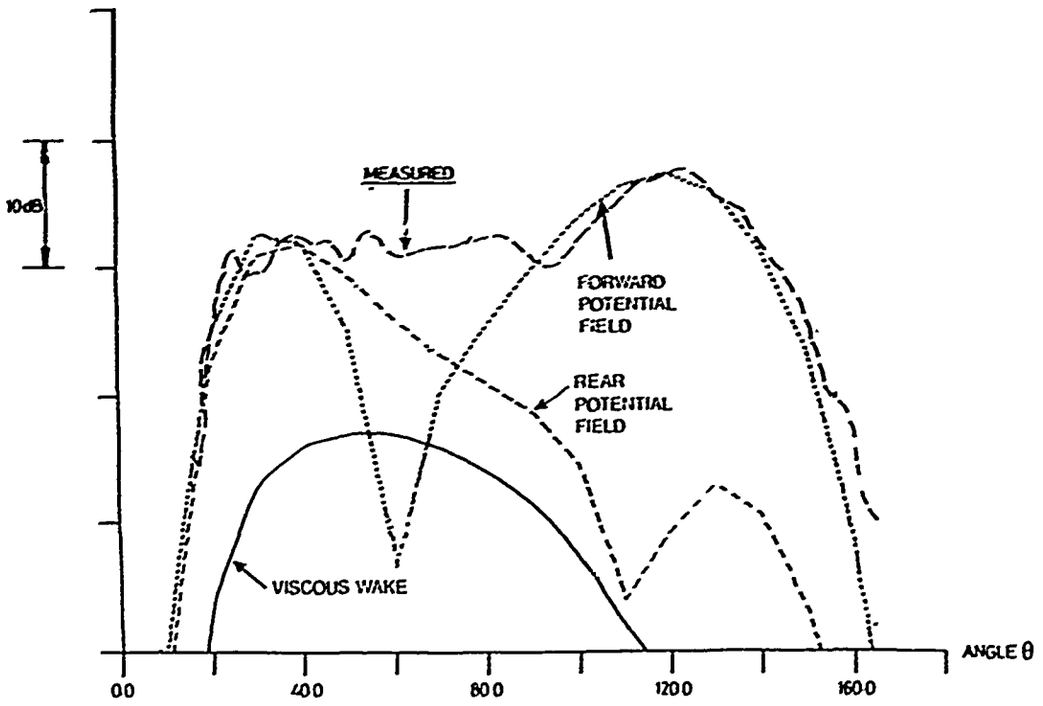


Figure 8. Gannet measurements vs. predicted wake and potential field interaction noise for the (2,1) interaction tone.

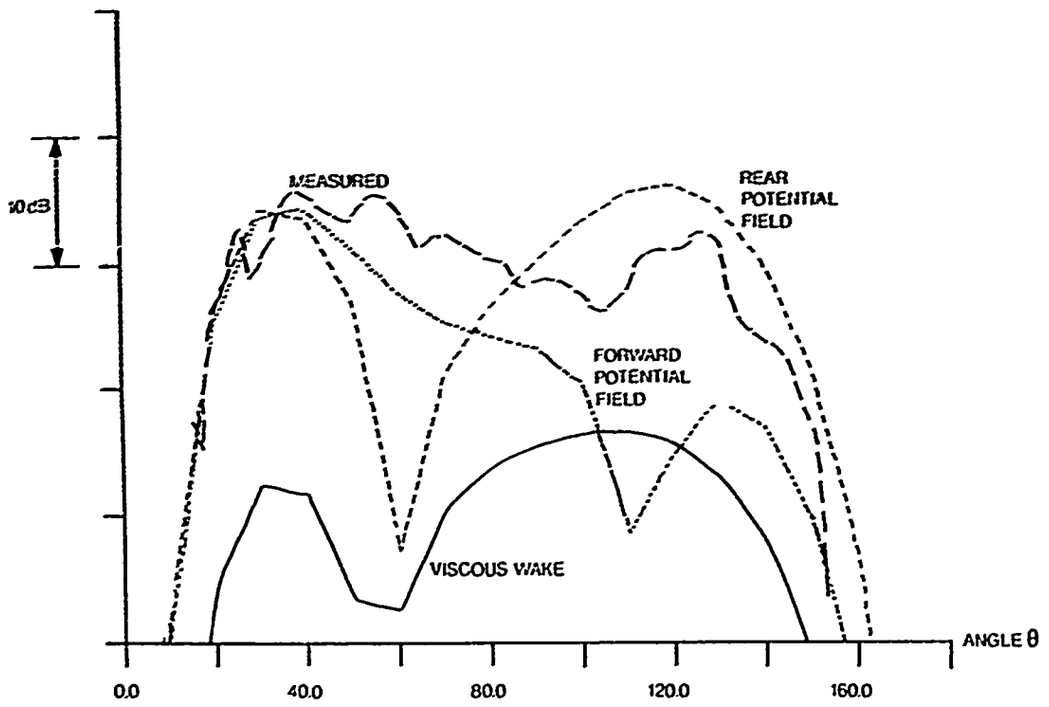


Figure 9. Gannet measurements vs. predicted wake and potential field interaction noise for the (1,2) interaction tone.

4. CONCLUSIONS

We have described a model for the calculation of unsteady velocity fields

scattered from airfoils in subsonic compressible flow at high reduced frequencies. The model covers interactions of leading edges or trailing edges with convected or nonconvected (potential) gusts. In addition, an iterative technique has been outlined by which the scattered field can be put in the form of an asymptotic series with successive terms decreasing by $O(1/\sigma)$, where σ is the reduced frequency.

Comparison with noise measurements taken from a Fairey-Gannet counter-rotation propeller has shown that the analysis produces extremely accurate results - in terms of both the absolute level and the far-field directivities - with no adjustment whatsoever of the theoretical predictions.

In addition, the comparison with measured data has shown that, for the Gannet, the downstream wake does not dominate the aerodynamic interactions, and that the potential flow field around each row generates significant and indeed dominant effects, both upstream and downstream.

REFERENCES

[1] Abramowitz, M. & Stegun, I. A. 1965 Handbook of Mathematical Functions. Dover, New York.

[2] Adamczyk, J. J. 1974 The passage of infinite swept airfoil through an oblique gust. NASA Contractors Report 2395.

[3] Amiet, R. K. 1975 Effects of compressibility in unsteady airfoil lift theories. Proc. Symposium Unsteady Aerodynamics. University of Arizona, 18 - 20 March, 1975.

[4] Bradley, A. J. 1986 A study of the rotor/rotor interaction tones from a contra-rotating propeller driven aircraft. AIAA Paper 86-1894.

[5] Crighton, D. G. 1977 Introduction to Wiener-Hopf methods in acoustics and vibration. David W. Taylor Naval Ship Research and Development Center Report 0112.

[6] Crighton, D. G. 1981 Acoustics as a branch of fluid mechanics. J. Fluid Mech. **106**, 261-298.

[7] Crighton, D. G. 1985 The Kutta condition in unsteady flow. Ann. Rev. Fluid Mech. 17, 411-445.

[8] Goldstein, M. E. 1976 Aeroacoustics. McGraw-Hill.

[9] Hanson, D. B. 1985 Noise of counter-rotation propellers. J. Aircraft 22, 609-617.

[10] Kemp, N. H. 1952 On the lift and circulation of airfoils in some unsteady flow problems. J. Aeron. Sci. 19, 713-714.

[11] Kemp, N. H. 1973 Closed form lift and moment for Osborne's unsteady thin-airfoil theory. A.I.A.A. J. 11, 1358-1360.

[12] Landahl, M. 1961 Unsteady Transonic Flow. Pergamon Press.

[13] Landau, L. D. & Lifshitz, E. M. 1959 Fluid Mechanics. Pergamon Press.

[14] Murray, J. D. 1974 Asymptotic Analysis. Oxford Univ. Press.

[15] Noble, B. 1958 Methods Based on the Wiener-Hopf Technique for the Solution of Partial Differential Equations. Pergamon Press.

[16] Osborne, C. 1973 Unsteady thin-airfoil theory for subsonic flow. A.I.A.A. J. 11, 205-209.

[17] Parry, A. B. 1988 Theoretical prediction of counter-rotation propeller noise. Ph.D. Thesis, University of Leeds.

[18] Parry, A. B. & Crighton, D. G. 1989 Prediction of counter-rotation propeller noise. A.I.A.A. Paper 89-1141.

[19] Sears, W. R. 1940 Some aspects of non-stationary airfoil theory and its practical application. J. Aeron. Sci. 8, 104-108.

[20] von Karman, T. & Sears, W. R. 1938 Airfoil theory for Non-Uniform motion. J. Aeron. Sci. 5, 379-390.

[21] Ward, G. N. 1955 Linearized Theory of Steady High-Speed Flow. Cambridge Univ. Press.

APPENDIX: THE KUTTA CONDITION

In our analysis the airfoils have been modelled as flat plates (finite or semi-infinite). As we have found, the pressure jump across the plate has inverse square root singularities at the edges of the plate. In order to alleviate the singularity at the trailing edge we must introduce a vortex sheet, extending to downstream infinity, across which the tangential velocities jump but the pressure is continuous. If the strength of the vortex sheet is fixed in order to cancel exactly the trailing edge singularity, then a Kutta condition is said to be satisfied.[8] The use of a Kutta condition in unsteady flow is a matter of controversy at the present. For the moment, however, we will assume that a Kutta condition is satisfied.

We look for a potential ψ_K which satisfies the Helmholtz equation (6)and is odd in Y. Then, downstream of the trailing edge, there is a jump in the potential across the vortex wake so that

$$\psi_K = \pm G e^{i\nu_K X} \quad \text{on} \quad Y = 0\pm, \quad X > 0, \tag{A1}$$

where ν_K and G are to be determined. Now $p^*_\kappa(X)$ is continuous across Y = 0 since $p_K(X)$ is continuous[9] so that (15) implies that $\nu_K = -\kappa$. Since there is

[8] A more detailed discussion of the Kutta condition in unsteady flow is provided by Crighton (1981, 1985).

[9] The definitions of p^*_K and p_K are the same as in §§2, 3 except that we have introduced a subscript K on those parameters relating to the velocity potential ψ_K.

no additional upwash the boundary condition is

$$\frac{\partial \psi_K}{\partial Y} = 0 \quad \text{on} \quad Y = 0, \quad X < 0. \tag{A2}$$

We now have, once again, a two-part boundary value problem which we will solve by using the Wiener-Hopf technique.[10] We then obtain, on $X < 0$,

$$\Delta p_K^*(X) = - \frac{2G\sqrt{(K+\kappa)}\, e^{iKX + i\pi/4}}{\sqrt{\pi}\sqrt{(-X)}}. \tag{A3}$$

This shows that the effect of the Kutta condition here is merely to remove the inverse square root singularity at the trailing edge, i.e. we select

$$G = \frac{i}{\sqrt{(K+\kappa)(K-\kappa^*)}} \tag{A4}$$

so that the sum of (32) and (A3) contains no term in $1/\sqrt{|X|}$.

A. B. Parry
Department of Mathematics
University of Strathclyde
Glasgow G1 1XH
Scotland, U.K.

[10] Crighton (1977, Chap. 9) shows how the Wiener-Hopf technique can be used to solve a trailing edge problem in unsteady compressible flow with a Kutta condition imposed.

M.K. PIDCOCK

Boundary problems in electrical impedance tomography

ABSTRACT

The reconstruction problem for Electrical Impedance Tomography is an extremely ill-posed non-linear inverse problem and the results obtained are highly sensitive to modelling and measurement errors associated with the technique. In this paper we continue our investigations into some simple problems which involve geometric errors and which can be solved by using perturbation theory. The sensitivity of the technique to these errors is made explicit and we describe one way of overcoming these difficulties which is suggested by the analysis.

1. INTRODUCTION

Electrical Impedance Tomography (EIT) is a technique of medical imaging which uses the contrast in the electrical conductivity of different body tissues to produce an image of the conductivity distribution within a part of the body. In many circumstances this image can be interpreted as a physical image and a number of potential medical applications of this technique are being investigated. The data used to obtain these images are measurements taken on the surface of an object of the electrical potential which are induced in the object by the application of known electrical currents to that surface.

Mathematically, the problem of EIT can be posed in the following way. Suppose that an object Ω with boundary $\partial\Omega$ consists of an isotropic Ohmic material with conductivity distribution σ. If ϕ is the electrical potential in Ω then

$$\nabla.(\sigma\nabla\phi) = 0 \qquad \text{in } \Omega$$

where

$$\phi \qquad \text{is known (measured) on } \partial\Omega$$

$$-\sigma \frac{\partial\phi}{\partial n} = j \qquad \text{is known (applied) on } \partial\Omega$$

It is important to note that one consequence of this model is that the applied current, j, satisfies

$$\int_{\partial\Omega} j \, ds = 0. \tag{1}$$

The aim of EIT is to determine σ from the boundary data. It has been shown [1-5] that this problem has a solution for many reasonable conductivity distributions and reconstruction algorithms have been described [6-11] which work well in simple cases and in the absence of errors. However, the reconstruction problem of EIT is a highly non-linear inverse problem and it has been demonstrated numerically [7] that the Fréchet derivative of the mapping which takes an applied current to a measured voltage has singular values which decay exponentially. The inversion procedure is therefore extremely sensitive to the possible errors associated with this technique.

There are a wide range of such errors which have to be considered. These range from basic deficiences in the mathematical model of the system to numerical errors introduced in the solution of equations in the reconstruction algorithm and data measurement errors defined by the instrument specification. Each has its own characteristics and a detailed study of them is essential if EIT is to become a useful diagnostic tool. The interaction of any reconstruction technique with a model of these errors represents one of the attractions of inverse problems.

In this paper we will consider just one type of error - that of imperfect knowledge of boundary shape. We will consider a class of simple problems

which can be solved using perturbation methods. This may be a little artificial but it does enable us to follow a parameter which characterises the boundary error in an explicit way. In EIT the electric current is applied to the surface via a number of electrodes positioned on the boundary but in this work we will assume that the current is adjustable at any point on the boundary.

In Section 2 we will describe the basic problem and the ideas behind our calculations and in Section 3 we will discuss a particular example of boundary error where the boundary is perturbed from its assumed shape in a simple way. In Section 4 we will extend our analysis to more general perturbations and suggest a scheme for identifying the parameters in the perturbation. We use polar coordinates (r,θ) throughout.

2. THE MODEL PROBLEM

Most of the studies in EIT have been concerned with the two-dimensional problem and a simple example often considered is that of distinguishing between two conductivity distributions σ_1, σ_2 on the unit disc, $0 \le r \le 1$ defined by

$$\begin{aligned} \sigma_1(r,\theta) &= 1 && 0 \le r \le 1 \\ \sigma_2(r,\theta) &= \sigma && 0 \le r \le R < 1 \\ &\quad 1 && R < r \le 1 \end{aligned}$$

It is easy to show that if the applied current $j(\theta)$ is given by

$$j(\theta) = \sum_{n=1}^{\infty} a_n \cos(n\theta) + b_n \sin(n\theta)$$

then

$$\phi(r,\theta;\sigma_1,j) = \sum_{n=1}^{\infty} \frac{r^n}{n}\left[a_n \cos(n\theta) + b_n \sin(n\theta)\right]$$

and

$$\phi(r,\theta;\sigma_2,j) = \sum_{n=1}^{\infty} \lambda_n r^n \left[a_n \cos(n\theta) + b_n \sin(n\theta)\right]$$

where

$$\lambda_n = \frac{1}{n}\left[\frac{1-\mu R^{2n} s^{2n}}{1+\mu R^{2n}}\right] \quad \text{with } s = \begin{matrix} r & (r>R), \\ R & (r<R) \end{matrix} \quad \text{and} \quad \mu = \frac{\sigma-1}{\sigma+1}\ .$$

As we are able to choose coefficients (a_n, b_n) in anyway, it is natural to ask if there are any combinations which are somehow better than others. Gisser et al [12] have suggested that if we wish to distinguish between σ_1 and σ_2 then the best normalised currents (i.e. $\|j\| = 1$) to use are those which maximise

$$\delta_j = \|\phi(r,\theta;\sigma_1,j) - \phi(r,\theta;\sigma_2,j)\|$$

where

$$\langle f,g\rangle = \frac{1}{\pi}\int_{\partial\Omega} fg\ ds \quad \text{and} \quad \|f\|^2 = \langle f,f\rangle.$$

In other words, we should try to maximise some average difference between the signals measured in the two cases. Other criteria for the choice of optimal currents have been proposed [13] and these will have different stability properties when interacting with the geometric errors considered in this paper. We have not yet studied these alternative currents and we confine our attention to those proposed in [12].

For the case of the problem described above, these optimal currents turn out to be trigonometric functions $j_m(\theta) = \cos(m\theta)$, $\sin(m\theta)$ for integer m, and the corresponding value of $\delta_m \equiv \delta_{j_m}$ is $\frac{2}{m}\frac{\mu R^{2m}}{(1+\mu R^{2m})}$. It is interesting to note that if this value of δ_m is less than the accuracy, E, of the measuring equipment, then σ_1 and σ_2 are not distinguishable using the current j_m since the value of δ_m could be an effect of noise. However, if $\delta_m > E$ then we can

distinguish between these two distributions. This is the basis of the calculations which we will report in the next section. We will calculate δ_m arising when these optimal trigonometric currents are applied to a body whose shape is thought to be the unit disc and which consists of material with a uniform conductivity distribution, σ_1. If the boundary were correctly known and all the measurements were accurate then the value of δ_m would be zero. The error in the boundary description will mean that a non-zero value of δ_m will be obtained and if $\delta_m > E$ we could, erroneously, infer that there is a circular anomaly of the form σ_2.

3. BOUNDARY DISTORTION

Consider the situation where the angular displacement of the drive electrodes is correct but the polar description of the boundary curve is $r = r(\theta)$ rather than $r = 1$. In this case we have that

$$\frac{\partial \phi}{\partial n} = \frac{1}{w(\theta)}\left[r(\theta)\,\frac{\partial \phi}{\partial r} - \frac{r'(\theta)}{r(\theta)}\,\frac{\partial \phi}{\partial \theta}\right] \tag{2}$$

and

$$\int_{\partial\Omega} j\,ds = \int_0^{2\pi} w(\theta) j(\theta)\, d\theta$$

where $w(\theta) \equiv \left[r^2(\theta) + r'^2(\theta)\right]^{1/2}$.

If $r(\theta) = 1 + \epsilon F(\theta)$ then we can write (2) in the form

$$\frac{\partial \phi}{\partial n}(1+\epsilon F(\theta),\phi) = \alpha_0 + \alpha_1 \epsilon + \alpha_2 \epsilon^2 + \alpha_3 \epsilon^3 + \ldots$$

where $\alpha_0, \alpha_1, \ldots$ are functions of θ and the partial derivatives of ϕ with respect to r and θ, evaluated at the point $(1+\epsilon F(\theta),\theta)$.

We can estimate the partial derivatives at this point by expanding these functions in a Taylor series about the point $(1,\theta)$. If we then write $\phi(r,\theta)$

in the form

$$\phi(r,\theta) = \phi^0(r,\theta) + \phi^1(r,\theta)\epsilon + \phi^2(r,\theta)\epsilon^2 + \phi^3(r,\theta)\epsilon^3 + \ldots$$

we find, finally, that

$$\frac{\partial\phi}{\partial n}(1+\epsilon F(\theta),\theta) = \beta_0 + \beta_1\epsilon + \beta_2\epsilon^2 + \beta_3\epsilon^3 + \ldots$$

where the functions $\beta_0,\beta_1,\ldots$ are complicated functions of θ and the partial derivatives of $\phi^0,\phi^1,\ldots$ with respect to r and θ evaluated at the point $(1,\theta)$.

Suppose that we have the simple case $r(\theta) = 1 + \epsilon\cos(\theta)$, i.e. $F(\theta) = \cos(\theta)$, then we find that

$$\int_{\partial n} j\ ds = \int_0^{2\pi}\left[1+2\epsilon\cos\theta+\epsilon^2\right]^{1/2} j(\theta)\ d\theta.$$

If we now try to apply the optimal currents $j_m(\theta)$ described earlier, we find that the condition (1) can be satisfied exactly only for $j_m(\theta) = \sin(m\theta)$, m an integer. So, for example, if we apply $j_1(\theta) = \sin(\theta)$ to the distorted boundary we find that at the point $(1,\theta)$

$$\beta_0 = \frac{\partial\phi^0}{\partial r_1}$$

$$\beta_1 = \frac{\partial\phi^1}{\partial r} + \cos\theta\,\frac{\partial^2\phi^0}{\partial r^2} + \sin(\theta)\,\frac{\partial\phi}{\partial\theta}$$

and similar, but more complicated, expressions for other β's.

Since ϕ satisfies Laplace's equation in Ω so too do $\phi^0,\phi^1,\ldots$. On the boundary, r = 1, we find that $\beta_0 = \sin(\theta)$ and $\beta_1 = \beta_2 = \ldots = 0$. After considerable computation it follows that

$$\begin{aligned} {}_1\Delta_1^\epsilon &= \|\phi^\epsilon(1+\epsilon\cos(\theta),\theta) - \cos(\theta)\| \\ &= \frac{\epsilon}{4}\left(1+\frac{8}{9}\epsilon^2 + \ldots\right) \end{aligned}$$

In general, if $j(\theta) = \sin(m\theta)$ then we find that

$$ {}_1\Delta_m^{\epsilon} = \frac{1}{|m^2-1|}\left[\frac{m^2+1}{2}\right]^{1/2} \epsilon + 0(\epsilon^3) \qquad m \neq 1 $$

We can generalise the type of perturbation slightly by considering boundaries defined by $r(\theta) = 1 + \epsilon \cos(k\theta)$ $(k = 1,2,\ldots)$. If we try to apply the optimal currents $j_m(\theta)$, we find that equation (1) restricts these currents to $\sin(m\theta)$ for m an integer and that

$$ k\ {}_k\Delta_m^{\epsilon} = \frac{\epsilon}{4} + 0(\epsilon^3) \qquad m=k $$

$$ = \frac{k}{|m^2-k^2|}\left[\frac{m^2+k^2}{2}\right]^{1/2} \epsilon + 0(\epsilon^3) \qquad m>k $$

$$ = \frac{k}{2}\left[\frac{1}{(m+k)^2} + \left[2 + \frac{1}{(k-m)}\right]^2\right]^{1/2} \epsilon + 0(\epsilon^3) \qquad m<k $$

In Table 1 we give values of ${}_k\Delta_m^{\epsilon}$ for a range of values of k and m. It is clear that there is a direct relationship between the detectable error in boundary shape and the measurement accuracy of the system. It is interesting to note that the observed behaviour gives a possible scheme for identifying displacements of this rather special type. It appears that if we apply currents of increasing spatial frequency all we need to do is to locate the dip in the values of Δ against frequency in order to identify k. Further comments on the interpretation of ${}_k\Delta_m^{\epsilon}$ can be found in Pidcock and Breckon [14].

k^m	1	2	3	4	5	6
1	0.25	0.53	0.28	0.19	0.15	0.12
2	3.02	0.25	1.02	0.53	0.36	0.28
3	3.77	4.51	0.25	1.52	0.77	0.53
4	4.68	5.01	6.01	0.25	2.01	1.02
5	5.64	5.84	6.26	7.51	0.25	2.51
6	6.61	6.76	7.01	7.51	9.00	0.25

Table 1. Values of $k\ {}_k\Delta_m^\epsilon$ for various values of k and m. All entries should be multiplied by ϵ.

4. GENERAL FOURIER PERTURBATIONS

It is interesting to note that for small ϵ, the relationship between Δ^ϵ and ϵ given the previous section is essentially linear and that it appears possible to perform simple experiments to determine the parameters of the perturbation once its general form is known. The relative ease of these tests encourages us to think that more general perturbations which can be described in terms of a Fourier expansion might be identified by a suitable series of such tests. Such an expansion should be very appropriate in the case of the human body where the relative smoothness of the body surface should lead to an economical description in terms of trigonometric series.

Consider, therefore the more general boundary perturbation given by

$$F(\theta) = \frac{1}{2}A_0 + \sum_{n=1}^{\infty} A_n \cos(n\theta)$$

If we try to apply one of the optimal boundary currents it is easy to see that (1) implies, once again that we must use $\sin(m\theta)$ for integer m. Following the analysis described in the previous section to first order in ϵ

we find that

$$\beta_0 = \frac{\partial \phi^0}{\partial r}(1,\theta) = \sin(m\theta)$$

$$\beta_1 = \frac{\partial \phi^1}{\partial r}(1,\theta) + F(\theta)\frac{\partial^2 \phi^0}{\partial r^2}(1,\theta) - F'(\theta)\frac{\partial \phi^0}{\partial \theta}(1,\theta) = 0$$

and it follows that

$$\phi^0(r,\theta) = \frac{r^m}{m}\sin(m\theta)$$

$$\phi^1(r,\theta) = -\sum_{k=1}^{\infty} \frac{r^k}{k} B_k^m \sin(k\theta).$$

where $B_k^M = \frac{1}{2}\left[(k-1)\, A_{|m-k|} + (k+1)\, A_{m+k}\right]$.

Hence, the potential actually measured differs from that expected if there were no boundary error by an amount

$$\delta\phi_m(\theta) = \phi^0(1+\epsilon F(\theta),\theta) + \epsilon\, \phi^1(1+\epsilon F(\theta),\theta) - \phi^0(1,\theta) + 0(\epsilon^2)$$

$$= \epsilon \sum_{k-1}^{\infty} c_k^m \sin(k\theta) + 0(\epsilon^2) \qquad (3)$$

where $c_k^m = \frac{1}{2k}\left[A_{|m-k|} - (2k+1)\, A_{m+k}\right]$.

We now have a possible scheme to determine the Fourier coefficients. It goes as follows. Apply a series of currents $\sin(m\theta)$ to the object and measure resulting voltages at a number of points. Use a Discrete Fourier Transform to express the measured voltage $\delta\phi_m(\theta)$ in terms of its Fourier components and use the above expression (3) as a system of linear equations to determine $\{A_k\}$. The implementation of this scheme and a detailed study of its numerical stability, together with an investigation into the effects of using only a finite number of electrodes, is the subject of future work in this area.

CONCLUSION

We have seen that the reconstruction problem for EIT is extremely sensitive to errors in the boundary shape. The analysis presented has, however, offered a possible way to overcome this problem. A study of this method will be presented elsewhere.

REFERENCES

[1] Kohn, R. V. and Vogelius, M. 1984 Commun. Pure Appl. Math. 37 289-298.

[2] Kohn, R. V. and Vogelius, M. 1984 SIAM-AMS Proc. 14 113-123.

[3] Kohn, R. V. and Vogelius, M. 1985 Commun. Pure Appl. Math. 39 644-667.

[4] Sylvester, J. and Uhlmann, G. 1986 Commun. Pure Appl. Math. 39 91-112.

[5] Sylvester, J. and Uhlmann, G. 1987 Annals of Math. 125 153-169.

[6] Breckon, W. R. and Pidcock, M. K. 1988 Some Topics on Inverse Problems ed P. C. Sabatier (World Scientific) 255-265.

[7] Breckon, W. R. and Pidcock, M. K. 1988 Information Processing in Medical Imaging ed C. N. de Graaf and M. A. Viergever (Plenum) 235-244.

[8] Breckon, W. R. and Pidcock, M. K. 1986 Proc. NATO-ASI 'Mathematics and Computer Science in Medical Imaging' 39 204-215.

[9] Yorkey, T. J. and Webster, J. G. 1987 Clin. Phys. Physiol. Meas. 8A 55-62.

[10] Barber, D., Brown, B. and Seager, A. 1985 Clin. Phys. Physiol. Meas. 6 109-121.

[11] Barber, D. and Seagar, A. 1986 Clin. Phys. Physiol. Meas. 8A 47-54.

[12] Gisser, D., Isaacson, D., and Newell, J. 1987 Clin. Phys. Physiol. Meas. 8A 39-46.

[13] Breckon, W. R., see this volume.

[14] Pidcock, M. K. and Breckon, W. R. Oxford Polytechnic Report.

M. K. Pidcock
Applied Analysis Research Group
Department of Computing and
Mathematical Sciences
Oxford Polytechnic
Headington, Oxford
England, U.K.

B.D. SLEEMAN

Interior and exterior inverse problems for the Helmholtz equation

1. INTRODUCTION

In this paper we discuss a number of recent developments relating to inverse problems for the Helmholtz equation. In particular we concentrate on the problem of determining the geometry of an unknown domain (e.g. vibrating membrane or scattering obstacle) from given data.

The paper is presented in two parts. In Part I we consider the classic inverse problem of determining an unknown domain Ω from a knowledge of the eigenvalues of the Laplacian defined in Ω. In §1 we survey the classic asymptotic estimates for the counting function $N(\lambda)$. Beginning with the fundamental results of Hermann Weyl we survey the most recent results for non smooth domains. In §2 we take up the Weyl-Berry conjecture regarding the asymptotics of $N(\lambda)$ for fractal domains. In particular we describe the recent contributions of Fleckinger and Lapidus.

Part II of the paper is concerned with the important inverse acoustic scattering problem. In §3 we formulate the direct scattering problem which provides the basic setting for the inverse problem. §4 discusses the central question of uniqueness of reconstruction of an unknown scattering obstacle from far field data.

In §5 we consider the "exterior" analogue of the asymptotics of $N(\lambda)$ by discussing the high frequency behaviour of the scattering phase $s(k)$. In particular we concentrate on the asymptotics of $s(k)$ for non-smooth domains, which complements the recent results of Melrose, and also fractal domains. This latter result provides the exterior analogue of the Weyl-Berry

conjecture. §6 returns to the practical problem of devising algorithms for the numerical reconstruction of the unknown scattering obstacle.

While this paper is largely expository, a number of the results are new and have not previously appeared in the literature. Theorem 2.3 is new as is Theorem 4.4 regarding uniqueness. The results embodied in Theorems (5.2), (5.3) and (5.4) are also new.

PART I

INTERIOR INVERSE PROBLEMS

§1 Geometry of the Counting Function

Let Ω be an arbitrary non empty bounded open connected set in $\mathbb{R}^n (n \geq 1)$ with boundary $\Gamma \equiv \partial\Omega$ and consider the eigenvalue problem

$$\begin{aligned} -\Delta u &= \lambda u \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \Gamma, \end{aligned} \tag{1.1}$$

where $\Delta \equiv \sum_{k=1}^{n} \partial^2/\partial x_k^2$ denotes the Dirichlet Laplacian in Ω. The parameter λ is said to be an eigenvalue of the problem (1.1) if there exists a $u \neq 0$ in $\overset{\circ}{H}_0^1(\Omega)$ satisfying $-\Delta u = \lambda u$ in the distribution sense. It is well known that the spectrum of (1.1) is discrete and consists of an infinite sequence of eigenvalues which may be ordered according to their multiplicity as

$$0 < \lambda_1 \leq \lambda_2 \ldots < \lambda_i \leq \ldots$$

where $\lambda_i \to \infty$ as $i \to \infty$.

In 1912 Weyl [47,48] established the classical result that

$$\lambda_i \sim C_n \left[\frac{i}{|\Omega|_n}\right]^{2/n} \quad \text{as } i \to \infty, \tag{1.2}$$

where $C_n = (2\pi)^2 (B_n)^{-2/n}$ depends only on the dimension 'n'. Here $|\Omega|_n$ denotes the n-dimensional Lebesgue measure or "volume" of Ω and B_n is the

volume of the unit ball in R^n.

Another way of estimating the asymptotic behaviour of the eigenvalues $\{\lambda_i\}_{i=1}^{\infty}$ is to work with the "counting function" $N(\lambda)$ defined as

$$N(\lambda) = \#\{i \geq 1, 0 < \lambda_i \leq \lambda\}, \text{ for } \lambda > 0. \tag{1.3}$$

Consequently Weyl's result can be restated as

$$N(\lambda) \sim (2\pi)^{-n} B_n |\Omega|_n \lambda^{n/2} \text{ as } \lambda \to \infty. \tag{1.4}$$

In a classic paper entitled "Can one hear the shape of a drum?" Kac [19] has asked the following question: Can someone with perfect pitch recover the precise shape of a drum just by listening to its fundamental tone and all the overtones? This question has motivated some important advances in the last two decades. In the first place, it is natural to ask whether the problem has a unique solution. Unfortunately, the answer appears to be no in general. Urakawa [44] has discovered two isospectral domains in R^n ($n \geq 4$) which are not isometric. Despite this it is possible to recover a lot of topological information about Ω from the spectrum of (1.1) and in particular from the counting function and other related functions. Indeed, if Γ is smooth (i.e. of class C^∞) then Seeley [36] and Pham The Lai [24] have shown that

$$N(\lambda) = (2\pi)^{-n} B_n |\Omega|_n \lambda^{n/2} + O(\lambda^{(n-1)/2}) \text{ as } \lambda \to \infty. \tag{1.5}$$

The proof of this result makes use of techniques from the theory of spectral transforms and of Fourier integral operators (c.f. Hormander [16]). More recently Ivrii [17] has shown that if Ω is a bounded domain with C^∞ boundary Γ and if $\bar{\Omega}$ does not have too many multiply reflected closed geodesics then

$$N(\lambda) = (2\pi)^{-n} B_n |\Omega_n| \lambda^{n/2} - C_n' |\Gamma|_{n-1} \lambda^{(n-1)/2} + o(\lambda^{(n-1)/2} \text{ as } \lambda \to \infty \tag{1.6}$$

where C_n' is a positive constant depending only on 'n'. There are a number of extensions of this result. For example Ivrii's result extends to the Neumann

problem, the only difference being that the minus sign of the second term on the right hand side of (1.6) is changed to a plus sign. Results are also known for the "impedance" or Robin boundary value problem [41] as well as for higher order positive elliptic operators with locally constant leading coefficients [12].

An alternative attack on the problem is to study the asymptotics as $t \to 0^+$ of the "partition function" (or trace of the heat semigroup)

$$Z(t) \equiv \int_0^\infty e^{-\lambda t}\, dN(\lambda) = \sum_{i=1}^{\infty} e^{-\lambda_i t} \tag{1.7}$$

which, when the integral exists, may provide more information than $N(\lambda)$. Thus for example if Γ is C^∞ and $n = 2$, McKean and Singer [31] and others have shown that

$$Z(t) = \frac{|\Omega|}{4\pi t} - \frac{|\Gamma|}{8(\pi t)^{1/2}} + \frac{1}{6}(1-h) + \sum_{i=1}^{\infty} C_i t^{i/2} \text{ as } t \to 0^+ \tag{1.8}$$

here h is the connectivity of Ω and the coefficients C_i are metric invariants. Indeed they are polynomials in the curvature of Γ and its derivatives.

Smith [42] has shown how to compute these coefficients using symbolic manipulation techniques.

If Γ is not smooth then neither (1.6) or (1.8) are expected to hold. For example if Ω has an outward pointing cusp then Waechter [46] has shown that (1.8) takes the form

$$Z(t) = \frac{|\Omega|}{4\pi t} - \frac{|\Gamma|}{8(\pi t)^{1/2}} + 0(t^{-\nu}),\ 0 < \nu < 1/2. \tag{1.9}$$

If we consider the Neumann problem then even the first term of (1.5) may not hold if the boundary Γ is "too long". This can be demonstrated in the following example due to Fleckinger and Metivier [11].

For a given positive number β define the set

$$\Omega_\beta = \left\{(x,y) \in R^2 | x \in (0,1), 0<y<1 + \sum_{j\in N} j^{-\beta} \psi_{I_j}(x)\right\}, \tag{1.10}$$

where $(I_j)j \in N$ is an infinite sequence of disjoint open intervals in (0,1) and where ψ_{I_j} denotes the characteristic function for the set I_j, i.e.

$$\psi_{I_j} = \begin{cases} 1, & x \in I_j \\ 0, & x \notin I_j \end{cases}$$

then for the Neumann problem we find that for $0 < \beta < 1/2$

$$N(\lambda) \approx \lambda^{1/2\beta} \text{ as } \lambda \to \infty \tag{1.11}$$

where $\approx$ means that there exist two positive constants C' and C" such that

$$C'\lambda^{1/2\beta} \leq N(\lambda) \leq C''\lambda^{1/2\beta}$$

for all λ sufficiently large.

In the following section we discuss the asymptotics of $N(\lambda)$ and $Z(t)$ in the extreme case when Γ is "fractal".

2. THE WEYL-BERRY CONJECTURE AND FRACTAL DOMAINS

In 1979 M. V. Berry [3] motivated by the study of the scattering of light by random surfaces conjectured that if Γ is "fractal" with Hausdorff dimension $H \in (n-1,n)$ then (1.6) takes the form

$$N(\lambda) = (2\pi)^{-n}B_n|\Omega|_n\lambda^{n/2} - C_{n,H}\mu_H(\Gamma)\lambda^{H/2} + o(\lambda^{H/2}) \text{ as } \lambda \to \infty, \tag{2.1}$$

here $C_{n,H}$ is a positive constant, depending only on n and H, and $\mu_H(\Gamma)$ denotes the H-dimensional Hausdorff measure of Γ.

Note that if Γ is sufficiently smooth then $H = n - 1$ and (2.1) reduces to (1.6). However in general Γ is very irregular and hence $H > n - 1$. Thus Berry's conjecture seems a reasonable one. However in (1986) Brossard and Carmona [4], through a series of illuminating examples, showed that (2.1) cannot be true in general and proposed that the Hausdorff measure (dimension) should be replaced by the less familiar Minkowski measure μ_δ (dimension δ).

In order to understand recent contributions to the Weyl-Berry conjecture we define the Hausdorff and Minkowski dimension as follows:-

Definition 2.1

Given $d \geq 0$, let $\mu_d(A) = \lim_{\epsilon \to 0^+} \left\{ \inf \sum_{i=1}^{\infty} r_i^d \right\}$, where the infimum is taken over all countable coverings of A by open balls $\{B_i\}_{i=1}^{\infty}$ of radius $r_i \leq \epsilon$. The number $\mu_d(A)$ in $[0,\infty)$ is called the d-dimensional Hausdorff measure of A and $H(A) = \inf\{d \geq 0, \mu_d(A) = 0\} = \sup\{d > 0, \mu_d(A) = \infty\}$ is called the Hausdorff Dimension of A.

Definition 2.2

Given $\epsilon < 0$ let $\Gamma_\epsilon = \{x \in R^n : d(x,\Gamma) \leq \epsilon\}$ be the ϵ-neighbourhood of Γ. For $d \geq 0$ let

$$\mu_d = \mu_d(\Gamma) = \lim_{\epsilon \to 0^+} \sup \epsilon^{-(n-d)} |\Gamma_\epsilon|_n \qquad (2.2)$$

be the d-dimensional upper Minkowski content of Γ. Then

$$D = D(\Gamma) = \inf\{d \geq 0, \mu_d(\Gamma) = 0\}$$
$$= \sup\{d \geq 0, \mu_d(\Gamma) = \infty\}$$

is called the Minkowski dimension of Γ.

If $0 < \mu_D < \infty$ then Γ is said to be Minkowski measurable and μ_D is called the Minkowski measure of Γ.

Examples

1. Let A be the set of rational numbers in [0.1], then $H(A) = 0$ and $D(A) = 1$.
2. Let A be the set $A = \bigcap_{i=0}^{\infty} K_i$, where K_i is the union of 2^i disjoint intervals of length α_i such that $\alpha_0 = 1$ and $\alpha_{i+1} < \alpha_i/2$. Then

$$H(A) = \lim_{i \to \infty} \inf \frac{\log 2^i}{\log(1/\alpha_i)}, \quad D(A) = \lim_{i \to \infty} \sup \frac{\log 2^i}{\log(1/\alpha_i)}.$$

If we have the classic 1/3-Cantor set then $\alpha_i = 3^{-i}$ and $H(A) = D(A) = \log 2/\log 3$.

3. Lapidus [25]

Let α be an arbitrary fixed positive number and let $\Omega \subset R^n$ be the bounded open set

$$\Omega = \bigcup_{i=1}^{\infty} (I_i \times J) \text{ where } J = (0,1)^{n-1},\ n \geq 1$$

and

$$I_i = ((i+1)^{-\alpha}, 1^{-\alpha})$$

Then

$$H(\Omega) = n - 1 \text{ and } D(\Omega) = n - 1 + (\alpha+1)^{-1}.$$

4. Brossard-Carmona [4]

Let $\Omega \subset \mathbb{R}^2$ be the countable disjoint union of all the small open cubes belonging to the successive generations defined as follows:-

Let $\{P_i\}_{i=1}^{\infty}$ be a nondecreasing sequence of positive integers. The 0-th generation contains 1 square of side 1. The 1-st generation contains 4 squares each of side 1/3 and is divided into $(P_1)^2$ congruent small squares. Similarly the i-th generation consists of $4\times5^{i-1}$ squares of side 3^{-i} and is divided into P_i^2 congruent small squares and so on.

Brossard and Carmona [4] show that irrespective of the sequence $\{P_i\}$ $H(\Omega) = \log 5/\log 3$. However if , for example, $P^i = [a^i]$ for same $a \geq 1$ and where [ ] indicates "integer part of" then $D(\Omega) = \log 5a^2/\log 3a$.

It is known and of course clear from the above examples that $H(A) \leq D(A)$. More than this, examples 3 and 4 show that while $H(A)$ is constant $D(A)$ is parameter dependent. The significance of this is that the Minkowski dimension is more sensitive to the "roughness" of the boundary of Ω than is the Hausdorff dimension.

Returning now to our main theme regarding the Weyl-Berry conjecture Lapidus and Fleckinger-Pelle [26] have proved

Theorem 2.1

If Ω is bounded and if $\partial\Omega$ is fractal with Minkowski dimension $\delta \in (n-1,n)$ then

$$N(\lambda) = (2\pi)^{-n} B_n |\Omega|_n \lambda^{n/2} + O(\lambda^{\delta/2}) \text{ as } \lambda \to \infty.$$

In addition this result has been extended by Lapidus [25] to more general elliptic operators and to the Neumann boundary value problem. More recently Fleckinger [10] has proved.

Theorem 2.2

If Ω is bounded and if $\partial\Omega$ is δ-Minkowski measurable with δ-Minkowski measure μ then

$$|N(\lambda) - (2\pi)^{-n} B_n |\Omega|_n \lambda^{n/2}| \leq C(n,\delta)\mu\lambda^{\delta/2} \text{ for all } \lambda \geq \lambda_0.$$

In order to give the reader an idea of the arguments used to establish results such as Theorems 2.1 and 2.2 we outline the proof of the following result concerning the asymptotics of the partition function $Z(t)$ for fractal domains in $\mathbb{R}^2$. This result is analagous to Theorem 2.2.

Theorem 2.3

If $\Omega \subset R^2$ is bounded and its boundary Γ is δ-Minkowski measurable with δ-Minkowski measure μ, then there exists a constant γ depending on δ so that

$$\left| Z(t) - \frac{|\Omega|}{4\pi t} \right| \leq \gamma\mu t^{-\delta/2} \qquad \text{for all } t \leq t_0.$$

In order to prove the theorem we need the following preliminary results.

(1) Dirichlet-Neumann Bracketing

Suppose Ω is a bounded open set in R^2, let $N_0(\lambda,-\Delta,\Omega)$ be the counting function for the Dirichlet Laplacian on Ω and $N_1(\lambda,-\Delta,\Omega)$ th counting function for the Neumann Laplacian on Ω. Then it is well known (c.f. Courant and Hilbert [9]) that the following proposition holds.

Proposition 2.1

If $\Omega' \subseteq \Omega$ then

$$N_0(\lambda,-\Delta,\Omega) \geq N_0(\lambda,-\Delta,\Omega'). \tag{2.3}$$

If Ω_1 and Ω_2 are two disjoint open sets in Ω with $\bar{\Omega} = \bar{\Omega}_1 \cup \bar{\Omega}_2$ then

$$\begin{aligned} N_0(\lambda,-\Delta,\Omega_1) + N_0(\lambda,-\Delta,\Omega_2) &\leq N_0(\lambda,-\Delta,\Omega) \\ &< N_1(\lambda,-\Delta,\Omega) \\ &< N_1(\lambda,-\Delta,\Omega_1) + N_1(\lambda,-\Delta,\Omega_2) \end{aligned} \tag{2.4}$$

(2) Polygonal Domains

Proposition 2.2 [45]

Let

(i) D be a polygonal domain with boundary ∂D

(ii) $P_1,\ldots,P_n$ be the vertices of ∂D and let W_i be the infinite wedge of angle γ_i with vertex P_i such that the boundary of the wedge contains the two edges adjacent to P_i.

(iii) Define for $y > 0$, $\gamma = \min \gamma_i$

$$B_i(y) = \{A \in W_i \mid d(A,P_i) < y\}$$

$$R = \frac{1}{2} \sup\{y \mid B_i(y) \cap B_j(y) = \phi \text{ for all } i \neq j, \bigcup_{k=1}^{n} B_k(y) \subset D\}$$

Then

$$\left|Z(t) - \frac{|D|}{4\pi t} + \frac{|\partial D|}{8(\pi t)^{1/2}} - \sum_{i=1}^{n} \frac{\pi^2 - \gamma_i^2}{14\pi\gamma_i}\right| \leq (5n+20\,\frac{|D|}{R^2})\frac{1}{\gamma} \times \frac{e^{-(R \sin \gamma/2)^2}}{16t}. \tag{2.5}$$

Note [5] that a similar but less precise result holds for the "Neumann" partition function together with a change of sign in the boundary term of (2.5).

Outline of the Proof of Theorem 2.3

To begin with we introduce a positive number ϵ_0 such that for all $\epsilon \in (0,\epsilon_0)$

$$\epsilon^{-(2-\delta)}|\Omega_\epsilon|_2 \leq 2\mu \tag{2.6}$$

and choose $P_o \in N$ such that $2^{-P_o} \leq \epsilon_o$. In (2.6) we define

$$\Omega_\epsilon = \{x \in R^2 | d(x,\partial\Omega) < \epsilon\} \tag{2.7}$$

where $d(.,\partial\Omega)$ denotes the Euclidean distance to the boundary $\partial\Omega$.

For each integer p we consider a tessellation of R^2 into congruent non-overlapping squares

$$Q_{\zeta_p}, \zeta_p \in Z^2 \text{ with side } \eta_p = 2^{-(p+P_o)}$$

Define

$$A_o = \left\{\zeta_o \in Z^2 | Q_{\zeta_o} \subset \Omega\right\}, \Omega_o' = \bigcup_{\zeta_o \in A_o} Q_{\zeta_o}$$

$$\Omega_o'' = \Omega \backslash \bar{\Omega}_o'.$$

$$A_1 = \left\{\zeta_1 \in Z^2 | Q_{\zeta_1} \subset \Omega_o''\right\}, \Omega_1' = \Omega_o' \cup \left[\bigcup_{\zeta_1 \in A_1} Q_{\zeta_1}\right],$$

$$\Omega_1'' = \Omega \backslash \bar{\Omega}_1',$$

$$A_p = \left\{\zeta_p \in Z^2 | Q_{\zeta_p} \subset \Omega_{p-1}''\right\}, \Omega_p' = \Omega_{p-1}' \cup \left[\bigcup_{\zeta_p \in A_p} Q_{\zeta_p}\right],$$

$$\Omega_p'' = \Omega \backslash \bar{\Omega}_p'.$$

We also define the boundary sets

$$B_p = \left\{\zeta_p \in Z^2 | Q_{\zeta_p} \cap \partial\Omega \neq 0, Q_{\zeta_p} \cap \Omega_p' = \phi\right\},$$

$$R_p = \bigcup_{\zeta_p \in B_p} Q_{\zeta_p}.$$

We now make the following observations and estimates

01) $\Omega_p'' \subset \Omega_{\epsilon_p}$ with $\epsilon_p = \sqrt{2}\eta_p = \sqrt{2}.2^{-(p+P_o)}$ where Ω_{ϵ_p} is defined by (2.7).

Furthermore $R_p \subset \Omega_{\epsilon_p}$.