

***Image reconstruction in electrical impedance
tomography***

Breckon, William Robert

1990

MIMS EPrint: **2008.48**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: <http://eprints.maths.manchester.ac.uk/>

And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

**Image Reconstruction in Electrical Impedance
Tomography
Ph.D. Thesis
William Robert Breckon
Oxford Polytechnic
May 1990**

Over the years I have been asked to provide copies of this thesis. When I wrote the thesis my ability, and the technology available at the time, made it difficult to include graphics electronically and I do not have the electronic sources for them. Indeed they were produced either by Fortran libraries long since neglected or hand drawn.

This limitation was a blessing in that requests for the thesis were directed to me personally rather than an anonymous server, and I was able to make personal contact with each person, often young graduate students new to the field, who requested a paper copy. I provided this in exchange for a brief account of why they were interested in EIT and what they intended to do. Some notable recipients of these paper copies were Andrea Borsic in Italy, and Manuch Soleimani in Iran. Whether despite or because of my encouragement and thesis they both went on to make significant contributions, both constructing EIT systems and implementing working reconstruction algorithms at Masters level in their own countries. At a time when these tasks were far from routine. I had the privilege to supervise them both for their PhDs and they continue to contribute to this and other areas of inverse problems.

With the advent of a photocopier in the School of Mathematics at Manchester capable of scanning hundreds of sheets of paper and emailing the portable document format file, my first thought was to scan this thesis. Looking over it it all seems very dated and we have moved on so much, but it is still requested by those new to the field. I hope the reader will gain some benefit and forgive any errors they find. Now that the thesis is available on the preprint server I will miss hearing from those starting out in EIT. So please feel free to contact me anyway and tell me what you are doing or planning to do in this area.

At the time of writing my contact details are on the web site <http://www.maths.manchester.ac.uk/~bl>, and this thesis can be found on the MIMS e-print server by searching <http://eprints.ma.man.ac.uk/>

Professor William Robert Breckon Lionheart
School of Mathematics
The University of Manchester
April 15th 2008

Abstract

Image Reconstruction Electrical Impedance Tomography

W.R. Breckon

This thesis is concerned with Electrical Impedance Tomography (EIT), a medical imaging technique in which pictures of the electrical conductivity distribution of the body are formed from current and voltage data taken on the body surface. The focus of the thesis is on the mathematical aspects of reconstructing the conductivity image from the measured data (the reconstruction problem). The reconstruction problem is particularly difficult and in this thesis it is investigated analytically and numerically. The aim of this investigation is to understand why the problem is difficult and to find numerical solution methods which respect the difficulties encountered. The analytical investigation of this non-linear inverse problem for an elliptic partial differential equation shows that while the forward mapping is analytic the inverse mapping is discontinuous. A rigorous treatment of the linearisation of the problem is given, including proofs of forms of linearisation assumed by previous authors. It is shown that the derivative of the forward problem is compact. Numerical calculations of the singular value decomposition (SVD) are given including plots of singular values and images of the singular functions. The SVD is used to settle a controversy concerning current drive patterns. Reconstruction algorithms are investigated and use of Regularised Newton methods is suggested. A formula for the second derivative of the forward mapping is derived which proves too computationally expensive to calculate. Use of Tychonov regularisation as well as filtered SVD and iterative methods are discussed. The similarities, and differences, between EIT and X-Ray Computed Tomography (X-Ray CT) are illuminated. This leads to an explanation of methods used by other authors for EIT reconstruction based on X-Ray CT. Details of the author's own implementation of a regularised Newton method are given. Finally the idea of adaptive current patterns is investigated. An algorithm is given for the experimental determination of optimal current patterns and the integration of this technique with regularised Newton methods is explored. Promising numerical results from this technique are given. The thesis concludes with a discussion of some outstanding problems in EIT and points to possible routes for their solution. An appendix gives brief details of the design and development of the Oxford Polytechnic Adaptive Current Tomograph.

Image Reconstruction

in

**Electrical Impedance
Tomography**

by

W.R. Breckon

**School of Computing and Mathematical Sciences
Oxford Polytechnic**

A thesis submitted to the Council for National Academic Awards in partial
fulfilment of the requirements
for the degree of

Doctor of Philosophy

May 1990

Copyright © 1992 by W. R. Breckon

All rights are reserved. No part of the contents of this book may be reproduced or transmitted in any form or by any means without written permission of the author, (which will usually be given).

Preface to second printing

There have been many changes since the first printing of this thesis in 1990. Both the institution and myself have changed our names. Oxford Polytechnic is now Oxford Brookes University. My wife Sarah and I both changed our surname to Lionheart before the birth of our daughter Katie in 1993.

At Brookes University EIT research has continued apace. We are now on our third design of adaptive current tomograph (OXBACT 3) which has produced promising *in vivo* images. This work has been generously supported by the Wellcome trust. we anticipate doing some clinical studies next year. Our hope is that we will be able to detect lung water. We are also constructing a three dimensional EIT system and a cylindrical tank. This work is supported by the Higher Education Funding Council for England.

On the mathematical side we have new algorithms (for example Kevin Paulson's POMBUS). I have been looking into the anisotropic EIT problem and have some results. Mike Pidcock with Sorin Ciulli and Simona Ispas has been investigating singularities in the potential on the boundary with realistic electrode models.

I have attached a list of publications and papers in print

William R. B. Lionheart

November 1995

School of Computing and mathematical Sciences

Oxford Brookes University

Gipsey Lane

Headington OX3 0BP

Tel (0)1865 483677 (office) 483682 (EIT lab) 483666 (fax)

Email p0054865@brookes.ac.uk

Recent Publications:

1990

- Breckon, W R, 'Image Reconstruction in Electrical Impedance Tomography', Oxford Polytechnic Doctoral Thesis (1990).
- Lidgey, F J, Vere-Hunt, M A, Toumazou, C, 'Developments in current driver circuitry', Proc. Third EC COMAC-BME meeting on EIT, Copenhagen (1990), 183-189.
- McLeod, C N, Murphy, D and Breckon W R, 'OXPACT: Developments in high speed data acquisition', Proceedings of Third EC COMAC-BME meeting on EIT, Copenhagen (1990).
- Paulson, K S and Pidcock M K, 'Solving symmetric matrix problems on rings of transputers', Rutherford Appleton Laboratory Report (1990).
- Paulson, K S, Breckon, W R and Pidcock, M K, 'The Importance of Electrode Modelling Electrical Impedance Tomography', Proc. Third EC COMAC-BME meeting on EIT, Copenhagen, 84-96, (1990).
- Paulson, K S, Breckon, W R and Pidcock, M K, 'Concurrent EIT reconstruction', Proc. Third EC COMAC-BME meeting on EIT, 136-143, Copenhagen (1990).
- Pidcock, M K, 'Recent developments in Electrical Impedance Tomography', Inverse Methods in Action, Ed. Sabatier P.C., 38-46, Springer-Verlag (1990).
- Pidcock, M K, 'Computational Issues in Electrical Impedance Tomography', Bulletin of the Hellenic Mathematical Society, Vol 31, 103-117 (1990).

1991

- Breckon, W R, Paulson, K S and Pidcock, M K, 'Parallelism in EIT Reconstruction', Information Processing in Medical Imaging, Progress in Clinical and Biological Research, Vol 363, Ed. Ortendahl, D.A. and Llacer, J., 187-196, Wiley-Liss (1991).
- Breckon, W R, 'Measurement and Reconstruction in EIT', Inverse Problems and Imaging, Research Notes in Mathematics 245, Ed Roach G., 130-140, Pitman(1991).
- Pidcock, M K, 'Boundary problems in Electrical Impedance Tomography', Inverse Problems and Imaging, Research Notes in Mathematics 245, Ed Roach G., 155-166, Pitman(1991).

1992

- Breckon, W R, Paulson, K S and Pidcock, M K, 'An iterative EIT reconstruction algorithm using optimal data', Proceedings of the EC workshop on EIT reconstruction algorithms, Toulouse, (1992).
- Breckon, W R, 'The problem of Anisotropy in Electrical Impedance Tomography', Proceedings of 14th International Conference of the IEEE Engineering in Medicine and Biology Society, Vol 5, 1734-1735, Paris, (1992).
- Denyer, C W, Lidgey, F J, McLeod, C N and Zhu, Q, 'High output impedance voltage controlled current source for Bio-impedance', Proceedings of 15th International Conference of the IEEE Engineering in Medicine and Biology Society, 1026-1027, Paris, (1992).
- Paulson, K S, 'Parallel Algorithms for Three Dimensional Electrical Impedance Tomography' Oxford Brookes University PhD Thesis, (1992).
- Lidgey F J, Zhu Q S, McLeod, C N and Breckon W R, 'Electrode current determination from programmable voltage sources', Clinical Physics and Physiological Medicine, Vol 13A, 43-47, (1992).
- Lionheart, W R, Paulson, K S and Pidcock, M K, 'Iterative Reconstruction Algorithms in Electrical Impedance Tomography', Proceedings of the IEE Conference on Electrical Impedance Tomography, October 1992.
- McLeod, C N, Lidgey, F J and Zhu, Q, 'Multiple drive Electrical Impedance Tomography', Proceedings of the IEE Conference on Electrical Impedance Tomography, October 1992.
- Paulson, K S, Breckon, W R, Pidcock, M K, 'Electrode modelling in Electrical Impedance Tomography', SIAM Journal of Applied Mathematics, Vol 52, 1012-1022, (1992).
- Paulson, K S, Breckon, W R, Pidcock, M K, 'Optimal Experiments in Electrical Impedance Tomography', IEEE Transactions on Medical Imaging, Vol 12, No 2, 681-686, (1993).
- Paulson, K S, Breckon, W R and Pidcock, M K, 'Optimal measurements in Electrical

Impedance Tomography 'Proceedings of 14th International Conference of the IEEE Engineering in Medicine and Biology Society, Vol 5, 1730-1731, Paris, (1992).

Paulson K S, Breckon W R and Pidcock M K, 'A hybrid phantom for Electrical Impedance Tomography', Clinical Physics and Physiological Medicine, Vol 13 A, 155-161,(1992).

Pidcock, M K and Paulson, K S, 'Current density distributions on electrodes', Proceedings of 14th International Conference of the IEEE Engineering in Medicine and Biology Society, Vol 6, 2386-2387, Paris, (1992).

Zhu, Q, 'High precision Electrical Impedance Tomography Instrumentation', Oxford Brookes University PhD Thesis, (1992).

Zhu, Q, Lidgey, F J and Vere-Hunt, M A, 'Improved wide-band, high CMRR instrumentation amplifier', Clinical Physics and Physiological Medicine, Vol 13 A, 51-57, (1992).

Zhu, Q, Breckon, W R, Lidgey, F J and McLeod, C N, 'A voltage driven Adaptive Current Electrical Impedance Tomograph', Proceedings of 14th International Conference of the IEEE Engineering in Medicine and Biology Society, Vol 5, 1704-1705, Paris, (1992).

Zhu, Q, Denyer C W, Lidgey, F J, Lionheart, W R and McLeod C N, 'A serial data acquisition architecture for continuous Impedance Imaging', Proceedings of 15th International Conference of the IEEE Engineering in Medicine and Biology Society, 1024-1025, Paris, (1992).

1993

Paulson, K S, Lionheart, W R and Pidcock, M K, 'Fast, non-linear inversion for Electrical Impedance Tomography', Information Processing in Medical Imaging, Ed. Barrett. and Gmitro, A F, 244-258, Springer (1993).

Paulson, K S, Lionheart, W R and Pidcock, M K, 'Optimal Experiments in Electrical Impedance Tomography', IEEE Journal of Medical Imaging, Vol 12, No 4, December, (1993)

Pidcock, M K, Lidgey, F J, Lionheart, W R, McLeod, C N, Paulson, K S and Zhu, Q, 'POMPUS-A Fast Reconstruction Algorithm for Electrical Impedance Tomography', Proceedings of ECAPT '93, Karlsruhe, (1993).

Zhu, Q, Lidgey, F J, Lionheart, W R, McLeod, C N, Paulson, K S and Pidcock, M K, 'A Novel Multiple-Drive EIT System', Proceedings of ECAPT '93, Karlsruhe, (1993).

Zhu, Q, Lidgey, F J and Su, W J, 'Second generation current mode instrumentation amplifiers' Proc. IEEE Symposium on Circuits and Systems, Chicago, Vol2, 1326-1328, (1993).

Zhu, Q, McLeod, C N, Breckon, W R, Lidgey F J, Paulson, K S and Pidcock, M.K, 'An Adaptive Current Tomograph using voltage sources', IEEE Transactions on Biomedical Engineering, Vol 40, No 2,(163-168), (1993).

1994

Denyer, C W, Lidgey, F J, McLeod, C N and Zhu, Q, 'High output impedance current source', Physiological Measurement, Vol 15, Supp 2A, 79-83.(1994).

Paulson, K S, Lionheart, W R and Pidcock, M K, 'A fast inversion for Electrical Impedance Tomography', Image and Vision Computing, Vol 12, No 6, 367-373, (1994).

Pidcock, M K, Paulson, K S and Lionheart, W R, 'Fast non-linear inversion for Electrical Impedance Tomography', Information Processing in Medical Imaging, Ed Barrett, H and Gmitro A, 244-258, Springer, (1994).

Pidcock, M K, Paulson, K S and Lionheart, W R, 'Optimal Experiments in Electrical Impedance Tomography' Transactions on Medical Imaging, Vol 12, No 4, 681-686, (1994).

Pidcock, M K, 'Analysing the Importance of Electrode Modelling in Electrical Impedance Tomography' Proc. ECAPT '94, 285-290, Oporto, (1994).

Zhu, Q, Denyer, C W, Lidgey, F J, Lionheart, W R and McLeod, C N, 'Development of a real-time adaptive current tomograph', Physiological Measurement, Vol 15, Supp 2A, 37-45 (1994).

1995

Paulson, K S, Lionheart, W R and Pidcock, M K, 'POMPUS: An optimised EIT reconstruction algorithm', Inverse Problems, Vol 11, 425 (1995).

Pidcock, M K, Kuzuoglu, M and Leblebicioglu, K, 'Analytical and semi-analytical solutions in Electrical Impedance Tomography I: Two dimensional problems', *Physiological Measurement*, Vol 16, 77-90 (1995).

Pidcock, M K, Kuzuoglu, M and Leblebicioglu, K, 'Analytical and semi-analytical solutions in Electrical Impedance Tomography II: Three dimensional problems', *Physiological Measurement*, Vol 16, 91-110 (1995).

Pidcock, M K, Ciulli, S and Ispas, S, 'Singularities of mixed boundary value problems for Electrical Impedance Tomography', to be published in *Physiological Measurement* (1995).

Pidcock, M K, Ciulli, S and Ispas, S, 'Aspects of electrode modelling in Electrical Impedance Tomography' to be published in *Computational Methods in Applied Mathematics* (1995).

Pidcock, M K, Ciulli, S and Ispas, S, 'Singularities of solutions to a mixed Neumann-Robin boundary value problem' submitted for publication in the *Journal of Mathematical Physics* (1995).

Pidcock, M K, Ciulli, S and Ispas, S, 'Numerical modelling of a mixed Neumann-Robin boundary value problem' submitted for publication (1995).

Lionheart, W R B Geometric Results in Anisotropic Electrical Impedance Imaging, submitted to *Inverse Problems* (1995)

The followin papers were given at IX International Conference on Electrical Bio-Impedance and European Concerted Action on Impedance Tomography in Heidelberg 26-30 Sept . We expect them to appear in a special edition of *Physiological Measurement*.

Paulson KS Pidcock MK McLeod CN An Impedance Measurement Probe for Live Tissue

McLeod CN Shi Y Denyer CW Lidgley FJ Lionheart WRB Paulson KS Pidcock MK Chest Impedance Using Trigonometric Current pattern

Shi Y Denyer CW McLeod CN Lidgley FJ High Speed Adaptive Current Setting - Calibration and in vivo results

Lionheart WRB Uniqueness of Solution for the Anisotropic Electrical Impedance Inverse Problem

Pidcock MK Ciulli S Ispas S Some Boundary value problems for electrical Impedance Tomogrphy

Image Reconstruction in Electrical Impedance Tomography

W.R. Breckon

This thesis is concerned with Electrical Impedance Tomography (EIT), a medical imaging technique in which pictures of the electrical conductivity distribution of the body are formed from current and voltage data taken on the body surface. The focus of the thesis is on the mathematical aspects of reconstructing the conductivity image from the measured data (the reconstruction problem).

The reconstruction problem is particularly difficult and in this thesis it is investigated analytically and numerically. The aim of this investigation is to understand why the problem is difficult and to find numerical solution methods which respect the difficulties encountered.

The analytical investigation of this non-linear inverse problem for an elliptic partial differential equation shows that while the forward mapping is analytic the inverse mapping is discontinuous. A rigorous treatment of the linearisation of the problem is given, including proofs of forms of linearisation assumed by previous authors. It is shown that the derivative of the forward problem is compact. Numerical calculations of the singular value decomposition (SVD) are given including plots of singular values and images of the singular functions. The SVD is used to settle a controversy concerning current drive patterns.

Reconstruction algorithms are investigated and use of Regularised Newton methods is suggested. A formula for the second derivative of the forward mapping is derived which proves too computationally expensive to calculate. Use of Tychonov regularisation as well as filtered SVD and iterative methods are discussed. The similarities, and differences, between EIT and X-Ray Computed Tomography (X-Ray CT) are illuminated. This leads to an explanation of methods used by other authors for EIT reconstruction based on X-Ray CT. Details of the author's own implementation of a regularised Newton method are given.

Finally the idea of adaptive current patterns is investigated. An algorithm is given for the experimental determination of optimal current patterns and the integration of this technique with regularised Newton methods is explored. Promising numerical results from this technique are given.

The thesis concludes with a discussion of some outstanding problems in EIT and points to possible routes for their solution. An appendix gives brief details of the design and development of the Oxford Polytechnic Adaptive Current Tomograph.

Acknowledgements

I would like to express my thanks to my Director of Studies, Mike Pidcock, and to my Supervisors David Barber and Chris Budd, for their help, criticism, encouragement and support over the four-and-a-half years this work has taken to complete. I would also like to express my gratitude to Oxford Polytechnic, and in particular the School of Computing and Mathematical Sciences for providing financial support and facilities. I am grateful also to all those working in the mathematics of medical imaging who have given freely of their time and expertise in helping me, in particular David Isaacson, Bob Kohn and Frank Natterer. To my mathematics teachers over the years who have given me inspiration, in particular Margaret Penrose, Jim Eels, Fran Burstall, Christopher Zeeman, Charles Pugh, Steve Smale and Mo Hirsch. To my colleagues in the EIT group at Oxford Polytechnic who have been, and continue to be, a pleasure to work with Kevin Paulson, Chris Mcleod, John Lidgley, Dale Murphy (honorary member) and 'Dynamo' Pete Furner.

My colleagues at Oxford Polytechnic have been a source of strength. My fellow postgraduates, Nic Wilson, Bob Hoyle, Caroline Skeels, Nigel Crook, Dave Shrimpton, Martin Chapman and Avril Smith have been a source of good-humour, help and advice. Thanks are due also to my teaching colleagues, and students, who have been so sympathetic to the conflicting demands on my time between thesis writing and teaching.

A special thanks is due to my dear friends who have celebrated with me my joy in success as well as helping me through the depths of despair. Amongst these Sarah, Jonathan, and Karen deserve a special mention for their contribution to helping me keep my head together, as do the Boat Dwellers of Weirs Mill Stream who are such good neighbours.

Contents

1	Introduction	1
1.1	Electrical Impedance Tomography	1
1.2	Impedance Measurement In Medicine	1
1.3	Medical Imaging	2
1.4	Description of EIT	3
1.5	Brief History	5
1.6	EIT at Oxford Polytechnic	9
1.7	Structure of this Thesis	10
2	Mathematical Preliminaries	12
2.1	Introduction	12
2.2	Regions	13
2.3	Current and Voltage	13
2.4	Simple Examples	14
2.4.1	Uniform disk	14
2.4.2	Concentric Anomaly	15
2.5	Partial Differential Equations	16
2.5.1	Sobolev Spaces	16
2.5.2	Boundary Conditions	19
2.5.3	Weak formulation	21
2.5.4	Existence , Uniqueness and Continuous Dependence . .	22
2.5.5	Transfer Impedance	23
2.6	Compact Linear Operators	24

2.7	Calculus in Banach spaces	27
2.7.1	Derivatives and Order of Convergence	27
2.7.2	Taylor's Theorem	28
3	Linearisation	30
3.1	Introduction	30
3.2	Approach	30
3.3	Choice of Space for γ	31
3.4	Direct Form	33
3.5	Interpretation as a source	37
3.6	Integral Form	37
3.7	Operator Form	39
3.8	Translation to Inverse scattering problem	39
3.9	Historical Note	42
4	Finite Element Modelling	43
4.1	Choice of Forward Modelling Technique	43
4.2	Theory of the Finite Element Method	44
4.2.1	Preliminaries	44
4.2.2	Approximation Space	45
4.2.3	The System Matrix	48
4.2.4	Application of Boundary Conditions	48
4.3	Specification of Conductivities	49
4.4	The NAG Finite Element Library	50
4.5	Implementation of the Forward Modelling Program	51
4.5.1	History of fwprob	51
4.5.2	Choice of Language	51
4.5.3	Program Details	51
4.5.4	Trouble Shooting	52
4.6	Mesh Generation, Numbering and Refinement	53
4.7	Numerical Calculation of the Derivative Matrix	55

5	Ill-posedness of the Inverse Problem	57
5.1	Ill-posedness	57
5.2	Existence of an Inverse	58
5.3	Continuity of the Inverse	59
5.4	The Linearised Problem	61
5.5	Singular Value Decomposition	62
5.6	Calderón Fields	65
5.7	Fourier Series for Calderón Fields	67
5.8	Ill-posedness of Linearised Problem	68
5.9	Numerical Calculation of SVD	69
5.9.1	Implementation	69
5.9.2	Results and Interpretation	70
5.10	Polar or Adjacent?	75
6	Reconstruction Algorithms	79
6.1	Introduction	79
6.2	Non-linear Data Fitting	80
6.2.1	How many minima?	80
6.2.2	Newton's Method for a critical point	81
6.2.3	Calculation of the second derivative	83
6.3	Regularised Newton Methods	84
6.3.1	The Levenberg-Marquardt Method	84
6.3.2	Using the SVD.	86
6.3.3	Iterative Methods	87
6.4	X-Ray CT	89
6.5	Radon Transform Inversion	92
6.5.1	Not a Radon Transform	92
6.5.2	Consistent Updates	94
6.6	Application of ART to EIT	95
6.7	Implementation of Regularised Newton Methods	97

6.7.1	Reconstructing the Moat Object	97
6.7.2	The Effect of Data Errors	97
6.7.3	The Positivity Constraint	98
7	Adaptive Methods	104
7.1	Which Measurements to Make?	104
7.2	Two-norm Optimal Currents	105
7.3	Algorithms for eigenfunctions	106
7.3.1	Power Method	106
7.3.2	Higher eigenfunctions	107
7.3.3	Critique of the Measurement Model	108
7.4	Reconstructing With Optimal Currents	109
7.5	Numerical Results	110
7.6	Point Optimal Currents	110
8	Conclusions and Further Work	113
8.1	Summary of this Work	113
8.2	The Future of EIT	113
8.3	Some Leads on the Problems	114
8.3.1	Electrode Modelling	114
8.3.2	Electrode Placement and Boundary Shape Errors . . .	116
8.4	Final Remarks	116
A	Design of an Adaptive Current Tomograph	117
A.1	Introduction	117
A.2	System Overview	118
A.3	Circuit Details	120
A.3.1	Motherboard	120
A.3.2	Electrode-Interface Board	125
A.4	Software Drivers	128
A.5	Problems with the Design	130

A.6 Calibration and Testing	131
A.7 Future Work	132

Chapter 1

Introduction

*The Way that can be put into words
Is not the Eternal Way;
The name that can be named
Is not the eternal name.*

Lao Tzu

1.1 Electrical Impedance Tomography

Electrical Impedance Tomography (EIT) is a technique for creating an image of the inside of a body by making electrical measurements at its surface. The image reflects the variation in conductivity or impedance within the body. The main application of EIT considered in this thesis is as a medical imaging technique for use in research and diagnosis. There are other applications and proposed applications of EIT to which much of this work is equally applicable. These include geological studies, non-destructive testing, archaeology, and industrial process monitoring. Certain species of weakly electric fish also use electrical location to hunt their prey and avoid obstacles in dark or murky water.

1.2 Impedance Measurement In Medicine

Impedance measurements have been used in several areas of medical diagnosis without actually attempting to image the impedance distribution. These

fall under the general heading of *impedance plethysmography* [89]¹. Measurements of the impedance of the thorax have been used to assess intrathoracic fluid content with a view to diagnosis of cardio-pulmonary diseases such as pulmonary oedema, pleural effusion and pneumothorax. These studies have been predominantly qualitative, rather than attempting to quantify intrathoracic volume changes. Thoracic impedance has also been used to monitor respiration - a technique called *Impedance pneumography*. A pair of strip electrodes, one on the neck and another on the upper abdomen, have been used to measure blood volume changes in the major arteries during the cardiac cycle. A crude model has been used to give an approximate measure of cardiac output. Electrodes attached to the head have been used to study cerebral haemodynamics, in particular, Tarassenko [89] applies this technique to the newborn.

1.3 Medical Imaging

An image of a cross section of the body is clearly a useful aid to medical diagnosis, treatment and research. There are several methods of *Tomographic Imaging*² which achieve this. In all these methods, energy is applied to the body in the form of a field or wave. This interacts with the tissues and the effect is measured. The first, and most widely known example, is X-Ray Computerized Axial Tomography (CAT or CT scanning). In this case, collimated X-Rays are directed at the body from a large number of different directions within a fixed plane. The attenuation of the X-Rays by the body is measured. An image of the attenuation function is then calculated by a computer (*reconstructed*) and the result displayed on a screen. A detailed Mathematical treatment of the reconstruction problem for X-Ray CT can be found in [44,72]. The other main methods of tomographic imaging used are Ultrasound, Nuclear Magnetic Resonance and Emission Computed Tomography. In Ultrasound an ultra high frequency sound wave is applied to the body. This may either be used as an 'echo sounder', similar to those used on ships, in which the delay between the emitted and reflected pulse is used to give a measure of depth, or as a true tomographic imaging technique similar to X-Ray CT. In Nuclear Magnetic Resonance (NMR) Imaging the patient is subjected to a strong static magnetic field. The rotation axes of the protons (Hydrogen atom nuclei) within the tissues take up a preferred orientation. If an alternating field is applied the protons will precess about the direction of the field. Some of the protons will absorb energy and be re-oriented. The energy absorbed by the proton is re-emitted when the field is removed. This signal depends

¹References may be found at the end of the thesis, starting on page 136.

²From the Greek *τομος* - slice. Hence tomography is imaging slices.

on the chemical environment of the proton. Spatial information is derived by changing the static magnetic field. An image of the tissues can then be reconstructed from these signals. In Emission Computed Tomography radioactive substances are introduced into the body. The radiation emitted from these substances passes through tissues and is measured using an array of suitable detectors. The two main methods of Emission Computed Tomography are Single Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET). In SPECT the decay of the radioactive substance results in the emission of a single photon of gamma radiation which is detected by a 'gamma camera' consisting of an array of collimators and detectors mounted on a moveable gantry. Different projection angles can be obtained by movement of the gantry. In PET the annihilation of an emitted positron results in a pair of photons with opposite momentum. These photons are detected by an annular array of detectors. In Emission Computed Tomography it is the concentration of the radio isotopes which is imaged, rather than the attenuation of the radiation by the tissues.

There are many variations on these methods, different scanning geometries and reconstruction algorithms. Each method has both advantages and disadvantages. X-Ray CT can achieve very high resolutions and produce accurate three-dimensional images, but this involves subjecting the patient to X-Rays. NMR scanners deliver high resolution and probably involve less risk to the patient than X-Ray CT but they require magnets capable of creating uniform fields of very high strength. This kind of equipment is expensive and physically large. Due to long relaxation times of protons NMR scanning is inherently slow. This makes real-time studies difficult. Emission Computed Tomography also involves large and expensive scanners. The resolution of Emission Computed Tomography systems is relatively low (of the order of 5mm) and scan times long (of the order of 15 minutes). However, the advantage is that the images are related to physiological processes rather than tissue structure. Ultrasound is relatively inexpensive and small, it is also safe and able to distinguish between soft tissues.

1.4 Description of EIT

In EIT an electric current is applied to the body via electrodes in contact with the skin. The current passes through the body and the resulting potential differences on the skin are measured. An alternating current is used, as a direct current would cause polarization of ions within the tissue. For sufficiently low current densities (below 1 mA/cm^2 [82]) tissues are Ohmic conductors. The reactive component of tissue impedance is typically 100 times less than the

resistive component at frequencies between 1kHz and 200kHz.

Body tissues show a wide range of variation in resistivity, as can be seen in Table 1.1. In particular soft tissues can easily be distinguished on the basis of impedance.

Tissue	Resistivity (Ωcm)	Comments
Cerebro-spinal fluid	65	
Blood	150	
Grey matter	250	
White matter	500-700	
Bone	1000-2000	newborn
	10000	adult
Muscle	200	transverse
	1000	longtitudinal
Fat	2000	

Table 1.1: *Some measurements of resistivity of human tissues taken from the literature. See [89, p122] and [11] for detailed references. Published sources differ by a factor of up to two.*

An Impedance Tomograph is essentially an elaborate Ohm meter. It applies currents and measures voltages. A number of electrodes are stuck to the skin of the patient and these are connected to the tomograph by leads. The tomograph has one or more current sources and one or more digital ‘volt meters’ (an analogue to digital converter — ADC). These are often connected to the electrode leads via a multiplexer, which is an electronic switch. This allows the same current source or ADC to be applied to any electrode. The tomograph is driven by a computer, either a dedicated microprocessor or external ‘host’ computer, which instructs the machine which patterns of current to apply and which voltages to measure. The measurements made are passed to the host computer which calculates an impedance image consistent with the measurements. The data collection phase can be made very rapid. It is possible to collect data at the rate of 25 frames-per-second which is the rate required to produce the illusion of continuous motion to a human observer. Higher rates may be achieved using parallel hardware. If reconstruction algorithms and hardware can match this speed EIT will be capable of following, for example, the cardiac cycle, in real-time. This fast data collection rate is a potential advantage for EIT over the methods discussed in the previous section. The problem of fast image reconstruction is addressed in this thesis and in [17].

An EIT system also has the potential advantages of low cost and small

physical size. The electronics required can be built in a box the size of a desktop computer for a few thousand pounds. A reasonably powerful host computer is required. The speed and memory capacity of the computer required is dictated by the size of data set collected, the accuracy required and the desired reconstruction rate. However, the cost of computer processing power is currently decreasing exponentially with time, so cost of the host computer need not be a limitation in the long term. It must be emphasised that the speed and size of NMR, CT and ultrasound scanners continues to improve and it is not clear in the long term whether EIT will show a significant advantage in this respect. One difference that will always remain between the various techniques is that they image different physical properties of tissue. As there is no particular relationship between the speed of sound and the electrical conductivity of a medium, Ultrasound and EIT will produce different images in general, and hence different clinical information. In this sense different imaging techniques may be regarded as complimentary rather than competing.

The main limitation on an EIT system is the resolution. This is limited by the size of the voltage data set collected which is determined by the number of electrodes. Typical systems under development still use only 32 electrodes. If EIT is to compete with other imaging modalities on the basis of resolution, vastly more electrodes will have to be used. The resolution and accuracy of the images is also constrained by the ill-posedness of the problem. This means that large improvements in the measurement accuracy are required to produce small improvements in the accuracy of the image. This, coupled with the inherent difficulty in making electrical measurements on anything as awkward and variable as the human body, means that EIT is difficult. Even if EIT cannot be developed to provide accurate high resolution images, it nevertheless has a useful role to play. All the techniques which use electrical impedance for physiological measurement could be improved by a more accurate knowledge of the cause of the impedance changes measured. Impedance tomography provides this information.

1.5 Brief History

The story starts several millennia ago with the evolution of fish which use electrical location. It is debatable whether this actually constitutes imaging. Presumably they interpret the impedance changes directly — however I doubt if this worried them unduly. The geologically motivated studies of Langer [57, 58] and Slichter [84] in 1933-36 provide possibly the earliest indications that *Homo sapiens* were endeavoured to probe their environment using electrical current. These treated the essentially one dimensional problem of determining

the depth and conductivities of horizontal layers in the earth.

The first papers on the use of electric current for medical imaging appear in the late 1970s. Henderson and Webster [43] put forward the idea of an ‘Impedance Camera’. This stands in the same relationship to impedance tomography as radiography does to CT. The idea was to create an approximately uniform electric field — forcing the current paths to be as straight as possible using ‘guard electrodes’. This was seen as analogous to the straight paths of X-Rays. As in X-Ray radiography, a transmission image was created. The method was tested using a system with an array of 100 active electrodes and 40 guard electrodes which was used to produce a transmission image of the human thorax. Price [75,76] in 1979 advanced the idea of ‘Electrical impedance tomography (ICT): A new CT imaging technique’, (since then there has always been disagreement on what to call this method). Price realised that a number of different ‘projections’ were needed for true tomography and also noted the insensitivity of exterior voltage measurements to conductivity changes in the interior. He advocated the use of guard electrodes to make the ‘beam of current’ have parallel stream lines — this would enable the use of a CT-like reconstruction algorithm.

Schomberg [81] also assumed that reconstruction could be achieved using X-Ray CT techniques. The analogy with X-Ray CT at this stage was both a useful insight and a distraction. As is made clear later in this thesis the forward projection in EIT (that is the determination of voltages resulting from a given current pattern applied at the surface of an object with given inhomogeneous conductivity) differs from its analogue in X-Ray CT in two important respects. It is non-linear, and even when linearised it is not a generalised Radon transform (in the sense of Gel’fand *et al* [31]). The fact that the forward mapping is not a Radon transform is manifested in the argument by Bates, McKinnon and Seagar [4]. They point out that a rotated cosine pattern of current, which produces a uniform field in a homogeneous two-dimensional disk, will produce the same boundary voltage profiles for a circularly symmetric conductivity distribution as a homogeneous one of a suitable conductivity. This means that an approach based on X-Ray CT will not succeed. Continuing this New Zealand tradition Seagar produced an excellent thesis [83] focusing mainly on circularly symmetric distributions on the disk. He showed that these could be reconstructed using data derived from a Fourier basis of currents, thus including the information found to be missing in [4]. He used Newton’s method to perform the reconstruction, recognising the non-linear character of the problem.

It is perhaps surprising then that in the early 1980s the first group to produce an *in vivo* tomographic impedance image used a reconstruction algorithm based on a back-projection argument derived by analogy with CT [11]. This

group, based at Sheffield and headed by Barber and Brown, used a sixteen electrode system with a single constant current source and single ADC. The current pattern used was to drive across adjacent pairs of electrodes, voltages being measured between other adjacent pairs. In contrast to the ideas put forward by Tasto and Schomburg [90], who advocated back-projection along current stream-lines, Barber and Brown back-project between equipotential lines. Despite the lack of theoretical justification this method produced recognisable images which the group have continued to refine and develop using the same basic technique. Some justification for this technique, formulated in a tighter mathematical framework, was given by Santosa and Vogelius [80] but this did not appear until 1988.

Apart from the question of algorithms there were two main controversies in the impedance imaging community in the early to mid 1980s. The first was whether to apply current and measure voltage or to apply a fixed voltage and measure current. Kim *et al* [51] built a machine which could apply multiple voltages and measure currents. Yorkey [96] also built such a system but with only one constant voltage source. The use of these systems, as far as the author is aware, was restricted to studies on phantoms and not applied to patients. Brown in [20] give the rationale for the choice made by the Sheffield group to apply current and measure voltage. The main reason is the presence of contact impedance on the electrodes. Brown advocates making what he calls *four electrode* measurements, rather than two electrode measurements thus eliminating the need to make voltage measurements on current carrying electrodes.

The second area of controversy was in the area of which current patterns (or indeed voltage patterns) to use. The Sheffield group firmly advocated their two current drives in adjacent positions while others, such as Tarassenko [89] advocated driving opposite pairs. Others still, such as Kim [53], advocated the use of guard electrodes to straighten current patterns, an idea which by the mid 1980s was distinctly unfashionable. Parallel to these developments by medical physicists and engineers, mathematicians were working on the problem, largely unaware of the medical applications. The main question addressed was uniqueness of solution, that is whether a conductivity distribution can be uniquely determined by boundary measurement. The earliest attempt to address this question was by one of the 'grandfathers' of partial differential equation theory, A.P. Calderón. He delivered a paper [22] in 1980 which is something of a legend in the literature of impedance tomography — it is very difficult to obtain as the proceedings of that Brazilian conference are nowhere to be found. It exists therefore mainly in the form of second or third generation photocopies of what appears to be an early draft. He considered the uniqueness problem and derived a formula for the derivative of a particular forward

mapping and proved that this derivative was injective. Calderón's motivation for studying the problem appeared to be electrical prospection. In the circles of the American Mathematical Society the uniqueness problem for EIT was colloquially referred to as 'the Calderón problem'. Uniqueness theorems were proved by Kohn and Vogelius [55] in 1985 and Sylvester and Uhlmann [87] in 1987. Later refinements of these results were published by Ramm [77] in 1988. It was probably 1986 before the American mathematicians working in this area became aware of the medical applications.

Development of reconstruction algorithms up to about 1986 proceeded largely in isolation from the mathematical community. They continued to be inspired by the many different techniques used X-Ray CT. The idea of a *sensitivity matrix* occurred to many investigators. This is a matrix of coefficients which relate the change in measured voltage (or current) to a localised change in conductivity. Opinions differed as to how to calculate this matrix and what to do with it. Some calculated the matrix by a perturbation technique [53,89], others used a more sophisticated approach, either based on a discrete resistor network [96] or on a continuous model for fields such as that used by Yamashita and Takahashi [94], Nakayama *et al* [95] or Murai and Kagawa [67]. These methods will be derived in a uniform context in Chapter 3. Both Kim and Tarassenko used the sensitivity matrix in apparently *ad hoc* variations on the iterative methods used in X-Ray CT. These both owe something to what, in the tomography literature, is referred to as Algebraic Reconstruction Technique (ART)(see [72] for details) and in the numerical analysis community as the method of Kaczmarz [49,72]. New variations are still published on this theme, for example the recent work of Kotre [56]. The sensitivity matrix, however formed, is ill-conditioned with respect to inversion, as will be shown in Chapter 5. Iterative algorithms, if halted short of convergence, have a regularising effect [88], which explains the success of the methods used by [53,89,56]. If a direct solution technique is employed, some explicit regularisation must be used, such as Tikhonov regularisation or truncation of singular function expansion. Both these methods have been applied. In [94] a single linear system was solved using Tikhonov regularisation. A segmented body model was used but no iteration was attempted. Murai and Kagawa [67] used the singular value decomposition. Yorkey [96] compared these methods and found the direct approach using Tikhonov regularisation best.

The first European workshop on EIT was held in Sheffield in 1986 [19]. This was the first large gathering of EIT enthusiasts from around the world. A particularly challenging paper was delivered by Isaacson [37] who advocated a completely different approach to the current patterns to be used and extended the idea of Seagar [83] of using a Fourier basis of currents. Isaacson advocated using current patterns which are eigenfunctions of a certain operator. This

had more in common with the discredited idea of guard electrodes than the fashionable four electrode measurement schemes and the idea was met with a certain amount of scepticism. Nevertheless, the group, based at Rensselaer Polytechnic Institute (RPI), New York, proceeded to build a system capable of applying this type of current pattern [74] which has now produced *in vivo* images of the human thorax.

1.6 EIT at Oxford Polytechnic

Nothing has yet been said of our own role in the development of the subject. The EIT project at Oxford Polytechnic started with Tarassenko joining the institution in 1985. He had completed a DPhil Thesis [89] under Rolfe at Oxford University and the John Radcliffe Hospital. The results of his thesis indicated the feasibility of using EIT to detect cerebral haemorrhage in the newborn. This work was continued at the John Radcliffe Hospital by Murphy [68]. Tarassenko realised that EIT reconstruction was an inverse problem which could be of interest to mathematicians. This challenge was taken up by Pidcock. The present author joined the group in September of 1985.

It was clear to the author from the outset that finite element modelling would be important, and that the problem was non-linear and ill-posed. Some kind of regularisation would be needed together with an iterative technique [12, 13].

The authors initial studies were into the ill-posedness of the linearised inverse problem. The results of these were presented in [14] and in more detail in Chapter 5. This work led to reconstruction algorithms which worked successfully on computer simulated data [15]. The adaptive current techniques of Isaacson [47] were an exciting development and led to an investigation of how such techniques could be used in conjunction with iterative reconstruction algorithms. These gave very promising initial results which were presented at the Second EC workshop on EIT [16]. At that meeting Murphy and Breckon started a collaborative venture to pursue these promising numerical results and build an Adaptive Current Tomograph. There was at the time only one such machine in existence, that at RPI. The machine, called OXPACT1 [69,70], is still under development and much has been learnt from its design.

In 1988 the team was joined by Paulson who has contributed significantly in two essential areas. He implemented both the finite element forward modelling code and several linear solvers for the inverse problem on a multi-processor system using Transputers, both of which constituted ground-breaking work in the use of Transputers, let alone EIT [17]. This points the way to the possibility of using non-linear full matrix techniques for real-time reconstructions. Paulson

has also investigated more accurate modelling of the boundary conditions at the electrode-skin interface, extending the work of Cheng [24]. An accurate forward model incorporating these more sophisticated boundary conditions will no doubt become essential for improved accuracy in reconstructions.

Our hope is now to design and build a second EI Tomograph, learning from the experience gained with OXPACT1. Our aim is to make a system of taking measurements on several planes of electrodes in order that adaptive techniques can be tried in three-dimensions, further it is hoped that this machine will be able to be used for clinical trials.

1.7 Structure of this Thesis

In Chapter 2 a review is given of the mathematical techniques and notations used later in the thesis. This chapter contains no original mathematics. However the discussion of the choice of space for the conductivity is new. In particular the problem of choosing a Hilbert space in which the set of feasible conductivities is an open set has not been addressed before, neither has the problem of evaluating the conductivity on the boundary. Chapter 3 extends the work of Calderón on calculation of the derivative of the forward mapping. This leads to a rigorous justification of the linearisations employed by other authors. The linearised problem is presented in both a direct and integral form, and the connection with the inverse scattering problem briefly explored. Chapter 4 gives details of the finite element method used to solve the forward problem and calculate the derivative matrix numerically. Two factors which make EIT reconstruction particularly difficult are the non-linearity and ill-posedness of the problem. While the non-linearity was exhibited in the Taylor series of Chapter 3, the ill-posedness is explored in Chapter 5. Two original analytical proofs of ill-posedness are given. These have not previously been published. A numerical study of the singular value decomposition of the derivative is presented. Details of singular values have already been published by the author, but contour plots of the singular functions have not previously been exhibited.

In Chapter 6 reconstruction algorithms are discussed. The reconstruction problem is formulated as a minimisation problem and an example of a two-dimensional contour plot of the objective function is given. A formula for the second derivative of the forward mapping is derived. Regularised Newton methods are discussed in detail and this leads to an analysis of reconstruction algorithms employed by other authors. It is seen that they all approximate to some form of regularised Newton method. In particular methods based on an analogy with X-Ray CT are explored. The author's own implementation of a

regularised Newton method is then presented. An investigation of the effect of data errors is reported and the problems of applying the positivity constraint discussed. Chapter 7 explores the use of adaptive current techniques. An algorithm is given for the calculation of a set of optimal currents for measurements of a given precision. A reconstruction algorithm is presented which integrated the adaptive measurement technique with Newton iterations. Finally an alternative point-optimal measurement system is suggested. Chapter 8 concludes the thesis with a discussion of directions of further work.

Chapter 2

Mathematical Preliminaries

And he who is versed in the science of numbers can tell of the regions of weight and measure but cannot conduct you thither.

Kahlil Gibran

2.1 Introduction

The study of Electrical Impedance Tomography is by necessity interdisciplinary. It requires a collaborative effort between medical physicists, medical engineers, electronics engineers, computer scientists and mathematicians. International workshops in this field (such as those reported in [19] and [21]) have demonstrated how fruitful such collaborations can be, but also highlighted the difficulties in communicating across the cultural divide between those trained in different disciplines. While this thesis resides firmly within the discipline of applied mathematics it is hoped that it will be of some use to workers in the field of EIT whose background is less mathematical.

This chapter aims to summarise notations, methods and models used later in the thesis and to indicate where further details may be found. The author would dearly love to write a tutorial on the mathematical prerequisites for EIT but that would be inappropriate to include in a doctoral thesis as well as impossible to achieve within the limits of time available. It is hoped that such information as is given will go some way to widening the possible audience for this work.

2.2 Regions

The physical domain under consideration is a region of space consisting of a material of variable electrical conductivity. The word ‘space’ here means two- or three- dimensional Euclidean space: R^2 or R^3 . Conditions will be placed on the region, which will be called Ω , to ensure that standard results about solutions of partial differential equations may be used (see [36] and [29]). The region will be a bounded, open, simply connected set. (*Simply connected* — connected with no ‘holes’). The boundary of the region, denoted by $\partial\Omega$, will be assumed to be smooth (although this condition can be weakened to include corners [36]). In the medical context Ω will be a human body and $\partial\Omega$ the surface of the skin.

2.3 Current and Voltage

Before giving more detailed mathematical definitions let us consider the model for electrical conduction in a body which is given to us by Maxwell’s equations. Suppose that the electrical potential (voltage) at a point x in Ω is $u(x)$. It is assumed that the body is an Ohmic, isotropic conductor. The conductivity at a point x is $\gamma(x)$. The tissues of a human body vary in conductivity (see for example [20] for a table of experimental results). Some tissues, such as muscle, are actually anisotropic [20]. Nevertheless the simplifying assumption of isotropy will be made throughout this thesis. The effects of anisotropy on EIT have not yet been investigated, although Kohn and Vogelius [54] exhibit two anisotropic distributions which cannot be told apart using EIT. From Ohm’s law the current density is a vector field J given by

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

If there are no electrical sources within Ω then by Gauss’s Law

$$\nabla \cdot J = 0$$

and hence

$$\nabla \cdot \gamma \nabla u = 0 \tag{2.1}$$

in Ω (see for example [60]). In this case Gauss’s Law is the equivalent of Kirchhoff’s Current Law for a continuum. Equation 2.1 is fundamental to the study of EIT. It is therefore convenient to introduce a notation for the differential operator involved:

$$L_\gamma u \equiv \nabla \cdot \gamma \nabla u. \tag{2.2}$$

If there are electrical sources within Ω then Equation 2.1 becomes

$$-L_\gamma u = q \quad (2.3)$$

where q is the current source density. For known γ (and q) Equation 2.1 (and Equation 2.3) are second order elliptic partial differential equations for u . For known u (and q) they are first order hyperbolic partial differential equations for γ .

In EIT an electric current is applied to $\partial\Omega$. If \mathbf{n} is the outward unit normal to $\partial\Omega$ and j is the flux density of the current applied to $\partial\Omega$ then

$$j = J \cdot \mathbf{n} \quad (2.4)$$

on $\partial\Omega$. The operator $\nabla_{\mathbf{n}}$ takes the normal derivative of a scalar field. As the normal only exists on $\partial\Omega$ the function $\nabla_{\mathbf{n}}u$ is only defined on $\partial\Omega$ although u is defined on Ω . Thus Equation 2.4 becomes:

$$j = -\gamma \nabla_{\mathbf{n}}u. \quad (2.5)$$

To find the actual current I passing into an area $A \subset \partial\Omega$ one takes the integral of the current density:

$$I = \int_A j.$$

Note that where no explicit references to the space variable x is made in an integral the surface or volume measure dV or dS will be omitted.

As EIT is a non-invasive technique the voltage resulting from the application of this current can only be measured on the surface of the skin. The restriction of u to $\partial\Omega$ will be denoted by $u|_{\partial\Omega}$.

2.4 Simple Examples

2.4.1 Uniform disk

At this stage it will help to have a concrete example. Let D denote the unit disk in the plane $D = \{(x, y) | x^2 + y^2 \leq 1\}$. We will assume that it consists of a material which has conductivity $\gamma = 1$. Equation 2.1 then becomes

$$\Delta u = 0 \quad (2.6)$$

where $\Delta \equiv \nabla \cdot \nabla$ is the Laplacian operator¹. Equation 2.6 defines u as a harmonic function. The real and imaginary part of any complex analytic

¹Using Δ , instead of ∇^2 is a kind of mathematicians shibboleth but it also has the attraction that its inverse (an *integral* operator) can be written as Δ^{-1} . It also carries connotations of the Laplace-Beltrami operator which is defined independently of coordinate systems and thus makes sense on curved manifolds.

function $f(z)$ satisfies this equation where the complex variable $z = x + iy$. An example of this is $f(z) = z^n$ for any positive n and $u(x, y) = \Re(x + iy)^n$. Given that the real and imaginary parts of f both represent valid potentials, let us now assume that complex potentials exist, on the understanding that real and imaginary parts actually represent separate potentials. We may then write $u_n = (x + iy)^n$ or in polar coordinates $u_n(r, \theta) = r^n e^{in\theta}$. Using the principle of superposition (that is the linearity of Δ) any linear combination of these solutions is also a solution. For sufficiently rapidly decaying coefficients c_k

$$u(r, \theta) = \sum_{k=0}^{\infty} c_k r^k e^{ik\theta} \quad (2.7)$$

will also be a solution (the rate of decay of coefficients gets special attention in the next section). The voltage on the boundary is easily calculated by setting $r = 1$:

$$v(\theta) = \sum_{k=0}^{\infty} c_k e^{ik\theta}. \quad (2.8)$$

The current density j on the boundary can be calculated by differentiating the right hand side of Equation 2.7 with respect to r and setting $r = 1$:

$$j(\theta) = \sum_{k=1}^{\infty} k c_k e^{ik\theta}. \quad (2.9)$$

If the Fourier coefficients c_k of v are known then u can be found at any point in the disk from Equation 2.7. Equation 2.9 gives j whose Fourier coefficients are $k c_k$. Conversely if the Fourier series for j is known the c_k can be found by simply dividing by k .

2.4.2 Concentric Anomaly

Perhaps the simplest example of a non-homogeneous conductivity is the concentric anomaly, known colloquially as the ‘blob in the middle’. This is the radially symmetric, piece-wise constant conductivity distribution

$$\gamma(r, \theta) = \begin{cases} \sigma, & \text{for } r < \rho \\ 1, & \text{for } \rho \leq r \leq 1 \end{cases} \quad (2.10)$$

The solutions of Equation 2.1 are given in [47]. If the current density on the boundary is given by $j(\theta) = \sum_{k=0}^{\infty} c_k e^{ik\theta}$ then for $r > \rho$

$$u(r, \theta) = - \sum_{k=1}^{\infty} \frac{c_k}{k} \frac{1 - \mu \rho^{2k} r^{-2k}}{1 + \mu \rho^{2k}} r^k e^{ik\theta} \quad (2.11)$$

where

$$\mu = \frac{\sigma - 1}{\sigma + 1}. \quad (2.12)$$

2.5 Partial Differential Equations

Equation 2.1 is a second order partial differential equation (pde) for u . Provided γ is bounded above and below, which means physically that there are no perfect conductors nor perfect insulators in the region, the equation is bounded and elliptic. The exact assumptions made about γ will be discussed in the next chapter. A space of functions must be chosen in which to represent the electrical potential u . The most general assumption is that u lies in the space of tempered distributions or ‘generalised functions’ as described in [29] or [78]. Derivatives of all orders are defined for distributions so Equation 2.1 certainly makes sense, however additional assumptions must be added so that the pdes involved have unique solutions. For this the notion of a Sobolev space must be introduced.

2.5.1 Sobolev Spaces

The idea of a Sobolev space is intuitively simple. A (generalised) function f lies in the Sobolev space H^k if $\partial^\alpha f / \partial x^\alpha$ is in L^2 for all multi-indices² α with $|\alpha| \leq k$, that is

$$\int |\partial^\alpha f / \partial x^\alpha|^2 < \infty, \quad \text{for all } |\alpha| \leq k.$$

Thus the condition that a function lies in H^k is a smoothness condition. When it is necessary to make the domain Ω of f explicit one writes $H^k(\Omega)$. It is possible to define Sobolev spaces based on L^p for $p \neq 2$. These are not Hilbert spaces and as an inner product will be needed in this thesis for the definition of singular functions the case $p \neq 2$ will not be used. The Sobolev space H^k is a separable Hilbert space with norm $\|\cdot\|_k$ given by

$$\|f\|_k^2 = \sum_{|\alpha| \leq k} \|\partial^\alpha f\|_{L^2}^2, \quad (2.13)$$

where $(\partial^\alpha f)(x) = \partial^\alpha f / \partial x^\alpha$. The definition just given applies for k a non-negative integer, however it is useful to generalise this notion further to include any real number. For functions defined on all of R^n this can be done elegantly using Fourier transforms. The relationship between the Fourier Transform and derivative of a function is

$$(\widehat{\partial^\alpha f})(\xi) = (2\pi i \xi)^\alpha \hat{f}(\xi) \quad (2.14)$$

²A multi-index α is an ordered set of non-negative integers $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_r)$. This provides a convenient notation for expression involving partial derivatives. In this notation $|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_r$ and $\frac{\partial^\alpha f}{\partial x^\alpha} = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_r^{\alpha_r}}$

where \hat{f} is the Fourier Transform of f and α is any multi-index. A function is said to be in $H^s(R^n)$ when $(2\pi i\xi)^\alpha \hat{f}(\xi) \in L^2(R^n)$ for any α with $|\alpha| \leq s$. This definition applies for all real s . The norm $\|\cdot\|_s$ on $H^s(R^n)$ is given by

$$\|f\|_s = \|(1 + |\xi|^2)^{s/2} \hat{f}(\xi)\|_{L^2}. \quad (2.15)$$

When s is an integer this formula gives a norm equivalent to the one given above in Equation 2.13. The definition of $H^s(\Omega)$ for a bounded domain Ω is more technical. The approach taken by Lions and Magenes in [63] is to define the Sobolev spaces for integer s by Equation 2.13 and then define the intermediate spaces by interpolation. Other approaches are possible and each gives a different insight into the nature of these spaces ³.

The negative order Sobolev spaces have an important interpretation as *dual spaces*. If X is a Banach space then X^* is the space of scalar valued continuous linear operators on X , that is $X^* = L(X, R)$. As an analogy, it is helpful to think of the finite dimensional version: if X is thought of as a space of column vectors then X^* is the space of row vectors. Usually a Banach space and its dual are completely different spaces although in a Hilbert space the Riesz representation theorem [1] gives a correspondance – for any $\alpha \in X^*$ there is an $a \in X$ with $\alpha(x) = \langle a, x \rangle$. For this reason it is customary to think of the elements in X^* as functions and by a standard abuse of notation use the same symbol $\langle \cdot, \cdot \rangle$ for the *dual pairing* as the inner product and write $\langle \alpha, x \rangle = \alpha(x)$. The negative order Sobolev spaces can either be defined by $H^{-s} = (H^s)^*$, or in the case of the Fourier transform definition, such non-convergent Fourier series only have meaning when multiplied by a smoother function and integrated, so the duality is included in the definition.

One more word about duals, if $A : X \rightarrow Y$ is a linear map between Hilbert spaces, there is a notion corresponding to the transpose for matrices, the *adjoint operator* A^* defined by

$$\langle A^*y, x \rangle_X = \langle y, Ax \rangle_Y \quad (2.16)$$

for all $x \in X$ and $y \in Y$.

It may be helpful to consider the Sobolev norms in terms used by electronics engineers. In these terms it becomes a very natural and familiar concept. The distribution f is thought of as a signal in the spatial rather than temporal domain. The H^s norm $\|f\|_s$ is then *the root-mean-square (r.m.s.) of the*

³The author's favorite is $\|f\|_s^2 = \int_\Omega (1 + \Delta)^s f$ which works for any manifold on which the Laplace-Beltrami operator is defined (any oriented Riemann manifold). The beauty of this definition is that it has no reference to local coordinates and yet it incorporates the geometry of the space. To calculate $(1 + \Delta)^s$ one diagonalises using the eigenfunctions of Δ which on the circle are the Fourier basis and on the sphere are the spherical harmonics.

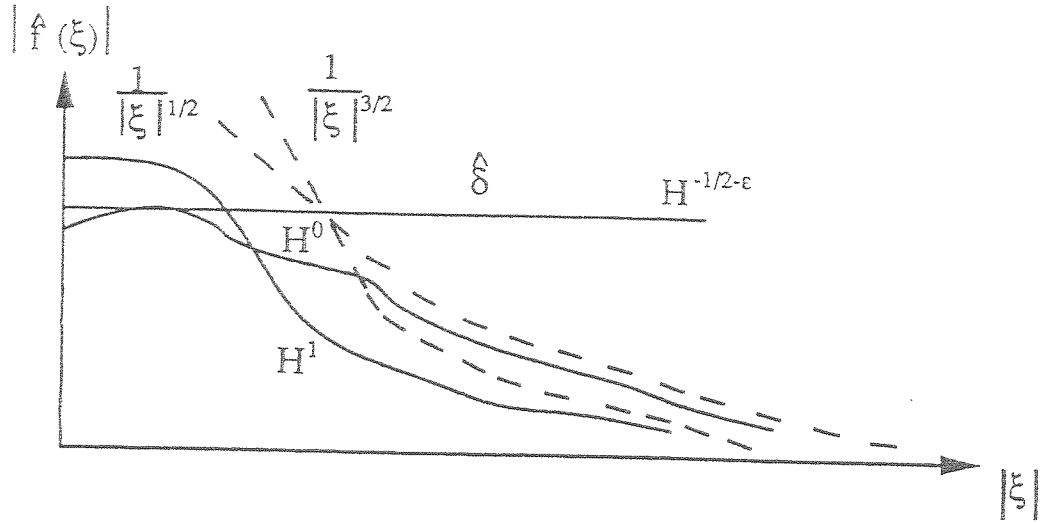


Figure 2.1: The frequency spectrum of functions in the Sobolev spaces $H^s(\mathbb{R})$ for various s . Note that it is only the asymptotic behaviour of the frequency spectrum which matters.

filtered version of f . If s is positive a *high pass* filter is used amplifying the higher frequencies according to an s -th power law. If s is negative a *low pass* filter is used, attenuating the higher frequencies according to an inverse $|s|$ -th power law. The space H^s consists then of all the signals with finite r.m.s. size when filtered in this way (see Figure 2.1).

Complete discussion of Sobolev spaces may be found amongst the references [29],[78],[36],[93] and [63] but their more important properties will be discussed here.

If $s > t$ then for a compact domain $H^s \subset H^t$. The inclusion mapping (sometimes called an embedding) has a property explained later in Section 2.6: it is a compact mapping. A particular case of this embedding tells us that

$$H^s \subset H^0 = L^2, \quad \text{for all } s \geq 0 \quad (2.17)$$

which means that distributions in a Sobolev space of positive order are also functions in the more conventional sense. If a distribution lies in a Sobolev space of negative order however, it only has physical significance when multiplied by a conventional function ('test function') and integrated over its domain. The negative order Sobolev spaces are the home of exotic beasts like the so called 'Dirac delta function' once thought to be a purely mythological animal. In one dimension, for example, the delta distribution δ_p centered at the point p on the real line \mathbb{R} just fails to make it into $H^{-1/2}(\mathbb{R})$. In fact it is easy to deduce from the fact that $\hat{\delta}_0$ is constant that $\delta_0 \in H^{-1/2-\epsilon}(\mathbb{R})$ for any $\epsilon > 0$ (see Figure 2.1).

The partial derivative operator ∂^α is a continuous map

$$\partial^\alpha : H^s \rightarrow H^{s-k} \quad (2.18)$$

provided $|\alpha| \leq k$. The result means intuitively that differentiation reduces the smoothness of a distribution.

The traditional notion of smoothness is that a function f is in C^k if $\partial^\alpha f$ exists (in the conventional sense) and is a continuous function whenever $|\alpha| \leq k$. Sobolev's lemma bridges the gap by guaranteeing that for an n -dimensional domain Ω

$$H^s(\Omega) \subset C^k(\Omega), \quad \text{whenever } s > k + n/2. \quad (2.19)$$

In particular taking $k = 0$ functions in $H^s(\Omega)$ are continuous provided $s > n/2$.

Continuous functions defined on Ω have natural restrictions to $\partial\Omega$. However this is not generally true even for functions in $L^2(\Omega)$ as these are not defined on a subset of Ω with measure zero. The Sobolev theory provides us with a condition for the existence of a natural restriction map. The restriction map (or trace operator) $\tau : C^0(\Omega) \rightarrow C^0(\partial\Omega)$ extends to a continuous map

$$\tau : H^s(\Omega) \rightarrow H^{s-\frac{1}{2}}(\partial\Omega) \quad (2.20)$$

provided $s > \frac{1}{2}$. This says intuitively that one half a degree of differentiability is 'lost' when one restricts to the boundary. A sketch illustrating intuitively how a function may be less smooth when restricted to the boundary of its domain is given in Figure 2.2.

2.5.2 Boundary Conditions

For a partial differential equation to have a unique solution, boundary conditions must be imposed. The two most commonly occurring boundary conditions are Dirichlet conditions and Neumann conditions. A Dirichlet condition is imposed by specifying the value of the function on the boundary whereas a Neumann condition is a specification of the normal derivative. In the case of the elliptic equation $L_\gamma u = s$ the natural Neumann condition is $-\gamma \nabla_n u = j$ for some distribution j . Physically then, a Dirichlet condition corresponds to specifying the voltage on $\partial\Omega$ whereas a Neumann condition corresponds to specifying a current density on $\partial\Omega$.

From physical considerations we see that if a certain pattern of voltages is established on $\partial\Omega$ then a current will flow through Ω and a potential u will be established. One might achieve this by connecting constant voltage sources to electrodes on $\partial\Omega$. The current drawn from, or sunk into, the voltage sources would then be determined by the resulting current density j . Conversely one

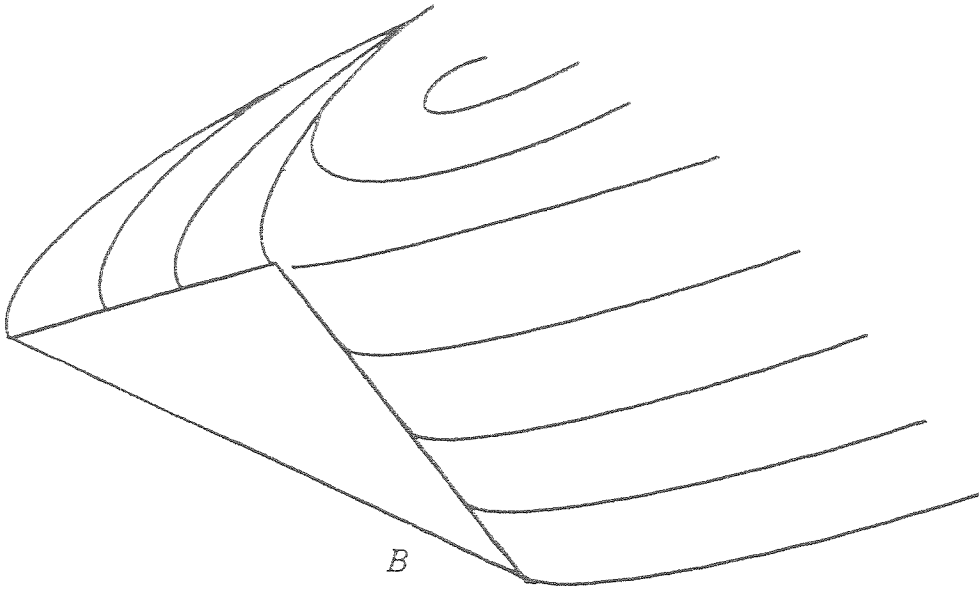


Figure 2.2: *Sketch of a function defined on the half-plane. It is differentiable at every point in the interior but fails to be differentiable on the boundary B .*

could attach constant current sources (and sinks) to establish a current density j over $\partial\Omega$. This would give rise to a potential u in Ω and v on $\partial\Omega$. However, with this arrangement there are two important points to note. Firstly that Kirchhoff's law applies and so

$$\int_{\partial\Omega} j = 0. \quad (2.21)$$

This means we can only apply Neumann conditions which satisfy this restriction. Attempts to violate Kirchhoff's law by setting current sources to an impossible j results in a 'fight' between the current sources in which they depart from their ideal behaviour. Secondly is that electric potential must always be measured with respect to some reference level, usually called 'earth'. If one simply specifies j the current sources will 'float' at some unknown average voltage. In practice one can earth some point in Ω , usually a special earth electrode on $\partial\Omega$. The voltage at this earth electrode can be adjusted to any convenient level. The simplest assumption mathematically is that the average voltage on $\partial\Omega$ is zero

$$\int_{\partial\Omega} v = 0 \quad (2.22)$$

which matches the condition placed on the current density. In a physical system any surplus current due to mismatch in the current sources will flow to earth via the earthed electrode.

Given that Equations 2.21 and 2.22 will be assumed throughout this thesis, it will be a useful convention to replace the usual Sobolev spaces $H^s(\partial\Omega)$ of

functions on the boundary by the subspace in each orthogonal to constants – that is the subspace of functions with average zero.

The physical situation shows that we cannot specify both current and voltage at the same point. We can however specify the current on a subset $S \subset \partial\Omega$ and voltage on its compliment $\partial\Omega - S$. Such a mixture of Dirichlet and Neumann data is called *mixed boundary data*. A set of boundary data will be called *sufficient* when it uniquely determines u in Ω . It is convenient in EIT also to refer to a set of boundary data as *complementary*, to a second set when the subsets on which the Dirichlet and Neumann data are defined are reversed. In the particular case that $S = \emptyset$ it may be assumed that the complementary data satisfies the average voltage zero condition Equation 2.22.

2.5.3 Weak formulation

The operator L_γ defined above is a second order differential operator. Looked at another way $L_\gamma u = 0$ is a first order partial differential equation in γ . It is useful to eliminate this occurrence of derivatives of γ so that less smoothness need be assumed for the conductivity. It is also convenient to find a formulation in which only first derivatives of u occur.

If v is any smooth test function then

$$\nabla \cdot (v\gamma\nabla u) = vL_\gamma u + \gamma\nabla v \cdot \nabla u. \quad (2.23)$$

Integrating, applying the Divergence Theorem and using Equation 2.3 yields

$$\int_{\partial\Omega} v\gamma\nabla_{\mathbf{n}} u = \int_{\Omega} \gamma\nabla v \cdot \nabla u - \int_{\Omega} vq. \quad (2.24)$$

If we assume that either $v|_{\partial\Omega} = 0$ or $\gamma\nabla_{\mathbf{n}} u = 0$ then Equation 2.24 becomes

$$\int_{\Omega} \gamma\nabla v \cdot \nabla u = \int_{\Omega} vq \quad (2.25)$$

The truth of 2.24 for all smooth v (possibly just those which vanish at the boundary) is *a priori* a weaker assumption than the truth of Equation 2.3 for $u \in C^2(\Omega)$ and $q \in C^0(\Omega)$. Moreover it is for a more general class of functions. Defining $D : H^1(\Omega) \times H^1(\Omega) \rightarrow R$ (called the Dirichlet form) by

$$D(v, u) = \int_{\Omega} \gamma\nabla v \cdot \nabla u \quad (2.26)$$

and writing $\langle \cdot, \cdot \rangle$ for the dual pairing between $H^1(\Omega)$ and $H^{-1}(\Omega)$ the weak form Equation 2.25 becomes

$$D(v, u) = \langle v, q \rangle. \quad (2.27)$$

Here u and v are distributions in $H^1(\Omega)$ and the source density q is in $H^{-1}(\Omega)$. If we wish to apply boundary conditions to Equation 2.27 then one simply restricts D to the space $X \times X$ where $X \subset H^1(\Omega)$ is the set of distributions satisfying the appropriate boundary condition. For example the Dirichlet problem is represented by $X = H_0^1(\Omega) = \{u \in H^1(\Omega) | \tau u = 0\}$ and the Neumann problem by $X = H_N^1(\partial\Omega) = \{u \in H^1(\Omega) | \gamma \nabla_{\mathbf{n}} u = 0, \int_{\partial\Omega} \tau u = 0\}$.

Before going further let us look at the physical background to the Dirichlet form. The first thing to notice is that its units are those of *power density*, that is the power dissipated per unit volume. When a current flows through a resistive medium power is dissipated through Ohmic heating. In a discrete circuit this would be V^2/R . The total power dissipated would be equal to the energy input. Taking $v = u$, Equation 2.27 is the continuum version of the discrete $V^2/R = IV$. Physically the simultaneous presence of two potentials u and v on Ω is not possible, but a knowledge of $D(u, u)$ for all u is sufficient to determine $D(v, u)$ for all u and v by application of the *polarisation identity*

$$D(v, u) = \frac{1}{4}(D(v + u, v + u) - D(v - u, v - u)) \quad (2.28)$$

which holds for any bilinear function.

In this section only Equation 2.3 has been discussed rather than the source free version Equation 2.1. In addition only zero Neumann and Dirichlet conditions have been discussed. Fortunately the two kinds of problem are interchangeable. Suppose as above X is the space of potentials satisfying the desired (possibly mixed) boundary conditions. Now let $f \in H^1(\Omega)$ be any function which agrees with the boundary data. The problem of solving $L_\gamma(u - f) = 0$ for u with trivial boundary conditions on $u - f$, that is $\tau(u - f)$ is zero on S the set where voltages were defined, and $\gamma \nabla_{\mathbf{n}}(u - f)$ is zero on the complement of S , is equivalent to the same trivial boundary value problem for $L_\gamma u = q$ where the source density is given by $q = -L_\gamma f$. For this to work the only restriction which must be imposed on the boundary conditions is that they agree with some function in $H^1(\Omega)$. The weakest hypotheses which guarantee this are $j \in H^{-1/2}(\partial\Omega)$ and $v \in H^{1/2}(\partial\Omega)$.

2.5.4 Existence , Uniqueness and Continuous Dependence

Existence and uniqueness of a solution, and the continuous dependence of that solution on the given data, are the Holy Trinity of properties which mathematicians seek to prove before they decide a problem is well posed. In the case of elliptic partial differential equations much work has been done in this area and conditions which guarantee these three properties can be found in

the literature (see [29,36,63,93]). The conditions can be stated for the weak formulation Equation 2.27 in terms of restrictions on the Dirichlet form D . A Dirichlet form D is said to be *bounded* if there is a constant C such that for all $u \in H^1(\Omega)$

$$C\|u\|_1^2 \geq D(u, u), \quad (2.29)$$

and *coercive* when there are constants c_0, c_1 so that for all $u \in H^1(\Omega)$

$$D(u, u) \geq c_1\|u\|_1^2 + c_0\|u\|_0^2. \quad (2.30)$$

In the literature hypotheses are made on γ to ensure that D is bounded and coercive. A sufficient condition is $\gamma \in L^\infty(\Omega)$ (that is $\gamma < C$ for some C) and $\gamma > c > 0$ for some constant c . This property will be called *property C*. As the domain Ω is bounded $L^\infty \subset H^0$ so property C makes sense in any Sobolev space of positive order. However, to be useful property C must be stable, that is, when a distribution with this property is perturbed slightly, the property still holds. In $H^0(\Omega)$ this is not the case. A stronger topology⁴ is needed. Clearly the subset of $C^0(\Omega)$ satisfying property C is an open set. By Sobolev's lemma (Equation 2.19) $H^s(\Omega) \subset C^0(\Omega)$ for $s > n/2$. For a two dimensional domain Ω we can take the space of conductivities to be the subset of $H^{1+\epsilon}(\Omega)$ satisfying property C , where $\epsilon > 0$. Without any danger of confusion this set will be denoted, also, by C .

Physically property C is simply that the conductivity is bounded above and below – that is the medium is neither perfectly conductive nor perfectly insulating at any point.

We now rejoin the standard theory of elliptic pdes armed with our bounded, coercive Dirichlet form $D : X \times X \rightarrow R$. First define the weak version of the operator L_γ , which will be called by the same name but is now a mapping from X to the dual space X^* . If $v \in X$ then we define $L_\gamma u = h \in X^*$ where $\langle h, v \rangle = D(u, v)$. The Lax-Milgram Lemma (see [29, 7.19]) guarantees that L_γ has a bounded, linear, inverse $G : X^* \rightarrow X$ such that for all $s \in X^*$ and $v \in X$

$$\langle v, s \rangle = D(v, Gs). \quad (2.31)$$

The existence of this mapping G , the Green's operator, tells us that the boundary value problems under consideration do indeed have a unique solution depending continuously on the data (in this case s).

2.5.5 Transfer Impedance

The system under observation in EIT can be thought of as a 'black box' which takes a current density j and returns a boundary voltage v . From the work in

⁴A stronger topology being one with more open sets.

the preceding section we can deduce that there is a continuous linear mapping from $H^{-1/2}(\partial\Omega)$ to $H^{1/2}(\partial\Omega)$ taking j to v . The usual name for the thing one multiplies the current by to get the voltage is ‘resistance’. To retain the link with this far simpler incarnation of Ohm’s law, the current to voltage mapping will be denoted by R_γ so $R_\gamma j = v$. The Dirichlet problem also has a unique solution depending continuously on the data, so we deduce that R_γ has a bounded inverse. This map R_γ is the transfer function of the black box hence the name transfer impedance operator. If one were able to make a complete set of measurements one would know R_γ and the reconstruction problem for EIT would be to invert the map $\gamma \mapsto R_\gamma$. Some interesting properties of R_γ can be found in [47]. In particular the map $\gamma \mapsto R_\gamma$ is monotone in the sense that if $\gamma_1 > \gamma_2$ then for all j we have $\langle j, R_{\gamma_1} j \rangle \leq \langle j, R_{\gamma_2} j \rangle$. This is a fancy way of saying if the resistance is increased more power is dissipated!

2.6 Compact Linear Operators

Infinite dimensional spaces will occur frequently in this thesis. These will be spaces of functions representing conductivities, currents and voltages. Linear transformations between such spaces, such as the transfer impedance operator will also feature. It is worthwhile to pause briefly to ask why it is necessary to consider infinite dimensional spaces as we can only ever make finitely many measurements and we can only manipulate finite dimensional vectors on a digital computer. There are many answers to this, perhaps the most compelling being that our physical models (on a macroscopic scale ignoring quantum theory) assume that space is a continuum and that physical quantities are defined everywhere. When we describe a physical situation in terms of a finite number of parameters we seek to do so in such a way that the discrete approximation to the continuous ‘reality’ can be made arbitrarily accurate by increasing the number of parameters. If we understand the behaviour of the continuous model we can then be assured that the behaviour of the discrete approximation will tend to that as the number of parameters increases.

Linear mappings between infinite dimensional Banach spaces such as H^s and C^k can differ alarmingly from their finite dimensional cousins. For example they may not be continuous. Those which are continuous are also called *bounded*. If $A : X \rightarrow Y$ is a linear map between Banach spaces X and Y then A is bounded when its *linear map norm* given by

$$\|A\|_{X,Y} = \sup_{x \in X} \frac{\|Ax\|_Y}{\|x\|_X}$$

is finite. For the finite dimensional case of $X = \mathbb{R}^n, Y = \mathbb{R}^m$ a linear map A is a matrix and $\|A\|_{X,Y}^2$ is the maximum eigenvalue of $\|A^T A\|$.

Amongst bounded linear maps the class of *compact mappings* share most of the properties of mappings between finite dimensional spaces. A mapping is compact⁵ when it takes closed sets in X to compact sets in Y . Details of the properties of such mappings can be found in books on functional analysis such as [78, page 98]. Intuitively a compact mapping can be thought of as one which loses information. Compact subsets of an infinite dimensional space can be thought of as rather small and flat – to illustrate this, the closed unit ball in a Banach space is only compact if the space is finite dimensional (see [78, page 17]). Thus a compact mapping ‘squashes’ the closed unit ball into a ‘smaller’ set. Such a mapping is likely to be hard to undo – a fact at the very heart of impedance tomography as will be shown in this thesis.

The space of bounded linear maps between Banach spaces X and Y will be denoted $L(X, Y)$. For $A \in L(X, Y)$ the following facts about compactness will be useful later (from [78, Theorem 4.18]):

1. If the dimension of the range $A(X)$ is finite then A is compact.
2. If A is compact and $A(X)$ is a closed subset of Y then the dimension of $A(X)$ is finite.
3. If $Y = X$ and A is compact then the multiplicity of any eigenvalue of A is finite.
4. If X is infinite dimensional and A is compact then A has no continuous inverse.
5. If $Y = X$, B is bounded and A is compact then AB and BA are also compact.

What was stated above about compact mappings being hard to undo is vindicated by 4. A compact mapping may have an inverse but it will not be continuous. Suppose one has a physical system with state space X and one can only make measurements on the system by observing a function $y \in Y$ where $y = Ax$ and A is compact. One cannot deduce the state x reliably even if A^{-1} exists for any error in y could produce an arbitrarily large error in $x = A^{-1}y$. A similar situation arises if the state x is controlled by $y = Ax$ — here a small error in the ‘setting of the controls’ y can produce arbitrarily large variations in the state x . This is the phenomenon called ill-posedness to which we shall return.

It is important to remember that the compactness of a mapping depends on the norms used to measure distance in its domain and range, just as the ill-posedness referred to above depends on how one measures error. The choice

⁵Some times also called completely continuous.

of spaces used to represent the physical quantities such as current, voltage and conductivity is far more than merely a mathematical nicety so that the theorems work. An ill-posed problem may be made well-posed by choosing a different norm, for say the state space, because this new norm is 'blind' to the error made in inverting the mapping. Two functions may be close in the H^{-1} norm and yet far away in terms of H^1 . In such a case one must look at the use to which the answer will be put. If it is important to get the high frequency components of the state correct then H^1 may provide an appropriate norm whereas the H^{-1} norm suppresses higher frequencies.

An example of this can be given using the transfer impedance operator. It forms the basis of one argument given for applying current patterns and measuring voltages in EIT, rather than the reverse.⁶ Most EIT systems which have been constructed use constant current sources to specify Neumann conditions while voltages, Dirichlet conditions, are measured. Two notable exceptions are the systems described by Kim in [53] and Yorkey in [96] which use constant voltage sources and measure current. The most general conditions on boundary current density j and boundary voltage v which guarantee finite power dissipation are $j \in H^{-1/2}(\partial\Omega)$ and $v \in H^{1/2}(\partial\Omega)$. The transfer impedance operator is then a bounded map $R_\gamma \in L(H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega))$ with a bounded inverse. However once one has passed to a discrete set of measurements it is debatable whether these fractional Sobolev norms are the natural norms to use, and they are certainly not trivial to calculate. Suppose then that one seeks a comparison using the humble L^2 (root-mean-square) norm. Since $L^2(\partial\Omega) = H^0(\partial\Omega)$ and $H^{-1/2}(\partial\Omega) \supset H^0(\partial\Omega) \supset H^{1/2}(\partial\Omega)$ one can consider the transfer impedance operator R_γ restricted to $H^0(\partial\Omega)$ as a map in $L(H^0(\partial\Omega), H^{1/2}(\partial\Omega))$. By composing this mapping with the compact embedding of $H^{1/2}(\partial\Omega)$ in $H^0(\partial\Omega)$ we obtain an $R_\gamma \in L(H^0(\partial\Omega), H^0(\partial\Omega))$ which is compact by 5 above. By 4 R_γ does not then have a continuous inverse. We would be unwise, therefore, to attempt to control the current pattern by setting voltages as a large L^2 error may result. On the other hand we may specify the currents to whatever L^2 precision our apparatus allows in the confidence that the voltages will vary continuously.

Compact mappings often occur as integral operators $A : L^2(\Omega) \rightarrow L^2(\Omega)$ defined by

$$(Af)(x) = \int_{\Omega} K(x, y)g(y) dy \quad (2.32)$$

The function K is called a *kernel* and if it satisfies

$$\int_{\Omega} \int_{\Omega} |K(x, y)|^2 dx dy < \infty \quad (2.33)$$

⁶This example was suggested by a remark made by David Isaacson at the first EEC workshop on EIT in Sheffield, 1986

then it is called a *Hilbert-Schmidt kernel*. In this case the operator (called a Hilbert-Schmidt operator) is a compact mapping. If an orthonormal basis $\{\phi_i\}$ is chosen for $L^2(\Omega)$ then A has an infinite matrix representation (a_{ij}) where

$$a_{ij} = \langle \phi_i, A\phi_j \rangle. \quad (2.34)$$

Since

$$\sum_{i,j} |a_{ij}|^2 = \int_{\Omega} \int_{\Omega} |K(x,y)|^2 dx dy \quad (2.35)$$

an alternative test for a Hilbert-Schmidt kernel is that

$$\sum_{i,j} |a_{ij}|^2 < \infty. \quad (2.36)$$

Using the example of the the homogeneous unit disk, the transfer impedance operator $R_1 : H^0(\partial D) \rightarrow H^0(\partial D)$ has a convenient matrix representation with respect to the standard Fourier basis. It is simply the diagonal matrix $a_{nm} = \delta_{nm} \frac{1}{n}$ which clearly satisfies Equation 2.36 and therefore R_1 is a Hilbert-Schmidt operator. The kernel in this case is the relevant Green's function.

2.7 Calculus in Banach spaces

In this section non-linear maps between function spaces and their calculus will be briefly reviewed. Details can be found in [28] or [1]. The main examples of such objects studied in this work will be maps which take conductivity distributions to voltages on the boundary. This will be made more explicit later. For the moment a general map $F : X \rightarrow Y$ between two Banach spaces will be considered. In fact, the most important example of such a map occurring in this thesis will only be defined on an open subset $U \subset X$, so in this case $F : U \rightarrow Y$.

2.7.1 Derivatives and Order of Convergence

The Fréchet derivative of F at a point x in U is denoted by DF_x . It is a bounded linear map $DF_x : X \rightarrow Y$ which gives the best affine approximation to F at x in the sense that

$$\lim_{h \rightarrow 0} \frac{F(x+h) - F(x) - DF_x(h)}{\|h\|} = 0. \quad (2.37)$$

When DF_x exists F is said to be differentiable at x . When DF_x is continuous as a function of x we say F is continuously differentiable, or $F \in C^1(X, Y)$. The

second derivative at x is a bilinear map $D^2F_x : X \times X \rightarrow Y$. If $G_a : X \rightarrow Y$ is defined by $G_a(x) = DF_x(a)$ for any $a \in X$ then the second derivative of F is defined using the derivative of G :

$$D^2F_x(a, b) = D(G_a)_x(b). \quad (2.38)$$

The function F is said to be twice differentiable at x when D^2F_x exists. Higher derivatives are defined similarly. When the Banach spaces in question are the familiar R^m and R^n the Fréchet derivative, when represented as a matrix in the standard basis, is simply the Jacobian matrix. Letting e_i be the i -th standard basis vector in R^m , and F_j the components of F in R^n we have

$$DF_x(e_i) = \left(\frac{\partial F_1}{\partial x_i}, \frac{\partial F_2}{\partial x_i}, \dots, \frac{\partial F_n}{\partial x_i} \right). \quad (2.39)$$

The matrix of the second derivative is the Hessian:

$$D^2F_x(e_i, e_j) = \left(\frac{\partial^2 F_1}{\partial x_i \partial x_j}, \frac{\partial^2 F_2}{\partial x_i \partial x_j}, \dots, \frac{\partial^2 F_n}{\partial x_i \partial x_j} \right). \quad (2.40)$$

In discussing approximation using derivatives it is convenient to use the concept of order of convergence. The Landau symbol $o(h^r)$ provides a useful notation⁷, it stands for any map defined in a neighbourhood of the origin of a Banach space X satisfying

$$\lim_{h \rightarrow 0} \frac{o(h^r)}{\|h\|^r} = 0 \quad (2.41)$$

Thus Equation 2.37 becomes

$$F(x + h) = F(x) + DF_x(h) + o(h). \quad (2.42)$$

2.7.2 Taylor's Theorem

If a map F is in $C^r(U, Y)$ then Taylor's Theorem states that for any $x \in U$ there is a neighbourhood V of the origin $0 \in X$ such that for $h \in V$

$$F(x + h) = F(x) + DF_x(h) + \frac{1}{2!}D^2F_x(h, h) + \dots + \frac{1}{r!}D^rF_x(h, \dots, h) + o(h^r) \quad (2.43)$$

The right hand side of Equation 2.43 is called the Taylor Series of order r of F at x . Taylors Theorem has a converse — any map which can be represented as

⁷There is a related and more common notation using O . A function $f(h) = O(h^r)$ whenever there are constants M and C such that $\|h\| < M \Rightarrow |f(h)|/\|h\|^r < C$. Clearly $O(h^{r+1}) \Rightarrow o(h^r)$.

a sum of symmetric k -linear maps (for $k \leq r$) plus a term $o(h^r)$ is in $C^r(U, Y)$ and the series is its Taylor series (see [1, page 92]). This method, as we shall see later, provides a convenient way of calculating derivatives. A map F which has continuous derivatives of all orders is called smooth, in that case one writes $F \in C^\infty(U, Y)$. When a smooth map is equal to its infinite Taylor series, that is the limit of the right hand side of Equation 2.43 as $r \rightarrow \infty$, it is called analytic and one writes $F \in C^\omega(U, Y)$.

Chapter 3

Linearisation

3.1 Introduction

The original statement of the linearisation of the forward problem appears in Calderón [22]¹. In this chapter, Calderón's techniques are used and elaborated upon to give linearised forms of the forward problem both in its direct and integral form. Calderón's result is extended to give a 'Neumann conditions constant' formulation. All proposed 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 Fréchet derivative of appropriately defined forward mappings.

3.2 Approach

If the conductivity γ is perturbed to $\gamma + \delta\gamma$ and yet one form of sufficient boundary data for u is kept constant, the complementary 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

¹Although this legendary Brazilian conference paper is somewhat hard to find

of L_γ . The reader, if unfamiliar with functional analysis, may care to skip the next section and may in subsequent sections be daunted by manipulations of differential and integral operators as though they were numbers. In that case it may be helpful to think of them as matrices, as they would be if we passed to some discrete approximation to the operators.

3.3 Choice of Space for γ

The standard theory of elliptic partial differential equations (such as that presented in Gilbarg and Trudinger [36]) requires the coefficients of the partial differential equations to be in $L^\infty(\Omega)$ which is a fairly weak assumption. To guarantee boundedness and coercivity of the operator L_γ we certainly need γ to be bounded above and away from zero, $C > \gamma > c > 0$ (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 relatively ‘nasty’ functions it has an extremely strong convergence criterion. We would certainly be unwise to compare images on the basis of their L^∞ distance — any two images could be made arbitrarily far apart in the norm by changing one pixel.

Natterer [71] suggests that the appropriate norm 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. In a Sobolev space of positive order, condition C is sufficient to guarantee boundedness and coercivity. Stability of condition C is required in this chapter as the standard theory of calculus in Banach spaces assumes functions to be defined on open sets. As stated in Section 2.5.4, this stability can be guaranteed by $\gamma \in H^{\frac{n}{2}+\epsilon}(\Omega)$, for $\epsilon > 0$.

In this chapter a bounded restriction operator will be needed. This corresponds to the existence of a natural trace operator in Sobolev spaces. As we have seen in 2.5.1 the space $H^s(\Omega)$ has a natural trace operator $\tau : H^s(\Omega) \rightarrow H^{s-1/2}(\partial\Omega)$. 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 at least $1/2$.

In a space of functions with a trace operator one can think of the function as being the sum of an interior component and a boundary component. This can be seen algebraically using the first isomorphism theorem of linear algebra. The kernel of τ is the set $\ker \tau = \{\gamma \in H^s(\Omega) | \tau\gamma = 0\} = H_0^s(\Omega)$. This is a

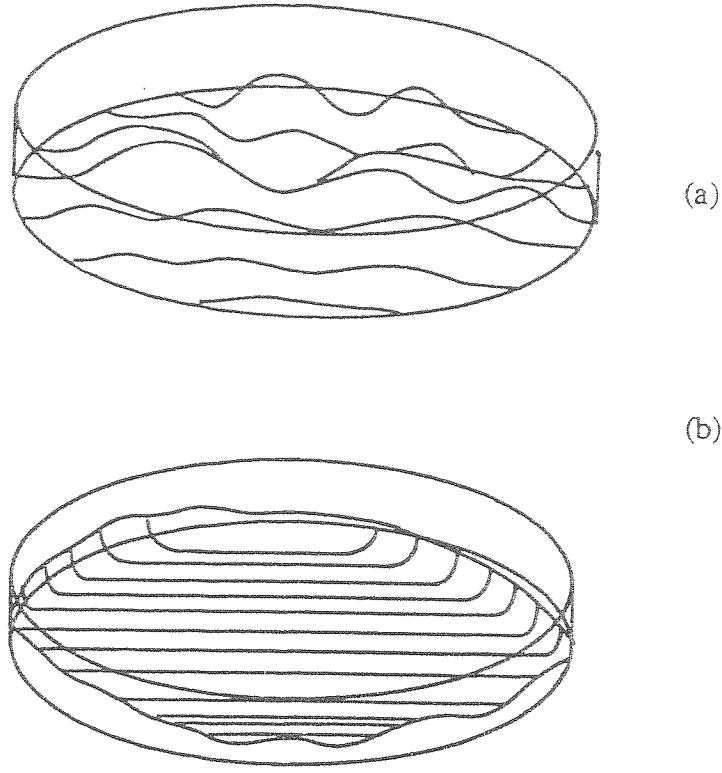


Figure 3.1: A sketch designed to give the impression of (a) γ_0 and (b) γ_∂ .

closed subspace of $H^s(\Omega)$ so has an orthogonal complement $H_0^s(\Omega)^\perp = \{\alpha \in H^s(\Omega) | \langle \alpha, \gamma \rangle = 0, \text{ for all } \gamma \in H_0^s(\Omega)\}$ which is naturally isomorphic to the image $H^{s-1/2}(\partial\Omega)$ of τ . The space $H^s(\Omega)$ can therefore be decomposed into a direct sum

$$H^s(\Omega) \oplus H_0^s(\Omega)^\perp \cong H^s(\Omega) \oplus H^{s-1/2}(\partial\Omega). \quad (3.1)$$

To simplify notation let \mathcal{C} denote the space of conductivities $\mathcal{C} = H^s(\Omega)$, the interior component by $\mathcal{C}_0 = H_0^s(\Omega)$ and boundary by $\mathcal{C}_\partial = H^{s-1/2}(\partial\Omega)$ which will be identified with $H_0^s(\Omega)^\perp$. The decomposition of a particular conductivity γ will be denoted by $\gamma = \gamma_0 + \gamma_\partial$. To visualise these components, one can think of γ_0 as being a function which vanishes on $\partial\Omega$ but, as far as the smoothness restriction allows, it agrees with γ on the interior of Ω . Conversely, γ_∂ agrees with γ on $\partial\Omega$ but, as much as an $H^s(\Omega)$ function satisfying that condition can, it vanishes in Ω (see Figure 3.1).

The open subset of \mathcal{C} consisting of those functions satisfying property \mathcal{C} will be denoted by \mathcal{C} .

3.4 Direct Form

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

$$G_\gamma^D : H_0^{-1}(\Omega) \rightarrow H_0^1(\Omega) \quad (3.2)$$

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

$$L_\gamma u = q, \quad u|_{\partial\Omega} = 0 \quad (3.3)$$

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

$$G_\gamma^N : H_N^{-1}(\Omega) \rightarrow H_N^1(\Omega) \quad (3.4)$$

solves

$$L_\gamma u = s, \quad \gamma \nabla_{\mathbf{n}} u = 0, \quad \int_{\partial\Omega} u = 0, \quad (3.5)$$

where $H_N^1(\Omega)$ is the kernel of $\gamma \nabla_{\mathbf{n}}$ in $H^1(\Omega)$.

In addition we will need the mapping

$$G_\gamma^{NB} : H^{-1/2}(\partial\Omega) \rightarrow H^1(\Omega) \quad (3.6)$$

which solves the inhomogeneous Neumann problem

$$L_\gamma u = 0, \quad \gamma \nabla_{\mathbf{n}} u = h, \quad \int_{\partial\Omega} u = 0. \quad (3.7)$$

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

Lemma 3.1 $\|L_\gamma\|_{X, X^*} \leq \|\gamma\|_C$ where X is either H_0^1 or H_N^1 and X^* is H_0^1 or H_N^1

Proof

$$\begin{aligned} \|L_\gamma u\|_{X^*} &= \sup_{w \in X \subset H^1} \frac{|\int_\Omega w L_\gamma u|}{\|w\|_{H^1}} \\ &= \sup_{w \in X} \frac{|\int_{\partial\Omega} w \gamma \nabla u \cdot \mathbf{n} - \int_\Omega \gamma \nabla w \cdot \nabla u|}{\|w\|_{H^1}} \\ &= \sup_{w \in X} \left| \int_\Omega \gamma \nabla w \cdot \nabla u \right| / \|w\|_{H^1} \\ &\leq \|\gamma\|_C \|u\|_{H^1} \end{aligned}$$

■

We are now ready to prove the following theorem:

Theorem 3.1 *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*

$$L_{\delta\gamma}u + L_\gamma\delta u = o(\delta\gamma) \quad (3.8)$$

For the Neumann constant case we also have

$$\delta\gamma\nabla_{\mathbf{n}}u + \gamma\nabla_{\mathbf{n}}\delta u = o(\delta\gamma). \quad (3.9)$$

The forward mapping defined by $F(\gamma) = u$ (which solves a given Neumann or Dirichlet problem) is in $C^\infty(C, H^1)$. For the Dirichlet case

$$DF(\gamma)\delta\gamma = -G_\gamma^D L_{\delta\gamma}u \quad (3.10)$$

and the Neumann case

$$DF(\gamma)\delta\gamma = -G_\gamma^N L_{\delta\gamma}u - G_\gamma^{NB} \delta\gamma\nabla_{\mathbf{n}}u \quad (3.11)$$

Proof

In all cases we have

$$L_{\gamma+\delta\gamma}(u + \delta u) = L_\gamma\delta u + L_{\delta\gamma}u + L_{\delta\gamma}\delta u = 0 \quad (3.12)$$

First consider the Dirichlet conditions constant case $\delta u|_{\partial\Omega} = 0$. Applying $G = G_\gamma^D$ to Equation 3.12 gives

$$(1 + GL_{\delta\gamma})\delta u = -GL_{\delta\gamma}u \quad (3.13)$$

Using standard series we have, at least formally,

$$\delta u = \sum_{k=1}^{\infty} (-GL_{\delta\gamma})^k u \quad (3.14)$$

which converges for $\|GL_{\delta\gamma}\| < 1$. Using Lemma 3.1

$$\|GL_{\delta\gamma}\| \leq \|G\| \cdot \|\delta\gamma\|_C \quad (3.15)$$

so convergence can be achieved by requiring

$$\|\delta\gamma\|_C < \frac{1}{\|G\|} \quad (3.16)$$

(the norm used for linear operators being the standard linear operator norm). The above series constitutes a Taylor series for F since, as $L_{\delta\gamma}$ is linear in $\delta\gamma$,

the k^{th} term is a k -linear symmetric map in $\delta\gamma$ (see 2.7.2). In particular we have

$$DF(\gamma)\delta\gamma = -G_\gamma^D L_{\delta\gamma}u \quad (3.17)$$

and

$$L_{\delta\gamma}u + L_\gamma\delta u = o(\delta\gamma) \quad (3.18)$$

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 $C_\partial \oplus C_\circ$ separately writing $\gamma = \gamma_\circ + \gamma_\partial$ as the sum of the components. The two partial derivatives can be calculated separately. In addition to the equation

$$L_{\delta\gamma}u + L_\gamma\delta u + L_{\delta\gamma}\delta u = 0 \quad (3.19)$$

we have its boundary equivalent

$$(\gamma + \delta\gamma)\nabla_{\mathbf{n}}(u + \delta u) = \gamma\nabla_{\mathbf{n}}u. \quad (3.20)$$

To calculate $\partial F/\partial\gamma_\circ$ we assume $\delta\gamma|_{\partial\Omega} = 0$. The boundary condition now reduces to $\gamma\nabla_{\mathbf{n}}\delta u = 0$ and the proof proceeds as before in Equations 3.13–3.16 with $G = G_\gamma^N$ and we have

$$\delta u = \sum_{k=1}^{\infty} (-G_\gamma^N L_{\delta\gamma})^k u \quad (3.21)$$

and thus

$$\frac{\partial F}{\partial\gamma_\circ}\delta\gamma_\circ = -G_\gamma^N L_{\delta\gamma_\circ}u. \quad (3.22)$$

On the other hand $\partial F/\partial\gamma_\partial$ can be calculated by assuming that $\gamma = 0$ in the interior $\overset{\circ}{\Omega}$. This leads to

$$L_\gamma\delta u = 0 \quad (3.23)$$

and

$$\gamma\nabla_{\mathbf{n}}\delta u = -\delta\gamma_\partial\nabla_{\mathbf{n}}u - \delta\gamma_\partial\nabla_{\mathbf{n}}\delta u. \quad (3.24)$$

The proof proceeds in a similar way to the interior case. Applying G_γ^{NB}

$$\delta u = G_\gamma^{NB}(-\delta\gamma_\partial\nabla_{\mathbf{n}}u - \delta\gamma_\partial\nabla_{\mathbf{n}}\delta u) \quad (3.25)$$

and rearranging we obtain

$$(1 + G_\gamma^{NB}\delta\gamma_\partial\nabla_{\mathbf{n}})\delta u = -G_\gamma^{NB}\delta\gamma_\partial\nabla_{\mathbf{n}}u. \quad (3.26)$$

Using standard series as before we have, at least formally,

$$\delta u = \sum_{k=1}^{\infty} (-G_{\gamma}^{NB} \delta \gamma_{\partial} \nabla_{\mathbf{n}})^k u \quad (3.27)$$

which converges for $\|G_{\gamma}^{NB} \delta \gamma_{\partial} \nabla_{\mathbf{n}}\| < 1$ which is ensured by $\|\delta \gamma_{\partial}\| < 1/(\|G_{\gamma}^{NB}\| \cdot \|\nabla_{\mathbf{n}}\|)$. Again, by Taylor's Theorem we have the derivative

$$\frac{\partial F}{\partial \gamma_{\partial}} \delta \gamma_{\partial} = -G_{\gamma}^{NB} \delta \gamma_{\partial} \nabla_{\mathbf{n}} u \quad (3.28)$$

and the result

$$\delta \gamma_{\partial} \nabla_{\mathbf{n}} u + \gamma \nabla_{\mathbf{n}} \delta u = o(\delta \gamma_{\partial}). \quad (3.29)$$

From the chain rule

$$DF(\gamma) \delta \gamma = -G_{\gamma}^N L_{\delta \gamma} u - G_{\gamma}^{NB} \delta \gamma_{\partial} \nabla_{\mathbf{n}} u. \quad (3.30)$$

From the converse to Taylor's Theorem in 2.7.2 we see that $F \in C^{\infty}(\mathcal{C}, H^1)$. ■

The proof in the Neumann case can be simplified by defining a new operator which combines G^N with G^B .

Define $\bar{L}_{\gamma} : H^1(\Omega) \rightarrow H_N^{-1}(\Omega) \oplus H^{-1/2}(\partial\Omega)$ by

$$\bar{L}_{\gamma} u = (L_{\gamma}, -\gamma \nabla_{\mathbf{n}} u). \quad (3.31)$$

This has a continuous inverse \bar{G}_{γ} given by

$$\bar{G}(q, j) = G^N q + G^{NB} j. \quad (3.32)$$

The Neumann case is now like the Dirichlet case with

$$\delta u = \sum_{k=1}^{\infty} (-\bar{G} \bar{L}_{\delta \gamma})^k u \quad (3.33)$$

and

$$DF(\gamma) \delta \gamma = -\bar{G}_{\gamma} \bar{L}_{\delta \gamma} u. \quad (3.34)$$

In both cases F is equal to its Taylor expansion in a neighbourhood of γ and is therefore an analytic mapping (this was pointed out by Calderón for the particular case $\gamma = 1$). 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.

3.5 Interpretation as a source

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

$$\nabla \cdot \gamma \nabla \delta u = -\nabla \cdot \delta \gamma \nabla u. \quad (3.35)$$

One interpretation of this is that the perturbed field δu is ‘caused’ by a distribution of current sources $-\nabla \cdot \delta \gamma \nabla u$. Equivalently, one could say that adding a source field $q = \nabla \cdot \delta \gamma \nabla u$ 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 of u , that is $\nabla u(p) \neq 0$. For simplicity we will take $\gamma = 1$. The source term is now $-\nabla \delta_p \cdot \nabla u$ 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 dipole. 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 $\|\delta u|_{\partial\Omega}\|_{H^{1/2}(\partial\Omega)} = O(1/\rho^2)$ where $\rho = \text{dist}(p, \partial\Omega)$. 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.

3.6 Integral Form

It is useful to re-formulate 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, which are 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 mapping as an integral operator.

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

$$\nabla \cdot (U \gamma \nabla V) = \gamma \nabla U \cdot \nabla V \quad (3.36)$$

and so using the divergence theorem

$$\int_{\partial\Omega} V \gamma \nabla_{\mathbf{n}} U = \int_{\Omega} \gamma \nabla V \cdot \nabla U. \quad (3.37)$$

From above

$$L_{\delta\gamma}u + L_{\gamma}\delta u = o(\delta\gamma) \quad (3.38)$$

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

$$\nabla \cdot w(\gamma \nabla \delta u + \delta \gamma \nabla u) = \gamma \nabla w \cdot \nabla \delta u + \delta \gamma \nabla w \cdot \nabla u + o(\delta\gamma). \quad (3.39)$$

Applying the divergence theorem

$$\int_{\partial\Omega} w(\gamma \nabla_{\mathbf{n}} \delta u + \delta \gamma \nabla_{\mathbf{n}} u) = \int_{\Omega} \gamma \nabla w \cdot \nabla \delta u + \delta \gamma \nabla w \cdot \nabla u + o(\delta\gamma). \quad (3.40)$$

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

$$\int_{\partial\Omega} \delta u \gamma \nabla_{\mathbf{n}} w = \int_{\Omega} \gamma \nabla \delta u \cdot \nabla w \quad (3.41)$$

and using the boundary condition $\gamma \nabla \delta u \nabla_{\mathbf{n}} + \delta \gamma \nabla u \nabla_{\mathbf{n}} = o(\delta\gamma)$ the result is

$$\int_{\partial\Omega} \delta u \gamma \nabla_{\mathbf{n}} w = - \int_{\Omega} \delta \gamma \nabla w \cdot \nabla u + o(\delta\gamma). \quad (3.42)$$

Notice the negative sign. If current is kept constant and the conductivity increased, more power will be dissipated.

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

$$V_{i,k} = \int_{\partial\Omega} u_i \omega_k \quad (3.43)$$

where ω_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

$$\delta V_{i,k} = \int_{\partial\Omega} \delta u_i \omega_k \quad (3.44)$$

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

$$\delta V_{i,k} = - \int_{\Omega} \delta\gamma \nabla w_k \cdot \nabla u_i \quad (3.45)$$

In this formulation the Neumann conditions were kept constant. As we have seen in Chapter 2, this is the most useful formulation for impedance measurement for both theoretical and practical reasons. For completeness, we must compare this to the problem investigated by Calderón 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 $j = -\gamma \nabla_{\mathbf{n}} u$ and

$$\delta j = -\delta(\gamma \nabla_{\mathbf{n}} u) = \delta\gamma \nabla_{\mathbf{n}} u + \gamma \nabla_{\mathbf{n}} \delta u \quad (3.46)$$

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

$$\int_{\partial\Omega} w \delta(\gamma \nabla_{\mathbf{n}} w) = \int_{\Omega} \delta\gamma \nabla w \cdot \nabla u + o(\delta\gamma) \quad (3.47)$$

which does not have the minus sign as the power dissipated will *increase* if the voltage is held constant and the conductivity increased.

3.7 Operator Form

It will be helpful in subsequent chapters to include the operators R_γ and G_γ^{NB} in the integral form of the derivative. In this case let $F : \mathcal{C} \rightarrow L(H^{-1/2}, H^{1/2})$ be the map $F(\gamma) = R_\gamma$. In this case $DF_\gamma(\delta\gamma)$ is the limit of the difference of transfer impedance operators. To obtain the matrix for such an operator we choose a basis j_1, j_2, \dots of currents then

$$\langle j_k, DF_\gamma(\delta\gamma) j_\ell \rangle = - \int_{\Omega} \delta\gamma \nabla(G_\gamma^{NB} j_k) \cdot \nabla(G_\gamma^{NB} j_\ell), \quad (3.48)$$

The properties of this operator are discussed in more detail in Chapter 5.

3.8 Translation to Inverse scattering problem

Although this approach is not explored fully in this thesis it is worth noting that there is a correspondence between the identification problem for the conductivity by boundary measurement and the problem of finding an unknown

refractive index from similar data. The correspondence is formalised in the following lemma (stated in [54]).

Lemma 3.2 *If $w = \gamma^{1/2}u$ and $\eta = -\Delta\gamma^{1/2}/\gamma^{1/2}$ then*

$$\nabla \cdot \gamma \nabla u = 0 \Leftrightarrow \Delta w + \eta w = 0. \quad (3.49)$$

Proof

$$\begin{aligned} \gamma \nabla u &= -w\gamma^{-1/2}\nabla\gamma/2 + \gamma^{1/2}\nabla w \\ \nabla \cdot \gamma \nabla u &= \nabla \cdot (\gamma^{1/2}\nabla w - \gamma^{-1/2}w\nabla\gamma/2) \\ &= \gamma^{-1/2}\nabla \cdot \gamma \nabla w/2 - \gamma^{-3/2}w|\nabla\gamma|^2/4 - w\gamma^{1/2}\nabla^2\gamma/2 \\ &= 0 \\ &\Leftrightarrow \\ 0 &= \nabla^2 w + \left(\frac{|\nabla\gamma|^2}{4\gamma^2} - \frac{\nabla^2\gamma}{2\gamma}\right)w \end{aligned}$$

which is equivalent to

$$\nabla^2 w + \eta w = 0 \quad (3.50)$$

where

$$\eta = \frac{|\nabla\gamma|^2}{4\gamma^2} - \frac{\nabla^2\gamma}{2\gamma} = -\frac{\Delta\gamma^{1/2}}{\gamma^{1/2}}. \quad (3.51)$$

■

In the new formulation, the function to be identified η appears with no derivatives. The identification problems for γ and η are closely related : η can be deduced from γ by definition and given η we must solve the elliptic partial differential equation

$$\Delta\gamma^{1/2} + \eta\gamma^{1/2} = 0. \quad (3.52)$$

This equation has a unique solution given sufficient boundary data for γ , for example, if γ is known on the boundary. This leads to another formulation of the linearized system to be solved to identify γ . Define

$$\mathcal{L}_\eta w = \Delta w + \eta w \quad (3.53)$$

and \mathcal{G}_η the inverse of $\mathcal{L}_\eta : H_N^1 \rightarrow H_N^{-1}$ then

Theorem 3.2 *If $F(\eta) = w$ where w solves $\mathcal{L}_\eta w = 0$ then $DF_\eta(\delta\eta) = -\mathcal{G}_\eta(\delta\eta w)$.*

Proof

Assume

$$\mathcal{L}_{\eta+\delta\eta}(w + \delta w) = 0 \quad (3.54)$$

that is

$$\Delta(w + \delta w) + (\eta + \delta\eta)(w + \delta w) = 0 \quad (3.55)$$

hence

$$(\mathcal{L}_\eta - \delta\eta)\delta w = -\delta\eta w. \quad (3.56)$$

Applying \mathcal{G}_η to both sides we get

$$(1 - \mathcal{G}_\eta \delta\eta)\delta w = -\mathcal{G}_\eta \delta\eta w \quad (3.57)$$

and as before

$$\delta w = -\sum_{j=1}^{\infty} (\mathcal{G}_\eta \delta\eta)^{j+1} w \quad (3.58)$$

which converges for $\|\delta\eta\| < \|\mathcal{G}_\eta\|^{-1}$. The result now follows

$$DF_\eta(\delta\eta) = \delta w = -\mathcal{G}_\eta(\delta\eta w) \quad (3.59)$$

■

As a corollary we have

Corrolary 3.1 *A first order approximation to the inverse problem is to find a $\delta\eta$ consistant with the Fredholm equation*

$$\delta w = -\mathcal{G}_\eta(\delta\eta w) \quad (3.60)$$

or defining g_η by $(\mathcal{G}_\eta f)(x) = \int_\Omega g_\eta(x, y) f(y) dy$

$$(\delta w)(x) = \int_{y \in \Omega} g_\eta(x, y) \delta\eta(y) w(y) dy \quad (3.61)$$

Identifying the coefficient η may be of interest in itself. It is hard to understand the physical significance of this quantity (its units are *per area*) but it may never the less be a useful quantity to image. One can see from its relation to the conductivity that it will yield an edge enhanced version of the conductivity image, the Laplacian operator being an edge detector. One may want to recover γ from η in a quantitative study. This would be reasonable as the mapping taking η to γ is a smoothing map. The reverse however would result in an amplification of noise inherent in the differentiation process. In Sobolev space terms, $\eta \in H^s$ means that $\gamma \in H^{s+2}$.

3.9 Historical Note

Both the integral and direct forms of formula for the derivative appear without proof in the EIT literature. The integral form can be traced to Geselowitz [32] and was well known in Japan, being cited by Nakayama et al [95] in 1981, Sakamoto and Kanai [79] in 1983 and by Murai and Kagawa [67] in 1985. Both Murai and Kagawa and Nakayama et al were also aware of a discrete version relating to resistor networks. In 1986, at the first European Workshop on EIT, Yorkey and Webster [97] described how the derivative matrix could be calculated using a discrete formulation. None of these authors gave or referred to a mathematical proof that the integral formula for the derivative was correct. Yamashita and Takahashi [94] give a matrix series derivation. Breckon and Pidcock [13] state that the derivative formula could be deduced rigorously as an extension of the work of Calderón [22].

The direct form exhibits itself in the correct assumption by Barber and Brown [5] that the term $L_{\delta\gamma}\delta u$ could be neglected to first order although no justification is given. It also appears in Yorkey's compensation method (see [96] and Section 4.7).

Chapter 4

Finite Element Modelling

The Buddha, the Godhead, resides quite as comfortably in the gears of a motorcycle transmission or the circuits of a digital computer as at the top of a mountain or in the petals of a lotus flower.

Robert Pirsig, *Zen and the Art of Motorcycle Maintenance*

4.1 Choice of Forward Modelling Technique

All methods proposed for EIT reconstruction require some method of calculating the potential u , or at least the electric field $E = \nabla u$, from given sufficient boundary conditions and conductivity. Analytical methods for solving these field equations are restricted to simple conductivity distributions. While domains with irregular boundaries (at least in two dimensions) can be handled to some extent by analytical methods using conformal mapping techniques, this is far from easy. The restriction on the complexity of the conductivity makes analytical methods useful only for calculating an initial approximation to the potential from a simple first guess for the conductivity. An example of this is the Barber-Brown [5] method. If one wishes to start from a more elaborate initial guess, or iterate after a first approximation to the conductivity has been calculated, it is necessary to use a numerical technique.

The two most readily available numerical methods for solving partial differential equations are the Finite Element Method (FEM) and the Finite Difference Method (FDM). In the FDM the potential is approximated by its values at the nodes of a regular rectangular mesh. In the FEM the domain is decomposed into irregular polyhedra or polygons (called finite elements) and the potential is approximated using finite element basis functions to interpolate between specified nodal values. The basis functions are polynomial within the

elements and the resulting approximation space is thus piecewise polynomial. It can be shown that the FDM is a special case of the FEM with a particularly simple mesh. The FEM has the advantage that irregular elements can be fitted accurately to the boundaries of irregularly shaped domains — such as human bodies. Also the size of the elements can vary within the domain to enable the fields to be calculated with greater accuracy in some regions. This is useful, for example, to give better accuracy in a region of higher field strength such as that near an electrode. The FDM has the advantage that the regular grids are easy to generate, the programs easier to write, and the results are easier to display. Another advantage is that fast general purpose multi-grid solvers are readily available for simple regular meshes but not for the more complicated finite element meshes.

Nevertheless, it was decided from the outset to use the FEM in this work. The use of this method had a precedent in EIT research in the work of Tarassenko [89] and also of Kim [52]. The later works of Yorkey [96,97] also use the FEM which has now become a familiar tool in the EIT reconstruction field. Following Tarassenko we choose to use the Finite Element Library of Greenough and Robinson [39] now distributed by Numerical Algorithms Group (NAG).

4.2 Theory of the Finite Element Method

4.2.1 Preliminaries

It seems appropriate here to give a brief summary of the aspects of the finite element method pertinent to this work. For a more detailed treatment the reader is referred to the extensive literature on the subject including works written specifically from an Engineering, Physics or Mathematics standpoint. Readers should choose a reference which appeals to their own perspective. For the mathematically inclined, Strang and Fix [85] and Zienkiewicz [99] were found to be particularly useful. For an introduction to the practicalities of writing finite element programs Greenough and Robinson [39, Vol 1] serves as a practical guide to methodology as well as a user's guide for the NAG Finite Element Library.

The essence of the Finite Element Method is to find the closest approximation to the solution of the weak form of a differential equation which lies in an finite dimensional approximation space of certain piecewise polynomial functions. This approximate solution can be found by solving a system of linear equations, the number of equations being equal to the dimension of the approximation space. In the case of the equation $L_\gamma u = q$ (with $u|_{\partial\Omega} = 0$)

recall from 2.5.3 that the equivalent weak form is

$$\int_{\Omega} (\gamma \nabla u \cdot \nabla w - qw) = 0$$

for all test functions w . If Dirichlet boundary conditions are to be imposed then the class of test functions is restricted to those which satisfy the boundary condition. Neumann conditions $-\gamma \nabla_{\mathbf{n}} u = j$ are imposed by replacing q by $q + f$ for any f satisfying the Neumann condition $-\gamma \nabla_{\mathbf{n}} f = j$ with $L_{\gamma} f = 0$. Integration by parts leads to the formula

$$\int_{\Omega} (\gamma \nabla u \cdot \nabla w - qw) = \int_{\partial\Omega} wj.$$

Neumann conditions can thus conveniently be thought of as (current) sources localised at the boundary whereas q represents interior sources. An approximation space can now be chosen in which to represent u .

4.2.2 Approximation Space

The starting point for constructing a finite element approximation space is to divide the domain into a finite number of irregular *elements* which in two dimensions will be polygons, typically triangles or quadrilaterals (see Figure 4.2.2), and in three dimensions could be 'bricks' (irregular cuboids) tetrahedra or triangular prisms. Each element has a number of *nodes* which include the vertices of the elements and possibly other points which may be interior or on the boundary. Each element must be *non-degenerate* in the sense of no coincident faces being parallel. Also, the partition of the domain into elements must be *consistent*, meaning that if a node lies on a face of more than one element it must be a node in both elements. A basis for the approximation space is formed by first considering a single element. The element is given a local coordinate system in which it is symmetrical. An example of a triangular element is illustrated in Figure 4.2.2. The local coordinates in this case are $\xi = (\xi_1, \xi_2)$. A *local basis* consists of a set of polynomials $N_1(\xi), \dots, N_n(\xi)$ in the local coordinates where n is the number of *degrees of freedom* associated with the element. To emphasise that a local basis function belongs to a particular element e it will be denoted by N_i^e . In simpler elements the number of degrees of freedom will be the same as the number of nodes in the element and in this case we have $N_i(\xi^j) = \delta_{ij}$ where ξ^j is the coordinate of the j^{th} node. A function $f(\xi)$ can then be approximated by $\hat{f}(\xi) = \sum_{i=1}^n f(\xi^i) N_i(\xi)$. The term *element* will be used to describe both the geometric shape, and the shape together with its basis functions. Higher order elements may have extra degrees of freedom associated with derivatives of f . If the first r derivatives were represented at each node, then the number of degrees of freedom would

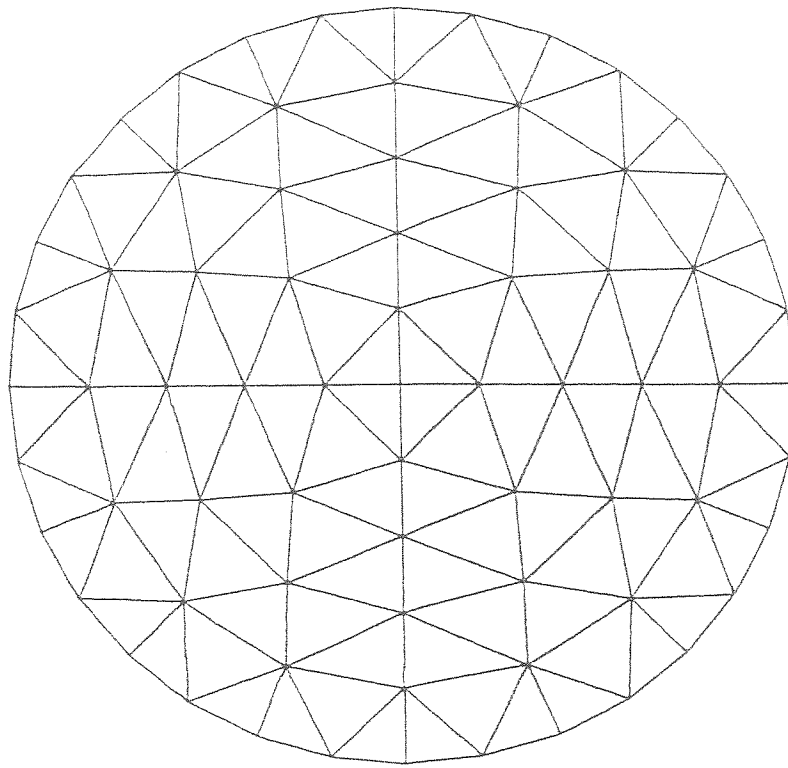


Figure 4.1: *An example finite element mesh using triangular elements*

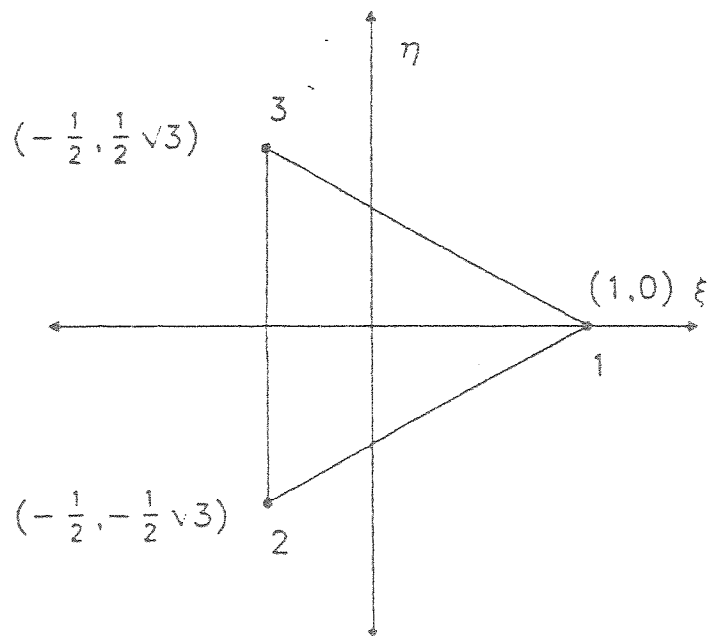


Figure 4.2: *A three noded triangular element in the local coordinate system.*

be $n\tau$ and the local basis $N_{i,j}(\xi)$, $1 \leq i \leq n$, $0 \leq j \leq \tau$. The approximation to $f(\xi)$ would then be

$$\tilde{f}(\xi) = \sum_{i=1}^n \sum_{j=0}^{\tau} f(\xi^i) \frac{\partial^j}{\partial \xi_j} N_{i,j}(\xi). \quad (4.1)$$

For the three-noded triangular element shown in Figure 4.2.2, the basis functions the basis functions are the first order polynomials

$$N_i(\xi) = a_i + b_i \xi_1 + c_i \xi_2. \quad (4.2)$$

For a finite element mesh \mathcal{M} the approximation space $S(\mathcal{M})$ is given by $\text{span}\{N_i^e | e \in \mathcal{M}, 1 \leq i \leq n(e)\}$. The local basis functions are chosen so that the first τ derivatives agree on the faces of the elements. This guarantees that $S(\mathcal{M}) \subset C^\tau(\Omega)$ and $\frac{\partial^\tau}{\partial \xi_r^\tau}$ will be piecewise constant and, as Ω is compact, bounded. Thus $S(\mathcal{M}) \subset H^{r+1}(\Omega)$. (Note that Sobolev's Lemma (see 2.5.1) guarantees that any function in C^τ is also in $H^{r+n/2-\epsilon}$. For $n = 2$ this just fails to show that $S(\mathcal{M}) \subset H^{r+1}(\Omega)$, but that for $n = 3$ it tells us that actually $S(\mathcal{M}) \subset H^{r+\frac{3}{2}-\epsilon}(\Omega)$ for some $\epsilon > 0$). As stated in Chapter 2, we are assuming that the power dissipated is finite and the weakest condition which guarantees this (at least within the L^2 framework) is $u \in H^1(\Omega)$. This suggests using $C^0(\Omega)$ elements to approximate the potential. For algorithms where the direction of the electric field is used explicitly it would be more appropriate to use $C^1(\Omega)$ elements.

The essential quality of a scale of approximation spaces is that they refine in such a way as to guarantee an arbitrarily accurate approximation. With finite element spaces there are two distinct methods of refinement. The mesh itself can be refined reducing the size of the elements or the order of the polynomials used can be increased increasing the number of degrees of freedom per element.

The convergence of the finite element approximation \tilde{u} to the solution u of a second order elliptic pde as the maximum element radius h and the degree of the polynomials used $k \geq 1$ can be summarised in a formula given in [85]:

$$\frac{\|u - \tilde{u}\|_s}{\|u\|_{k+1}} \leq Ch^{k+1-s} \quad (4.3)$$

where s is non-negative. As a particular example, let us consider a finite element space using three-node, triangular elements of order $k = 1$ such as those in Figure 4.2.2. In this case, the L^2 error in the finite element approximation to the potential has order of convergence $O(h^2)$ as $s = 0$. The error in the gradient of the solution, as measured by the norm $\|\cdot\|_1$, is only $O(h)$. Notice that the left hand side of Equation 4.3 is the *relative* error, compared with the

the Sobolev $k + 1$ norm of the actual potential. The inequality can be applied locally to show that in a region where any of the first $k + 1$ derivatives of u are large, however the *absolute error* $\|u - \tilde{u}\|$, will be larger.

4.2.3 The System Matrix

The finite element equivalent of the operator L_γ is called the *system matrix*. It is simply a representation of the Dirichlet form in the finite element basis. If A is the system matrix then

$$a_{k\ell} = \int_{\Omega} \gamma \nabla N_k \cdot \nabla N_\ell. \quad (4.4)$$

To calculate the system matrix it is necessary to calculate the inner product of the gradients of the basis functions and perform the integration. The usual approach is to treat each element in turn changing variables to the local coordinates ξ . It is necessary, therefore, to calculate the Jacobian $J = \partial \xi / \partial x$. The *element matrix* for the element e is then A^e where

$$a_{k,\ell}^e = \int_e \gamma(J \nabla N_k^e) \cdot \gamma(J \nabla N_\ell^e) (\det J)^{-1} d\xi. \quad (4.5)$$

The integration is typically performed numerically using a quadrature rule. The system matrix is then *assembled* from sums of the element matrices.

The system matrix inherits the properties of the Dirichlet form — it is symmetric and positive definite. In the finite element basis it is also *banded* — that is entries in the matrix more than b away from the leading diagonal are zero; here b is called the semi-bandwidth. The bandwidth is determined by the numbering of the nodes and is equal to the maximum difference in numbers of nodes within the same element. This banded structure can be exploited to speed numerical inversion of the system matrix and to reduce computer memory requirements. Analytically, the bandedness is due to L_γ being a local operator.

To represent the trivial Neumann conditions problem

$$L_\gamma u = q, \quad \gamma \nabla_n u = 0 \quad (4.6)$$

in finite element form, u and q must be represented as vectors of nodal values \underline{u} and \underline{q} , then

$$A \underline{u} = \underline{q}. \quad (4.7)$$

4.2.4 Application of Boundary Conditions

Boundary conditions can now be applied to Equation 4.7. For Neumann conditions this is very similar to the continuous case. To apply the condition

$-\gamma \nabla_{\mathbf{n}} u = j$ a function f is found with $-\gamma \nabla_{\mathbf{n}} f = j$ and the trivial Neumann problem for $s = L_{\gamma} f$ is solved. Finding such an f is very simple in the finite element representation — it is simply a vector \underline{f} supported only at boundary nodes. In the finite element representation, Neumann conditions are simply sources at the boundary. Dirichlet conditions can be applied by one of two methods. The first is simply to eliminate the components of \underline{u} corresponding to boundary nodes where the potential is known. This is very straight-forward when the boundary condition is zero — that line in the linear equation is simply deleted. For more complicated conditions it can destroy the banded structure of the system matrix. The other way commonly used is the Payne-Irons ‘big spring’ method. Details can be found in [39].

4.3 Specification of Conductivities

Previous authors using finite elements to model the potential in EIT (for example Yorkey [96] and Tarassenko [89]) have modelled the conductivity as piecewise constant on elements. While it is reasonable to approximate the conductivity distribution of the body as piece-wise constant by assuming that the individual organs are homogeneous it is not possible to make organ boundaries lie only on the boundaries between elements without *a priori* knowledge of the positions and shapes of the organs. As this information is not available it would be necessary to include the positions of the elements as free variables to be fitted to the observed measurements. That path may well prove fruitful but it is not the one investigated in this thesis.

The basis functions in the finite element model used to calculate the potentials were already available and have useful properties for approximating the conductivity. In addition the assumption that the conductivity lies in a given finite element space is a form of regularisation for ill-posed problems (see Natterer [73]). In this thesis the assumption for the analytical work has been $\gamma \in \mathcal{C} \subset H^{n/2+\epsilon}(\Omega)$. It is necessary to ask how this assumption is related to the representation of γ in a finite element space. Let us consider the particular case of a two dimensional domain with piece-wise linear elements. The first derivatives of the nodal basis functions for this finite element space will be in $H^{1/2-\epsilon}$ as they are characteristic functions of domains with sufficiently regular boundaries (as shown in [71]). We can conclude that the nodal basis functions themselves lie in $H^{3/2-\epsilon}$. For any $\epsilon < 1/2$ this is consistent with the assumed smoothness of the conductivity.

Another issue associated with the approximation space used for the conductivity is the stability of boundedness property (more especially property \mathcal{C} as the mesh is refined. A result which sheds some light on this topic (and

indeed why the FEM works at all!) is the inverse estimate of Suli [86]. If $v \in H_0^1(R^2)$ and v_h is its finite element approximation in a finite element space of piece-wise linear elements with radius h , then

$$\|v_h\|_\infty \leq C \log 1/h \|v_h\|_{H_0^1}. \quad (4.8)$$

This shows that the upper and lower bounds of the finite element approximation cannot be controlled by a bound on the H^1 norm. However the growth of the upper or lower bounds will be at worst logarithmic in the element size.

4.4 The NAG Finite Element Library

When faced in 1985 with the question of how to write a forward modelling program for EIT the following considerations were important:

- A capability for solving two and three dimensional problems was required.
- A solver must be incorporated in a larger EIT reconstruction program.
- Both Neumann and Dirichlet boundary conditions may need to be applied.
- An interface with Fortran 77 was required to facilitate use of other numerical software libraries.
- A large selection of element types was needed.

The NAG finite element library [39] satisfied these criteria. The library consists of a collection of Fortran subroutines which perform fairly low level operations, such as matrix arithmetic, basis function calculation, calculating quadrature points, system matrix assembly and solution of linear equations. There is also a set of example programs illustrating the use of the library to solve specific (two dimensional) problems. These examples illustrate the data structures used to represent the finite element mesh and the system matrix, and algorithms for matrix assembly and solution. The library contains no mesh generation, node numbering or graphical display routines.

4.5 Implementation of the Forward Modelling Program

4.5.1 History of fwprob

A flexible forward modelling subroutine called `fwprob` has been implemented using the NAG Finite Element Library. This code has evolved considerably over its four years of life. The first version was implemented on a PR1ME minicomputer in early 1986. It was ported to an Apollo Domain workstation in 1987 and then again to Sun 360 in 1988. Later the same year it was transferred to a Sun 386i. This version has now been ported by Paulson to a Transputer system which uses the Sun 386i as a host [17]. Paulson's version, written in Parallel FORTRAN, uses concurrent programming techniques to effect a considerable increase in speed.

4.5.2 Choice of Language

With the benefit of hindsight, the use of a more modern programming language than Fortran 77 would have been an advantage. In particular, more flexible data structures (such as structures in C or records in Pascal) would have made the representation of finite element meshes much clearer. Also it is the nature of finite element programs that objects (for example elements) have both data and operations associated with them. This suggests that an object oriented programming approach would be advantageous. An expedient choice of programming language at the time of writing would be C++, combining the features of Fortran essential to the numerical analyst such as adjustable array sizes and double precision arithmetic, with facilities for object oriented programming. In an effort to retain the portability of Fortran while adding some extensions (such as included source files and global constants) the standard Unix macro package `m4` is used as a pre-processor for all code.

4.5.3 Program Details

In the documentation for the NAG-FEL [39, Vol. 2,3.2], an example program is given to solve a second order elliptic partial differential equation in divergence form, with Neumann and Dirichlet boundary conditions. This was used as a starting point for the development of `fwprob`. The first modification needed was to allow for a conductivity specified in the finite element space, rather than as a function given in terms of the global coordinates. Further modifications have been mainly concerned with increasing the speed of execu-

tion. Those readers wishing to understand the detailed workings of `fwprob` are recommended to read [39, Vol. 2,3.2] so as to understand the programming conventions and methodology used.

It was found that using the example program assembly of the system matrix took significantly longer than the solution of the linear system. This was due to the large number of arithmetic operations and subroutine calls required to calculate the quantities

$$s(e, i, j, k) = \int_e N_i^e \nabla N_j^e \cdot \nabla N_k^e \quad (4.9)$$

for each element $e \in \mathcal{M}$. These quantities depend only on the mesh and not on the conductivity or boundary conditions and thus can be calculated once for a given mesh and stored. Subroutine `precalc` calculates the array `s`.

The first call to `fwprob` assembles the system matrix. It then calculates the Choleski factor of this matrix and solves for the potential given the boundary conditions described in the arrays `bdcnd` and `bval`. On subsequent calls if the system matrix is unchanged the logical variable `recalc` can be set to `false`, and `fwprob` simply assembles the right hand side `rhs` (which is the array representing the source term q) according to the boundary conditions and performs forward and back substitution.

4.5.4 Trouble Shooting

It seems worth recording some of the most common conditions which can cause the routine `fwprob` to fail, as these will no doubt also be encountered by others implementing finite element programs for EIT. The most common error message generated comes from the Choleski factorisation routine `CHOSOL` indicating that the system matrix is not positive definite. The most common cause for this is that the infimum of the conductivity is too small. This is the numerical equivalent of a violation of condition \mathcal{C} . The smallest feasible conductivity will depend on machine precision, but an insulating region can be represented by this value.

A second possible error is that the boundary conditions are specified incorrectly. If no Dirichlet conditions are specified, the matrix will have a rank deficiency of one and again will not be positive definite. The trivial Dirichlet condition of ‘earthing one node’ can be specified by deleting the column in the system matrix corresponding to that node. Other Dirichlet conditions can be included using the Payne-Irons technique.

4.6 Mesh Generation, Numbering and Refinement

While regular meshes on rectangular domains are trivial to generate, circular domains present something of a problem. The program `rmesh` generates three-noded triangular meshes for a circular domain. The input to the program is a sequence of decreasing radii r_i and numbers of nodes n_i on the concentric circles. The numbers of nodes must satisfy $k(n_i - n_{i+1}) = n_i$ for some integer k . The algorithm reduces the number of triangular elements in successive concentric annuli by collapsing every k^{th} element to a radial line (see Figure 4.2.2). As stated in Equation 4.3, the relative accuracy of the finite element approximation depends on the maximum radius of the elements. If a uniformly accurate approximation is required the size of the elements must decrease in proportion to the size of the first two derivatives of u on the element, as measured by $\|u\|_{H^2(e)}$. Ideally the mesh would be ‘tuned’ to a particular set of boundary conditions. Supposing that the conductivity were uniform, a current pattern of $j = \cos k\theta$ would have $\partial^2 u / \partial r^2 = (k-1)r^{k-2} \cos k\theta$ (for $k \geq 2$). In this case uniform approximation would be guaranteed by an element radius given by $h(r) = Cr^{2-k}$. As a general principle, the elements should be smaller near the boundary.

The program `rmesh` initially generates nodes ordered lexicographically in the polar coordinates : $(r_1, \theta_1) \leq (r_2, \theta_2) \iff (r_1 < r_2) \text{ or } (r_1 = r_2 \text{ and } \theta_1 \leq \theta_2)$. This produces a semi-bandwidth equal to $n_1 + 1$, where n_1 is the number of nodes on the boundary. To reduce the bandwidth of the system matrix the nodes were re-ordered according to the ordering $(x_1, y_1) \leq (x_2, y_2) \iff (x_1 < x_2) \text{ or } (x_1 = x_2 \text{ and } y_1 \leq y_2)$, that is lexicographically in (x, y) . This ordering reduces the bandwidth by a factor of up to π , the ratio of circumference to diameter. Because the nodes in these circular meshes do not ‘line up’ along lines of constant y a few changes in this numbering scheme may be required to minimise the band width. Paulson has written a version of `rmesh` which swops pairs of nodes to achieve this.

The boundary of a cross section of the human body is not circular (see Figure 4.3). However it is relatively smooth and has an approximate axis of symmetry (left-right). In the Sobolev theory of derivatives, smoothness is reflected in rapidity of decay of the Fourier transform. We can expect therefore that a Fourier series for the surface shape of a (star-shaped¹) body cross-section, represented as $r(\theta) = \sum_{k=0}^N c_k e^{ik\theta}$, to be economical. That is,

¹*Star-shaped* is used here in the mathematical sense of there being one point which can be connected to all others by a line lying within the body. We will leave the impedance location of truly star shaped bodies to fish!

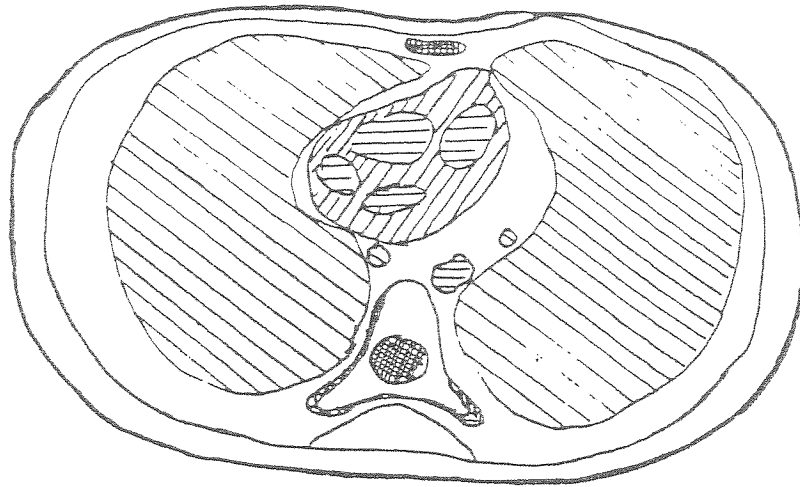


Figure 4.3: *A cross section of the human chest*

the error in our approximation will decrease rapidly with increasing N . For such shapes we can easily modify the circular mesh generation program to provide reasonably regular meshes provided their deviation from a circle is not too extreme. The author has produced a version of `rmesh` which replaces the concentric circles with curves whose radial coordinate is given by a Fourier series.

It is useful to have the capability to refine an existing finite element mesh for two reasons. One is to verify the experimental stability of the approximation scheme as $h \rightarrow 0$. This applies both to forward modelling and to inverse problem solution. The other reason is that it is desirable to represent the conductivity on a coarser mesh than the potential, in such a way that the finer mesh is a regular refinement of the coarser. The mesh for the conductivity should have no more nodes than there are independent degrees of freedom in the voltage measurements. On the other hand, the mesh for the potential needs to represent the potential accurately near the boundary where the field strength is high. The requirement that the mesh for the potential is a refinement of that for the conductivity facilitates the calculation of the conductivity at the nodes of the potential mesh, as required by `fwprob`.

A mesh refinement program was written which divides each triangular element into four smaller triangles, all similar to the original. This ensures that the radii of the elements decrease uniformly with increasing number of elements. This is the most obvious way to refine a mesh and its disadvantage is that it multiplies the number of elements by four. Therefore the number of elements increases exponentially. An algorithm with a linear increase would be useful.

The mesh refinement code `refine` is somewhat more complicated than one

might expect. The algorithm is as follows:

```

input old mesh
foreach element do
    calculate global coordinates of new nodes
    add new nodes to list of nodes
    add new elements to element topology array
    if newboundary nodes are created then
        add them to list of boundary nodes
enddo
sort nodes lexicographically in (x,y)
remove redundant nodes from list of nodes
remove redundant nodes in list of boundary nodes
output newmesh

```

Of these steps the addition of boundary nodes was most difficult requiring a case-by-case analysis of which original nodes were on the boundary. The new boundary nodes will have slightly incorrect coordinates and can be moved radially to fit the boundary.

4.7 Numerical Calculation of the Derivative Matrix

In subsequent chapters the derivative of the forward mapping DF_γ will be required. This matrix, sometimes called the ‘sensitivity matrix’ has been calculated by a number of other workers in the field. Kim [52] and Tarassenko [89] use a ‘perturbation technique’, that is they use a finite difference approximation to the derivative

$$DF_\gamma(\delta\gamma) \approx \frac{F(\gamma + \delta\gamma) - F(\gamma)}{\|\delta\gamma\|}. \quad (4.10)$$

These matrices were used in various ways to give approximate linear reconstruction algorithms. The disadvantage of such a perturbation technique is that a different finite element system must be solved to calculate $F(\gamma + \delta\gamma)$ for each conductivity change $\delta\gamma$. This requires inversion of a different system matrix each time. The cost of this technique would be prohibitive, except in the case where the derivative matrix is precalculated for a given initial guess γ for the conductivity and only one linear step is required. The integral formula for the derivative 3.45 gives a more efficient method of calculating DF_γ numerically. Only one system matrix need be assembled and its Choleski

factors calculated. All the potentials required in the derivative can then be calculated. Furthermore, the integrals of inner products of potential gradients required in 3.45, can be calculated readily from the stored coefficients $S(e, i, j, k)$ of Equation 4.9.

Yorkey [96] [97] calculated the derivative by a similar method. He also used another method in which the finite element model was replaced by a discrete resistor network. He then derived a discrete version of the integral form of the derivative using something he calls the ‘compensation theorem’. The idea is to calculate what change is required in an interior resistor to compensate for a change in voltage at the boundary. This is, in fact, a discrete version of the interpretation of $L_{\delta\gamma}u$ as a source (see Section 3.5). Yorkey reports that superior results were obtained in reconstructions using the compensation method.

In the author’s own calculations of DF_γ using the finite element method, it has been found that there are many possible pitfalls for the unwary. The most difficult part was ensuring that the boundary conditions for the measurement fields were correct. It was easy to calculate the derivative of a forward problem with a slightly different set of measurements from those intended. To detect human error, two methods were used for verifying the subroutine (called `calder`²) which calculates the derivative matrix. The first was to perform the lengthy calculation of the derivative by the perturbation method Equation 4.10 and verify that the result converges to the matrix calculated by the integral formula. This served to verify that the routine `calder` had indeed calculated the derivative of the particular forward mapping used in the program. Another much quicker test exploits the fact that

$$DF_1(1) = \int_{\Omega} \nabla w_k \cdot \nabla u_i = \int_{\partial\Omega} u \omega_k \quad (4.11)$$

(in the notation of Section 3.6) which provides a check on the row sums of the derivative matrix.

A possible reason for Yorkey’s finding is that it may have been easier for him to get the details of the boundary conditions correct using the discrete formulation — the two forms should be equivalent.

²A pun of CALculate DERivative and CALDERon.

Chapter 5

Ill-posedness of the Inverse Problem

*A monk asked Joshu, a Chinese Zen master:
'Does a dog have the Buddha-Nature or not?'
Joshu answered 'Mu.'¹*

Traditional Zen Koan.

5.1 Ill-posedness

Given a general map $F : X \rightarrow Y$, for Banach spaces X and Y consider the problem of solving the equation $F(x) = y$ when y is known and we seek x . Hadamard [41] defines such a problem to be *well-posed* if it satisfies the following conditions:

1. A solution exists.
2. The solution is unique.
3. The inverse mapping is continuous.

A problem is *ill-posed* if any of the above conditions is violated. Solving $F(x) = y$ for x is often called an *inverse problem* although this term is rather ill-defined itself. It is usually applied when the problem of evaluating F , that

¹Joshu's answer, meaning neither 'yes' nor 'no' indicates that the question was inappropriate.

is the *forward problem* is a classical or at least well known problem, for which it is known that F exists and is continuous. The term inverse problem is also usually reserved for those problems which are ill-posed in the sense of Hadamard.

As an example, solving Poisson's equation would not generally be considered an inverse problem, even though it does involve inverting the Laplacian. On the other hand, determining γ in Equation 2.3 from a knowledge of the transfer impedance operator R_γ , it would generally be agreed, is an inverse problem. In this thesis the term 'the forward problem' generally refers to the problem of evaluating the mapping $F(\gamma) = R_\gamma$. Sometimes, more precisely, it will refer to a semi-discrete problem $F(\gamma) = (R_\gamma j_k)$ for some finite set of currents j_k or to the discrete problem $F(\gamma) = (\langle \chi_l, j_k \rangle)$ where the χ_l are a set of measurement functions — such as the characteristic functions of electrodes. In the discrete problem the data are simply a matrix representation of the transfer impedance operator, that is a *transfer impedance matrix*.

5.2 Existence of an Inverse

Following the Hadamard conditions, the first question to ask is whether F^{-1} exists, where F here is the continuous mapping $F(\gamma) = R_\gamma$. For a mapping to have an inverse it must be injective and surjective. If F were surjective all linear operators in $L(H^{-1/2}, H^{1/2})$ would be possible transfer impedance operators for some $\gamma \in \mathcal{C}$. This is plainly not the case. Firstly, to qualify as a bona fide transfer impedance operator, $R \in L(H^{-1/2}, H^{1/2})$ must be invertible. Secondly the restriction of R to H^0 must be self adjoint and positive definite. Thirdly certain restrictions on the relationship between the maxima and minima of j and $R_\gamma j$ must apply (see [12]). However, we may simply restrict our attention to the range $R(\mathcal{C})$, that is assume that we already have a genuine transfer impedance operator for at least some γ . There is an obvious practical problem here that if we have measured the transfer impedance operator it will have some error and therefore no longer be in the range. Ignoring this difficulty for the moment, let us turn to the existence of F^{-1} on the range : $F^{-1} : F(\mathcal{C}) \rightarrow \mathcal{C}$. Restated, this is the injectivity of F : Do two conductivities produce the same transfer impedance operator?

This question was unsolved when the author began working on the project in autumn of 1985, but results were soon published in the literature. Kohn and Vogelius [55] answered the question in the affirmative for piece-wise analytic conductivities. Later Sylvester and Uhlmann [87] proved the same result for smooth conductivities (in fact their proof holds in a certain weighted Sobolev space). The technique used by Sylvester and Uhlmann hinges on proving

that there are solutions of Equation 2.1 which have similar properties to the Calderón fields described below but for a more general conductivity. These fields have the right asymptotic properties for the inverse scattering problem discussed in Section 3.7 to be solved.

A more recent result of Ramm [77] proves the result for $\gamma \in H^3(\Omega)$. The proof given here is more satisfying by virtue of its simplicity. The result hinges on proving that the set $\{\nabla u \cdot \nabla v\}$, where u and v range over all the $H^2(\Omega)$ solutions of $L_\gamma u = 0$, is complete in $H^0(\Omega)$. This result also depends on the Sylvester and Uhlmann fields. These special fields only exist in dimension greater than or equal to three, so they do not apply to the two dimensional case.

The discrete, finite data version of the problem may or may not have a unique solution depending on how the discretisation and data set are chosen. The existence of a solution may also be in question if the data have some error, or the representation chosen for the conductivity is not sufficiently general. To some extent these difficulties can be addressed by choosing the best fitting solution satisfying some additional criterion such as a least norm condition.

5.3 Continuity of the Inverse

In a discrete problem the most interesting aspect of ill-posedness is the violation of the third Hadamard condition that the inverse is not continuous.

In the case of the EIT reconstruction problem with complete data it is known that the inverse exists. The condition that the inverse is not continuous, can be expressed for the EIT inverse problem thus: there is a bound $M > 0$ such that given any accuracy $\epsilon > 0$ there is a conductivity perturbation δ with $|\delta| > M$ such that $\|F(\gamma + \delta) - F(\gamma)\| < \epsilon$. That is, there is an arbitrarily large δ such that whatever the accuracy of our measurements we cannot distinguish between γ and $\gamma + \delta$ by boundary measurement.

Whereas the existence of F^{-1} has received much attention in the literature, very little has been said about its lack of continuity, although Alessandrini [2] proves that it has some weaker form of continuity. In a special case, it is simple but instructive to construct an example to show that F^{-1} is not continuous.

Theorem 5.1 *For the case of $\Omega = D$, the unit disk, the forward mapping $F : \mathcal{C} \subset H^0 \rightarrow L(H^{-1/2}, H^{1/2})$ given by $F(\gamma) = R_\gamma$ has no continuous inverse.*

Proof

To prove this theorem it is sufficient to find a sequence $\gamma_k \in \mathcal{C} \subset H^0$ such that $\|\gamma_k - 1\| \geq \text{const} > 0$ and yet $R_{\gamma_k} - R_1 \rightarrow 0$. The construction of such a sequence is based on the observation that the distance of an object from ∂D has a greater effect on its detectability using boundary measurements, than does its conductivity contrast. Let

$$\gamma_k(r, \theta) = \begin{cases} 1, & 1/k \leq r \leq 1 \\ k+1, & 0 \leq r < 1/k \end{cases} \quad (5.1)$$

then clearly $\|\gamma_k - 1\|_0 = \sqrt{\pi}$. The norm $\|R_{\gamma}\|_{L(H^{-1/2}, H^{1/2})}^2$ is the maximum eigenvalue of $R_{\gamma}^* R_{\gamma}$. In this case, the spectrum of R_{γ} is known (see Section 2.4.2) and

$$\|R_{\gamma_k} - R_1\| = \frac{1 - \mu_k k^{-2}}{1 + \mu_k k^{-2}} - 1 \quad (5.2)$$

where $\mu_k = k/(k+2) \rightarrow 1$ as $k \rightarrow \infty$. Now

$$\lim_{k \rightarrow \infty} \mu_k k^{-2} = 0 \quad (5.3)$$

so we have

$$\lim_{k \rightarrow \infty} \|R_{\gamma_k} - R_1\| = 0 \quad (5.4)$$

as desired. ■

The same theorem works for smoother conductivities. To prove the result in H^s for $s > 0$ it is necessary to construct a sequence of smooth ‘bump’ functions with similar properties to γ_k . Rather than calculating the spectrum for these smooth conductivities it is sufficient to ‘sandwich’ them between the piece-wise constant functions used above and use the monotonicity property of F (see Figure 5.1). Let $\tilde{\gamma}_k$ be a smooth function with $4\gamma_{k/2} \geq \tilde{\gamma}_k \geq \gamma_k$ then by monotonicity $R_{4\gamma_{k/2}} \geq R_{\tilde{\gamma}_k} \geq R_{\gamma_k}$. Now

$$\|\tilde{\gamma}_k - 1\|_s > \|\gamma_k - 1\|_s > \|\gamma_k - 1\| = \sqrt{\pi} \quad (5.5)$$

and

$$\lim_{k \rightarrow \infty} R_{\tilde{\gamma}_k} - R_1 = \lim_{k \rightarrow \infty} R_{\gamma_k} - R_1 = 0. \quad (5.6)$$

So the following theorem has been proved:

Theorem 5.2 *For the case of $\Omega = D$, the unit disk, the forward mapping $F : \mathcal{C} \subset H^s \rightarrow L(H^{-1/2}, H^{1/2})$ given by $F(\gamma) = R_{\gamma}$ has no continuous inverse.*

■

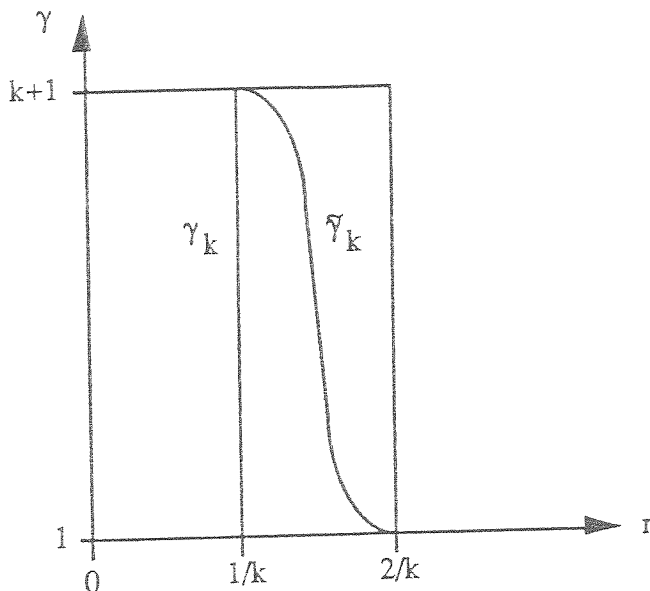


Figure 5.1: The smooth function $\tilde{\gamma}_k$ sandwiched between piecewise constant conductivities.

This theorem captures something of the essence of why the EIT reconstruction problem is ill-posed. *Small objects in the centre are very hard to detect however big their conductivity contrast.* Another way to think of this is that the Greens operator G^B which solves the Neumann problem for Equation 2.1 is a smoothing operator. It is hard to get high frequency components in the current pattern to propagate into the centre. This means that information about conductivity gradients in the centre is difficult to obtain.

5.4 The Linearised Problem

An affine approximation to the inverse problem is to solve

$$F(\gamma + \delta\gamma) = F(\gamma) + DF_\gamma(\alpha), \quad (5.7)$$

for α . Then α is an approximation to $\delta\gamma$. This problem will be referred to loosely as ‘the linearised problem’. More specific information can be obtained about the conditioning of the linearised problem than the original non-linear problem. However the relationship between the ill-posedness of the linear and non-linear problems is not straightforward. If the affine problem were well-posed for some γ and the range of the derivative DF_γ were closed, the inverse function theorem would guarantee that F was invertible in a neighbourhood of $F(\gamma)$ and the inverse is differentiable (and hence continuous). The fact that

F^{-1} is not continuous indicates that the range of F is not closed or that DF_γ is not invertible for any γ . On the other hand, knowledge that the continuous affine problem is ill-posed tells us little about the conditioning of the non-linear problem.

Theorem 5.2 gives an example which shows that F^{-1} is discontinuous at some point. This leads one to suspect that this is generally the case. It would be satisfying to prove that F was a compact mapping, however no proof of this is known to the author. However it is proved in Theorem 5.3 that DF_1 is compact.

If the conductivity space is discretised as well as the measurements, the derivative is a matrix. If the matrix is invertible then the inverse function theorem guarantees that the discrete F will also be invertible. However the invertibility of a matrix with accurate data and exact arithmetic does not imply that the linear system can be inverted for real data using a fixed precision computer. One must look at measures of the conditioning of the linear system.

5.5 Singular Value Decomposition

In this section the linearised problem is studied in detail. The key question is: *To what extent is the derivative of the forward mapping invertible?* Equivalently one could ask *How ill-posed is the linearised inverse problem?* The main tool used to pursue these questions is the Singular Value Decomposition (SVD)[35]. Let us briefly review this technique.

If X and Y are separable Hilbert spaces and $A : X \rightarrow Y$ is a bounded linear map then a *Singular Value Decomposition* of A is a set $\{\phi_k \in X, \psi_k \in Y, \lambda_k \in \mathbb{R} | k \in \mathbb{N}\}$ where the $\{\phi_k\}$ forms an orthogonal basis for X and $\{\psi_k\}$ forms an orthogonal basis for Y satisfying

$$\begin{aligned} A\phi_k &= \lambda_k\psi_k \\ A^*\psi_k &= \lambda_k\phi_k. \end{aligned}$$

The ϕ_k and ψ_k are called *right* and *left singular functions* and the λ_k *singular values*. In the case where X and Y are finite dimensional the singular functions may be called *singular vectors*. This extends the concept of an orthogonal eigensystem from symmetric operators to general operators between completely different spaces. For example it does not make sense to talk of the eigenfunctions of a linear map between a space of conductivities and a space of voltages since the two cannot be compared. The existence of an SVD can be deduced by applying standard spectral theory to the map $\tilde{A} : X \oplus Y \rightarrow X \oplus Y$

defined by

$$\tilde{A}(x, y) = (A^*y, Ax).$$

In particular, a compact operator has an SVD. Again a particular case of this situation is when X and Y are finite dimensional. An SVD, when it exists, is unique up to ordering of the singular functions and their singular values, and the sign of the singular values. By convention the singular values are chosen to be non-negative and the singular functions are ordered so that the corresponding singular values are decreasing.

Let us consider the finite dimensional case where A can be represented by a matrix. The SVD gives immediate access to some important numerical parameters of A . In particular, the norms $\|A\| = \max_k \lambda_k^2$, $\|A^{-1}\| = 1/\min_k \lambda_k^2$ and the condition number $\kappa(A) = \|A\| \|A^{-1}\|$. For system of linear equations $Ax = y$ the condition number measures the accuracy to which the system may be solved. The value of $\kappa(A)$ is always at least one and a large condition number indicates that the equations will be difficult to solve accurately. If A is well scaled one can expect to be able to calculate x to within an accuracy of $k - |\log_b \kappa(A)|$ base- b digits on a machine with k -digit base- b arithmetic (See Dongarra [26]). Another more geometric characterisation of the condition number is that it measures the relative distance to the set of singular matrices [35]:

$$\kappa(A) = \inf_{A+E \text{ singular}} \frac{\|E\|}{\|A\|}. \quad (5.8)$$

It is important to emphasise that the condition number, like the SVD, is dependent on the norm used.

In general the SVD provides a basis in which the operator A is diagonalised. That is, introduce the (possibly infinite) matrix representation of A in the basis ϕ_k for X and ψ_k for Y is

$$\Lambda = [\delta_{ij} \lambda_i].$$

One may compare this situation with differential operators, which can be diagonalised using the Fourier transform (FT). Indeed the SVD provides us with an extension of the FT, the singular function transform SFT. If $x \in X$ and $y \in Y$ the (left and right) SFTs are given by $\langle x, \phi_k \rangle_X$ and $\langle y, \psi_k \rangle_Y$.

In linear signal processing one deals, in the main, with systems which are governed by pseudo-differential operators, that is, systems whose transfer function can be diagonalised using Fourier transforms. The SFTs are the Fourier transform and inverse Fourier transform. In this context, a *band-limited* signal is a function with a compactly supported FT. A *filter* is a pseudo-differential operator and its frequency response the ordered singular values. The *impulse response* of a system is the response to a delta distribution input. In image processing similar techniques are used. The time varying signal is

replaced by an image considered as a function of two spatial variables. Images can be band-limited and one can apply spatial filters (the Laplacian operator for example acts as an edge enhancer). The response of an imaging system to a point image, that is a delta distribution, is called the *point spread function* (see Section 3.5 for an example of the point spread function for an EIT system).

In a more general linear system these Fourier transform techniques can be extended using the singular value decomposition. Given a function $s : R^+ \rightarrow R$ one can filter the transfer function A by defining

$$s(A) : X \rightarrow Y, \quad s(A)f = \sum_k s(\lambda_k) \langle f, \phi_k \rangle_X \psi_k. \quad (5.9)$$

A linear inverse problem, such as the reconstruction problem for a tomographic imaging system, involves the identification of $x \in X$ from a knowledge of $y \in Y$ using the relation $Ax = y$. Typically our knowledge of y is not complete. It may be known that the error e in y satisfies $\|e\|_Y < \epsilon$. This gives a corresponding uncertainty in x which can be understood in terms of the SVD. Let M be the smallest integer such that $\lambda_{M+1} < \epsilon$. Define the SFT of x as

$$\alpha_k = \langle \phi_k, x \rangle_X.$$

then we have no information about $\alpha_{M+1}, \alpha_{M+2}, \dots$. Put another way we can reliably deduce the component of x lying in the span of the first M left singular functions. The SFT of $y, \beta_k = \langle \psi_k, y \rangle_Y$ contains useful information but the components β_k for $k > M$ contain none. If we decide to increase the accuracy of our system by making more linearly independent measurements, this will only add useful information if the new measurements have non-trivial components in the span of ψ_1, \dots, ψ_M . The most economical set of measurements to collect would be β_1, \dots, β_M . Later in this chapter an example is given of this in the context of data collection in EIT.

The decreasing sequence $\{\lambda_k\}$ measures the maximum amount of information which can be gained by adding an additional measurement. The function $M(\epsilon) = \text{maximum } k \text{ with } \lambda_k > \epsilon$ measures the number of identifiable degrees of freedom in x for a given measurement error in y . The SVD therefore summarises the information loss in the transfer function, or the *ill-posedness* of the inverse problem. The following two cautions must be observed in its interpretation. The SVD depends on the norms (and hence the inner products) used on the spaces X and Y . Different norms will result in differently conditioned problems. For example adding higher order derivatives or differences in the norms (that is going to a higher order Sobolev space) will typically result in a less ill-posed problem. This is the result of applying an *a priori* smoothness condition on the solution. The second caution is that it is not sufficient to look at the decay of the singular values to determine the ill-posedness of the

problem. One must also ascertain whether the question asked (such as: *Has this brain got a haemorrhage?*) can be answered when projected onto the first M singular functions.

5.6 Calderón Fields

One example where it is possible to calculate explicitly the SVD of the linearised Forward problem for EIT is afforded by using certain fields devised by Calderón. To construct these fields take a vector $\omega \in R^n$. Let $\omega^\perp \in R^n$ be any vector orthogonal to ω with $|\omega| = |\omega^\perp|$. The vectors $\zeta_+ = \pi(i\omega + \omega^\perp)$ and $\zeta_- = \pi(i\omega - \omega^\perp) \in C^n$ now have the properties that $\zeta_\pm^2 = 0$, $\zeta_+ + \zeta_- = 2\pi i\omega$ and $\zeta_+ \cdot \zeta_- = -2\pi^2|\omega|^2$. The Calderón potentials are now defined as

$$u_\pm(x) = e^{\zeta_\pm \cdot x}.$$

These are complex solutions of Laplace's equation:

$$\Delta u_\pm = 0.$$

They also have the particularly useful property that

$$\nabla u_+ \cdot \nabla u_- = -2\pi^2|\omega|^2 e^{2\pi i\omega \cdot x}.$$

Now consider the particular EIT forward problem with

$$\gamma \nabla_{\mathbf{n}} u = \gamma \nabla_{\mathbf{n}} u_+$$

and

$$\gamma \nabla_{\mathbf{n}} v = \gamma \nabla_{\mathbf{n}} u_-.$$

One small point which deserves clarification at this point is that the potentials are now complex. How can one apply complex boundary conditions? The simple answer is to apply real and imaginary parts separately and then combine the resulting measurements. The data measured are

$$d_\omega(\gamma) = \int_{\partial\Omega} u \gamma \nabla_{\mathbf{n}} v$$

for each ω . The derivative $D(d_\omega)$ at $\gamma = 1$ is given by

$$D(d_\omega)\delta\gamma = - \int_{\Omega} \delta\gamma \nabla u_+ \cdot \nabla u_- = 2\pi^2 \int_{\Omega} \delta\gamma |\omega|^2 e^{2\pi i\omega \cdot x} dx.$$

Extending $\delta\gamma$ by zero outside $\bar{\Omega}$ we can consider the above integral to be over all of R^n . This yields the Fourier transform

$$D(d_\omega)\delta\gamma = 2\pi^2 |\omega|^2 \widehat{\delta\gamma}.$$

This has diagonalised the linearised forward mapping and, provided care is taken to define which spaces and norms are used, constitutes the SVD.

A Hilbert space is required for an SVD. The space used by Calderón, $L^\infty(\Omega)$, will not do. A first choice might be $L^2(\Omega)$. This space is spanned by the Fourier basis $\{e^{2\pi i \omega \cdot x}\}$ and the Fourier transform is a unitary automorphism. We would be measuring the distance between pictures using the L^2 norm which is frequent practice in medical imaging although its use is questioned [71] and it is justified usually by convenience. The space of measurements is parameterised by $\omega \in R^n$. As Ω is bounded it is possible (possibly after a change of scale) to use the discrete Fourier transform and $\omega = \underline{k} \in Z^n$. This corresponds to extending $\delta\gamma$ periodically. The data consists of sequences parameterised by n -tuples of integers. If we take the $L^2(\Omega)$ norm on the conductivity and the $L^2(Z^n)$ norm on the data, the mapping is unbounded as it stands. Restricting the domain to $H_0^2(\Omega)$ but using the $L^2(\Omega)$ norm results in a densely defined mapping. The SVD of this mapping is:

$$\begin{aligned}\phi_{\underline{k}}(x) &= e^{2\pi i \underline{k} \cdot x} \\ \psi_{\underline{k}, \underline{m}} &= \delta_{\underline{k}, \underline{m}} \\ \lambda_{\underline{k}} &= 2\pi |\underline{k}|^2.\end{aligned}$$

The data space is somewhat artificial as it is would not be possible to collect data in this form. Consideration of this point will be deferred while the conditioning of this map is investigated.

Taking a finite subset of the data, $\delta D(d_{\underline{k}})\delta\gamma$, for $|\underline{k}| \leq N$ yields a matrix with condition number N^2 and the conditioning of the problem worsens with increasing number of measurements. However in this case the situation is not too bad. The higher frequency components become easier to find.

A different formulation would be to take the conductivity space as $H_0^2(\Omega)$. Recall that this space is the closure in $H^2(R^n)$ of the set of smooth functions supported on Ω , and in particular these distributions vanish on $\partial\Omega$. This amounts to assuming that the conductivity has not changed on the boundary. The change in mean conductivity is also invisible to this mapping as $\widehat{\delta\gamma}(0)$ is in the kernel. Taking the conductivity space now to be $H_0^2(\Omega)$ modulo constants with the norm $\|f\|^2 = \sum_{k \in Z^n} |k|^2 \hat{f}(k)$ (which is equivalent to the standard norm) the set $\{|k|^2 e^{2\pi i k \cdot x}\}$ forms an orthogonal basis and a set of left singular functions where the singular values are all now 1. The linearised forward mapping is now an isometry between these spaces and is thus well conditioned with respect to inversion. The derivative $D(d_{\underline{k}}) : H_0^2/\text{constants} \rightarrow L^2(Z^n)$ is now an injective map with closed range and the inverse function theorem guarantees that $D(d_{\underline{k}})$ is invertible in a neighbourhood of $\gamma = \text{constant}$. This is a weaker result than the uniqueness theorems of Sylvester and Uhlmann [87] but it is worth noting for its simplicity and intuitive appeal.

For the case of non-constant γ and domains of at least three dimensions, Sylvester and Uhlmann [87] give an analogue of the Calderón fields. For $\zeta \in C^n$ such that $\zeta \cdot \zeta = 0$ they show there are solutions of $L_\gamma u = 0$ given by

$$u(x) = (\gamma(x))^{-1/2} e^{i\zeta \cdot x} (1 + R(x, \zeta)) \quad (5.10)$$

where the function R satisfies

$$\|R(\cdot, \zeta)\|_{L^\infty(\Omega)} \rightarrow 0 \text{ as } |\zeta| \rightarrow \infty, \zeta \cdot \zeta = 0. \quad (5.11)$$

It can be seen that as the spatial frequency gets large, these tend to the Calderón fields. The result of Ramm [77] relies on the fact that for $\gamma \in H^3$, the gradient of R has the same asymptotic property

$$\|\nabla R(\cdot, \zeta)\|_{L^\infty(\Omega)} \rightarrow 0 \text{ as } |\zeta| \rightarrow \infty, \zeta \cdot \zeta = 0. \quad (5.12)$$

5.7 Fourier Series for Calderón Fields

It is interesting to calculate the Fourier Series for the Calderón fields at the boundary. As the linearised inverse problem is well posed with Calderón data it would seem that one could measure Fourier data and synthesise the Calderón data from it. In this section Ω is taken to be the unit disk $D \subset R^2$.

Let $\omega = (\alpha, \beta) \in R^2$, and consider a point $x = (\cos \theta, \sin \theta) \in \partial D$. We have

$$u_\pm = e^{\alpha(\pm \sin \theta + i \cos \theta) + \beta(\mp \cos \theta + i \sin \theta)} \quad (5.13)$$

$$= e^{(i\alpha \mp \beta)e^{\mp i\theta}} \quad (5.14)$$

$$= \sum_{k=0}^{\infty} \frac{(i\alpha \mp \beta)^k}{k!} e^{\mp ik\theta} \quad (5.15)$$

which is the Fourier series for u_\pm .

An important observation about Equation 5.15 is that to synthesise the Calderón data for any spatial frequency ω from Fourier data, one needs the Fourier data of all frequencies. Put simply, band-limited Fourier data can not be translated to perfectly band limited Calderon data and reconstructed to give a band limited conductivity. This is a fundamental limitation of impedance imaging. Limits to spatial resolution in the measurements (principally the finite number of electrodes) lead to a degradation of all spatial frequency components of the reconstructed conductivity.

Suppose that one can only apply band limited current patterns — that is we can only specify the first K frequency components up to a precision ϵ . We can then calculate ϕ_ω to the same accuracy provided $|\omega| \leq N$ where N is such

that $N^K/K! \leq \epsilon$. This indicates that the number of frequency components of $\delta\gamma$ which can be calculated accurately is $N(K) = \sqrt[K]{K!}$. The ratio of number of frequency components of $\delta\gamma$ which can be recovered from data limited to the first K frequency components is $N(K)/K$, which tends to zero as K goes to infinity.

This argument, although admittedly rather loose, shows the difficulty of synthesising Calderón data from Fourier data and gives some insight into the ill-posedness of the inverse problem. To increase the number of frequency components of the conductivity recovered to a certain accuracy one has not only to increase the number of frequency components of the data applied, but also to increase their accuracy.

5.8 Ill-posedness of Linearised Problem

In this section the ill-posedness of the linearised problem is proved. More precisely, the derivative is a compact mapping:

Theorem 5.3 ² *Let $F : \mathcal{C} \rightarrow L(H^0(\partial\Omega))$ be $F(\gamma) = R_\gamma$ and $A = DF_\gamma : \mathcal{C} \rightarrow L(H^0(\partial\Omega))$ then A is compact.*

Proof

Define $K_{k,\ell} \in H^0(\Omega)$ by

$$K_{k,\ell} = \gamma \nabla(G_\gamma^{NB} j_k) \cdot \nabla(G_\gamma^{NB} j_\ell) \quad (5.16)$$

where j_1, j_2, \dots are the eigenfunctions of R_γ with eigenvalues $\lambda_1 > \lambda_2 > \dots$. The range of A is contained in $L_c(H^0(\partial\Omega))$ the set of compact operators between $H^0(\partial\Omega)$ and itself. This space is naturally isomorphic to the tensor product $H^0(\partial\Omega) \otimes H^0(\partial\Omega)$ when the former is given the Frobenius norm ($\|R\|^2 = \text{trace}(R^*R)$). The inclusion $L_c(H^0(\partial\Omega)) \hookrightarrow L(H^0(\partial\Omega))$ is continuous so it is sufficient to prove $A : H^0(\Omega) \rightarrow H^0(\partial\Omega) \otimes H^0(\partial\Omega)$ is compact. The proof continues by showing that A is a Hilbert-Schmidt operator:

$$\sum_{k,\ell} \int_\omega K_{k,\ell}^2 \leq \|\gamma\|_0^2 \sum_{k,\ell} \|G_\gamma^{NB} j_k\|_1^2 \|G_\gamma^{NB} j_\ell\|_1^2 \quad (5.17)$$

but

$$G_\gamma^{NB} j_k = G^{DB} R_\gamma j_k = \lambda_k G^{DB} j_k \quad (5.18)$$

²This theorem is dedicated to the man sitting next to me, on an airliner above Arizona, on my birthday in 1988. Without the sheet of paper he gave me this theorem might not have been recorded as the margin on the in-flight magazine was too small.

Name	Number of Nodes	Number of Elements	Type of Elements
rmesh	73	112	3 node triangular
wlrmesh	305	544	3 node triangular

Table 5.1: *Finite Element meshes used for calculation of derivative matrix*

Letting $\|G^{DB}\| = C$ we have

$$\sum_{k,\ell} \int_{\omega} K_{k,\ell}^2 \leq \sum_{k,\ell} \lambda_k^2 \lambda_{\ell}^2 \|G^{DB} j_k\|_1^2 \cdot \|G^{DB} j_{\ell}\|_1^2 \quad (5.19)$$

$$\leq C^2 \|\gamma\|_0^2 \sum_{k,\ell} \lambda_k^2 \lambda_{\ell}^2 \leq C^2 \|\gamma\|_0^2 \left(\sum_k \lambda_k^2 \right) \left(\sum_{\ell} \lambda_{\ell}^2 \right) \quad (5.20)$$

which is finite, by the compactness of R_{γ} (see Section 2.6). \square

The compactness of the derivative shows that the linearised forward mapping is not continuously invertible. The theorem of Ramm ([77]) shows that the set $\{K_{k,\ell}\}$ is complete in H^0 and the derivative is injective, so it does have an inverse, but that inverse is not continuous.

The proof of Theorem 5.3 relies on the compactness of the transfer impedance operator. Recall that this is only true because we chose to measure currents using the 0-order Sobolev norm. Had we chosen the much less stringent $-1/2$ -th order norm the situation may have been different. This would allow us to apply very ‘spikey’ current patterns which would still be less smooth even after the smoothing treatment by G^{NB} . This would make the linearised problem better posed.

5.9 Numerical Calculation of SVD

5.9.1 Implementation

The derivative matrix was calculated using the finite element method described in Chapter 4. Two meshes were used (see Table 5.1), both discretisations of a two dimensional unit disk. They were created using the program round as described in Section 4.6. The singular value decomposition was calculated using the Fortran subroutine dsydc from the Linpack library [26]. It is important to use a purpose designed SVD routine rather than using a general purpose eigenvalue routine applied to $A^T A$ [34]. This avoids amplifying truncation errors in the calculation of the product $A^T A$. The time taken to compute the SVD

is $O(n^3)$ where n is the largest dimension of the matrix. This prohibited the use of significantly finer meshes with the computing resources available (the Sun workstations). Finer meshes could be tackled using parallel processing hardware [10] but this was not attempted.

5.9.2 Results and Interpretation

As anticipated, the decay of the singular values is rapid. For `wimesh` the logarithm of the singular values is given in Figure 5.2. These results were for sixteen point electrodes. The current patterns driven were trigonometric, the current at electrode l (for $0 \leq n \leq 15$) being given by

$$I_l = \begin{cases} \cos k(\pi/8)^l \\ \sin k(\pi/8)^l \end{cases} \quad (5.21)$$

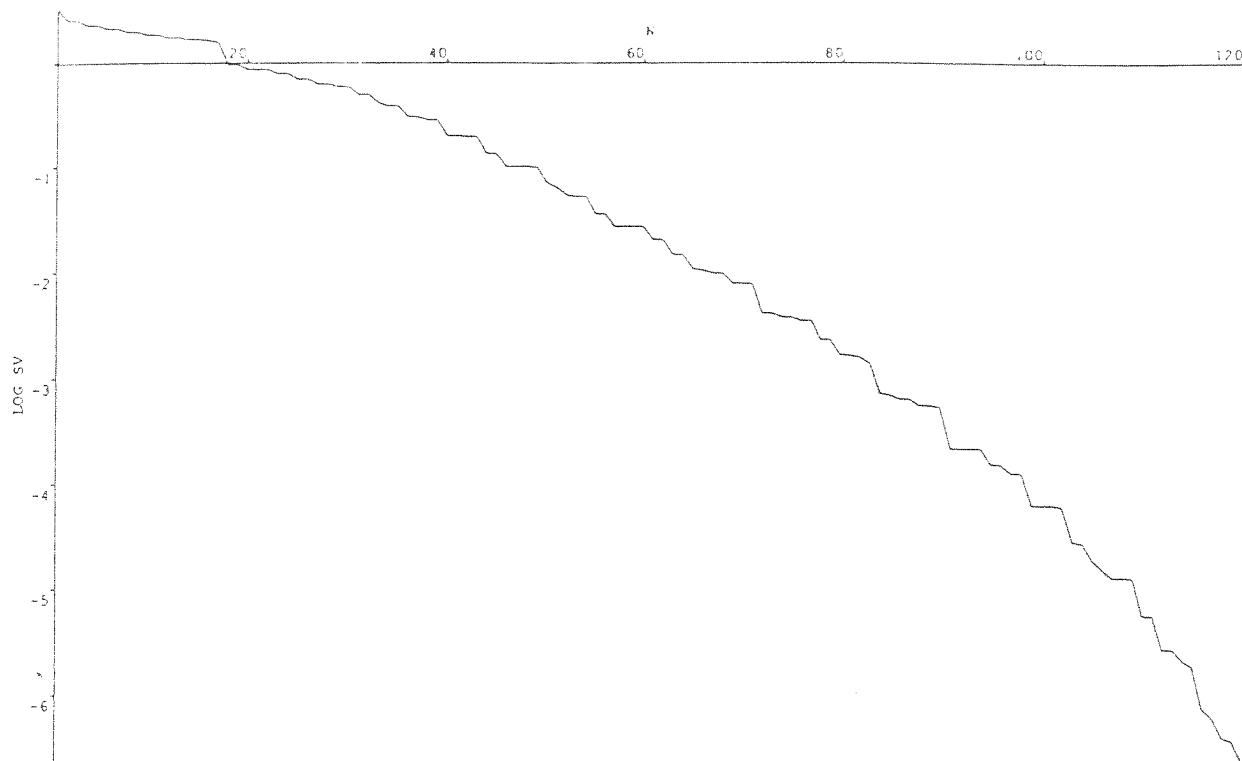
where k ranges from 1 to 16 for the sine terms and 1 to 15 for the cosine. The voltage measurement data set taken was designed to be symmetric with respect to rotations mapping electrodes to electrodes. Hence the reference level used was the average of the voltages over all the electrodes for that current pattern. If the voltage at electrode l for current pattern k was v_{lk} then the measurements used were

$$V_{lk} = v_{lk} - \sum_{i=0}^{16} v_{ik}. \quad (5.22)$$

This scheme gives the exact measurement fields the same sixteen-fold symmetry as the potentials. The derivative matrix, and hence the singular functions, will inherit this symmetry. Deviations from this are possible in the numerical results as the mesh itself will not have sixteen-fold symmetry. This can be used as a check on numerical results. The first 30 singular functions are shown in Figure 5.3. Some particularly interesting later examples are shown in Figure 5.4.

Computation times on a Sun workstation were approximately 15 minutes elapsed for `rmesh` and 2 hours for `wimesh`. Meshes or measurement schemes resulting in larger matrices were found either to cause the SVD routine not to converge or not to be feasible within the limits of memory available. The accuracy of the singular functions and values were checked by application of the matrix A to each singular function. The error $\|A\phi_k - \lambda_k\psi_k\|$ was less than 1×10^{-10} in all cases where the SVD was successfully computed.

On a logarithmic scale (Figure 5.2), the singular values decay faster than a linear function. This indicates decay of order $O(e^{-p(n)})$ for p a polynomial of degree two or more. This indicates extreme ill-posedness of the inverse problem. By contrast the inverse problem for X-Ray CT has singular values which



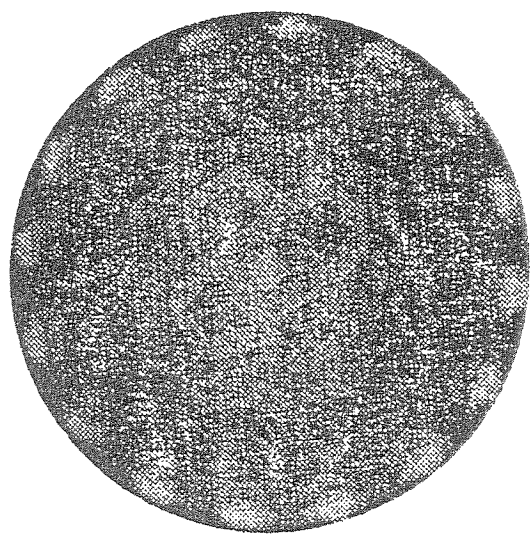
W1MESH, 16 point electrodes

Figure 5.2: *Logarithmic singular values* (\log_{10}), for w1mesh . .

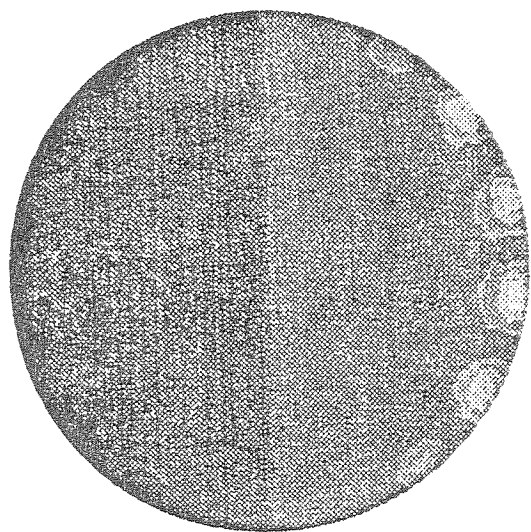
decay with an order $O(1/\sqrt{n})$. The other noticeable feature is the pattern of 'jumps' in the values. This phenomenon is not completely understood, and perhaps will not be until an analytic form for the singular value decomposition is found. However some qualitative understanding of the phenomenon can be obtained by comparing the jumps with patterns of the singular functions.

As a general principle, conductivity perturbations further from the boundary result in smaller voltage changes. This explains the observation that those singular functions whose maxima and minima are closer to the centre correspond to smaller singular values. The first singular function consists mainly of conductivity changes near the electrodes — clearly this results in the largest voltage changes. Singular functions 2-16 represent conductivity changes with increasing orders of rotational symmetry (see Figure 5.5). The possibility for orthogonal functions of this type is exhausted by singular function 17 which exhibits a dramatic change with structure in the centre. This corresponds to the first jump in the singular functions. The pattern continues with singular functions which have rotationally symmetric structures towards the centre, the order of symmetry increasing until the next dramatic change from singular value 39 to 40, when the order of symmetry again decreases in favour of the structure appearing still closer to the centre.

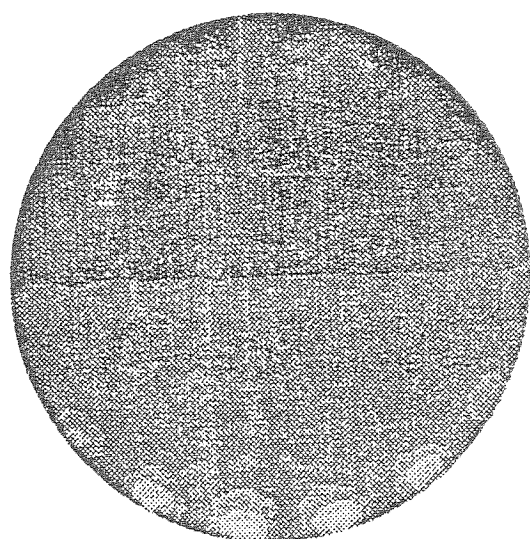
Figure 5.3: *The first 30 singular functions.*



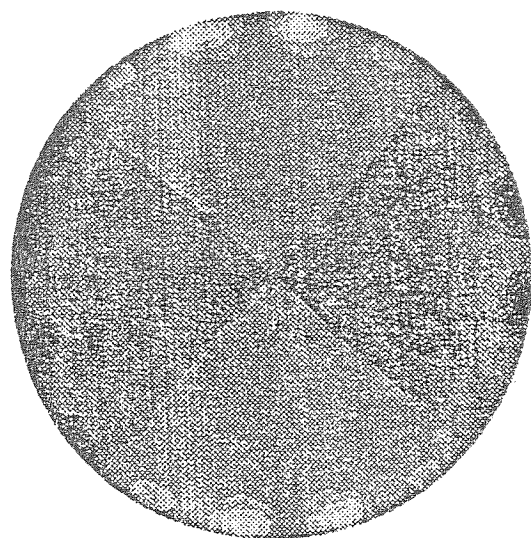
1



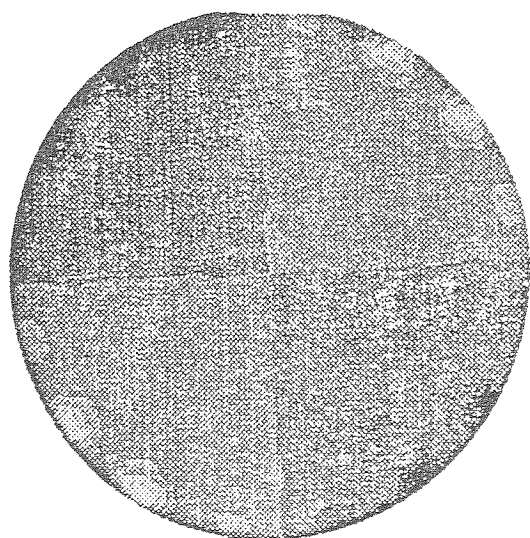
2



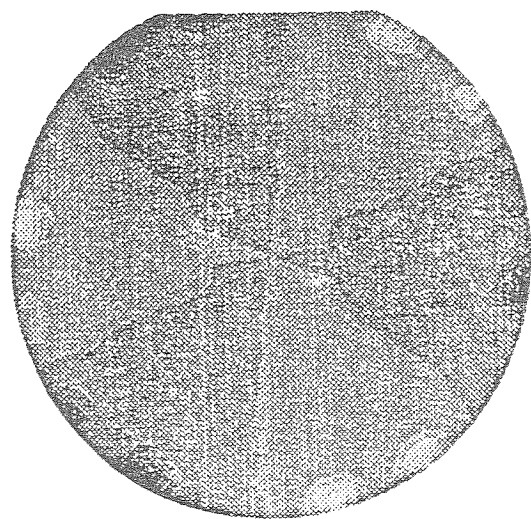
3



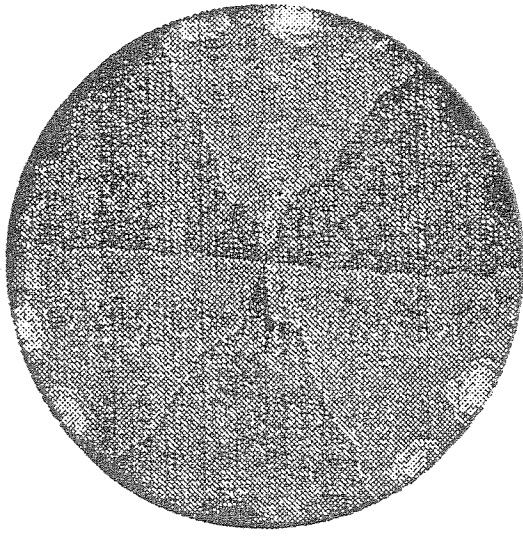
4



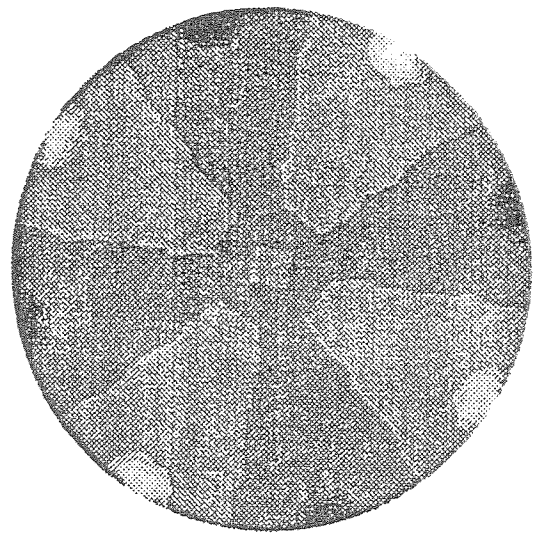
5



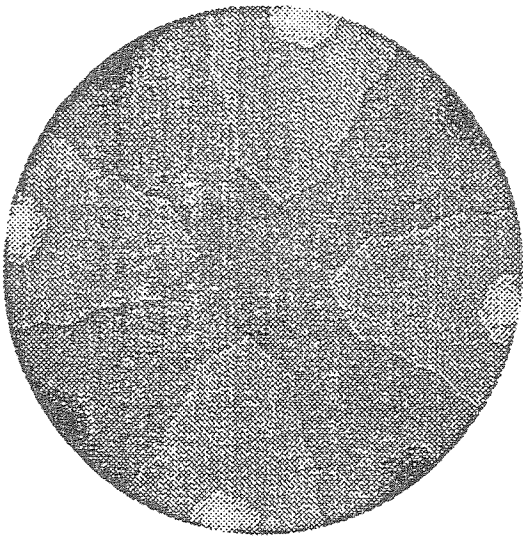
6



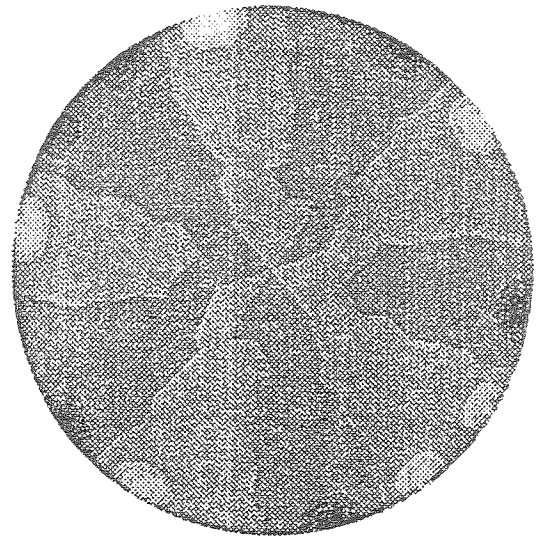
7



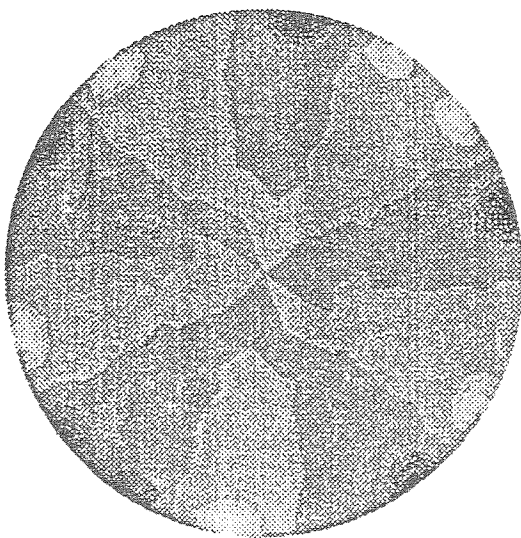
8



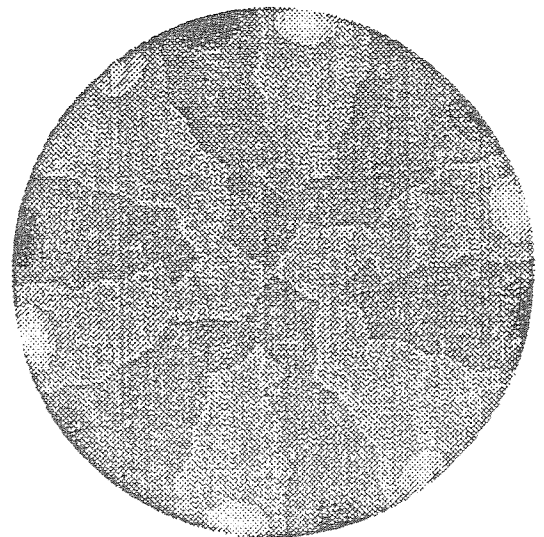
9



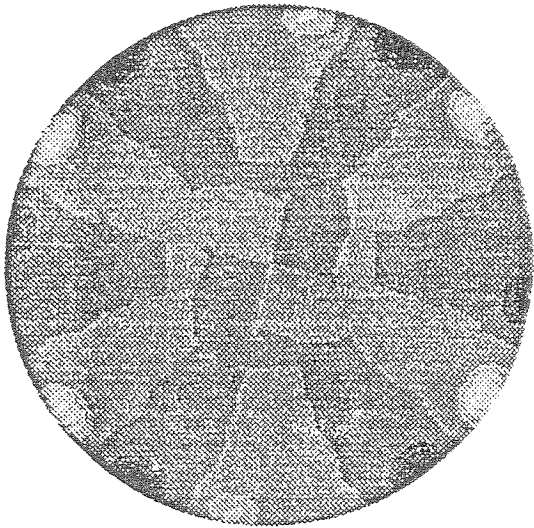
10



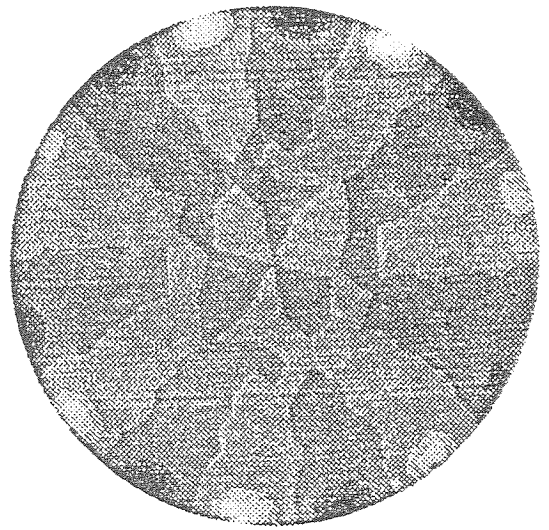
11



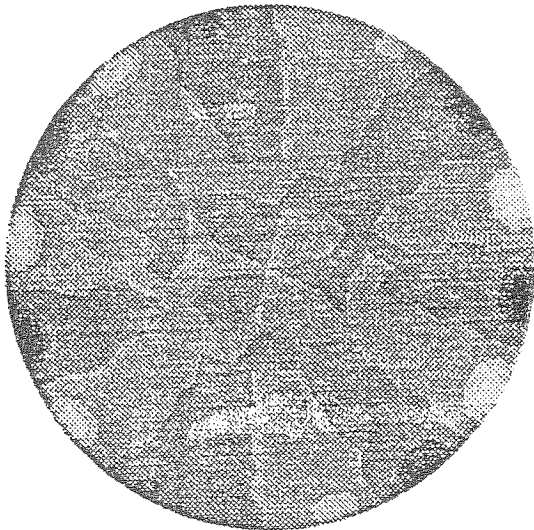
12



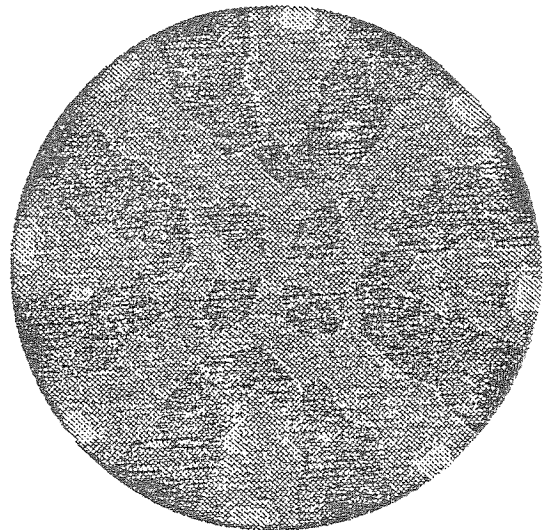
13



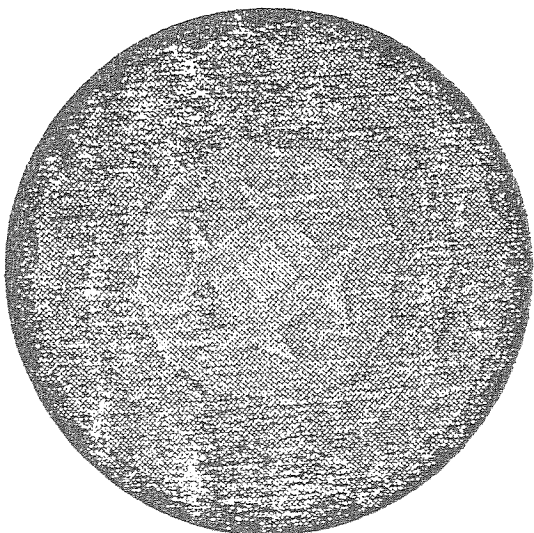
14



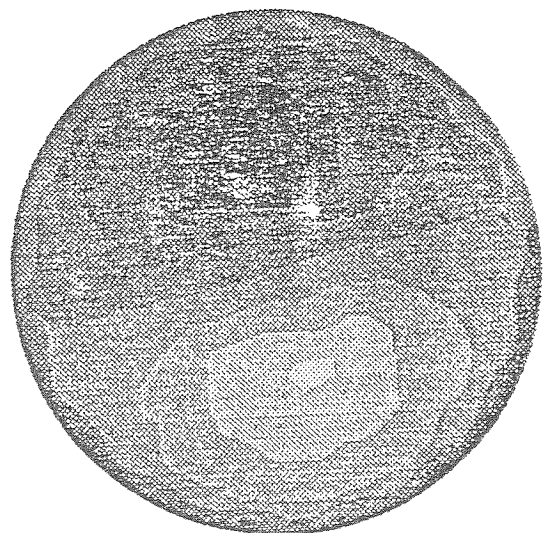
15



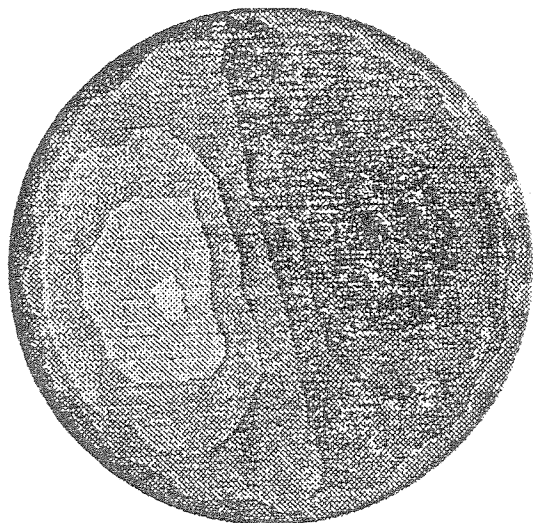
16



17



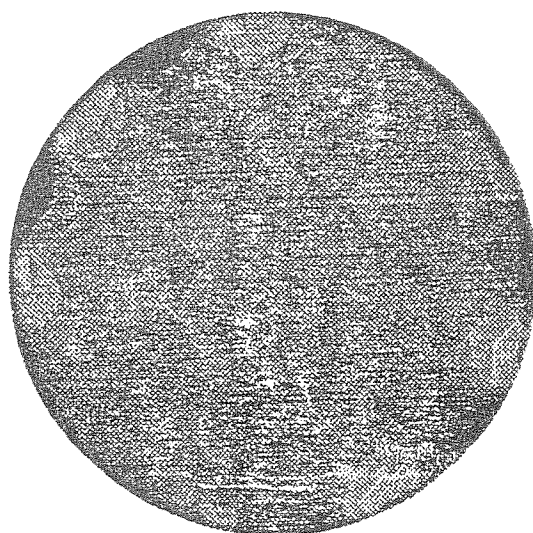
18



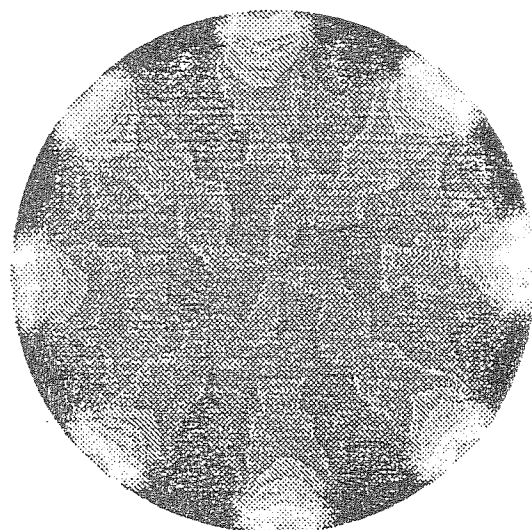
19



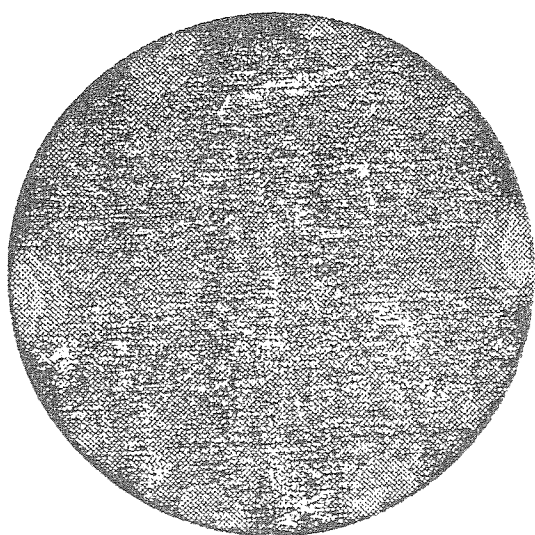
20



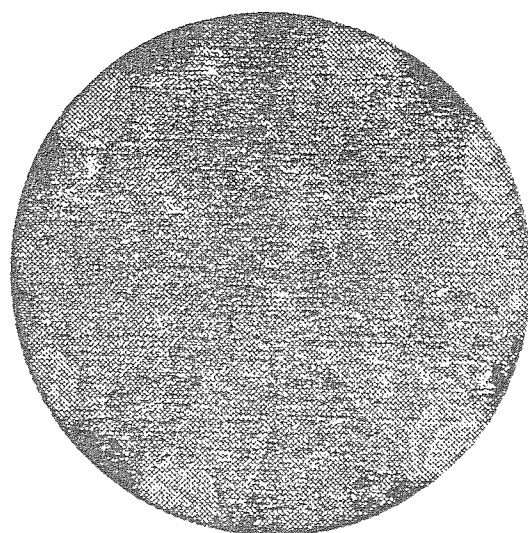
21



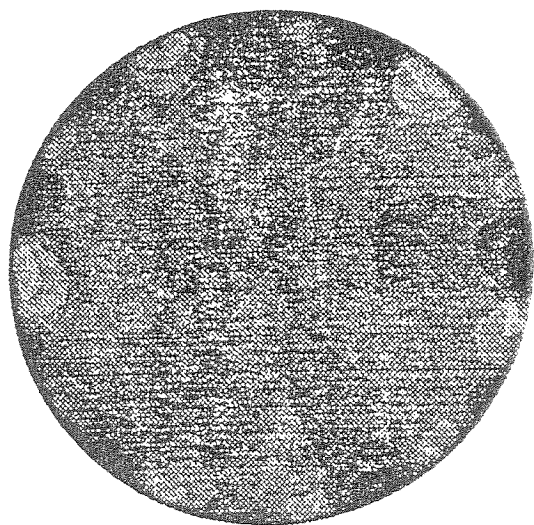
22



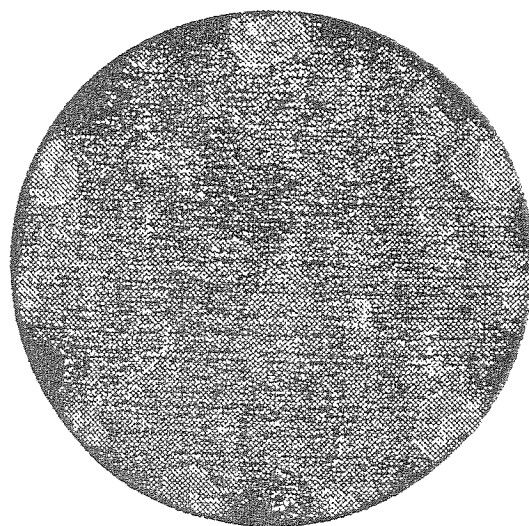
23



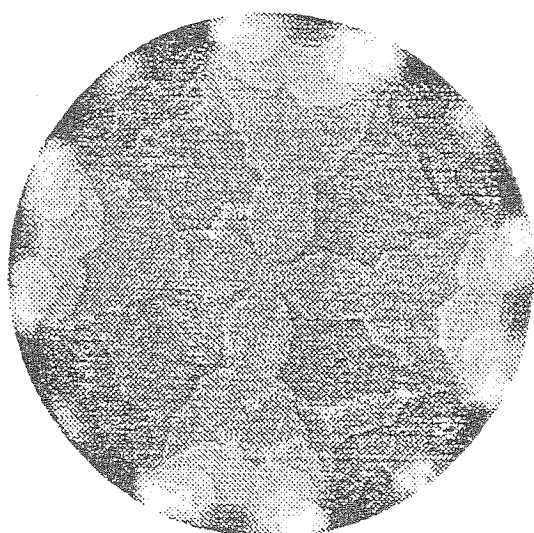
24



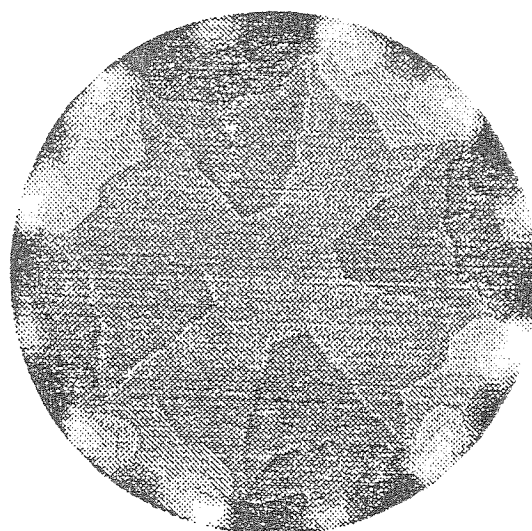
25



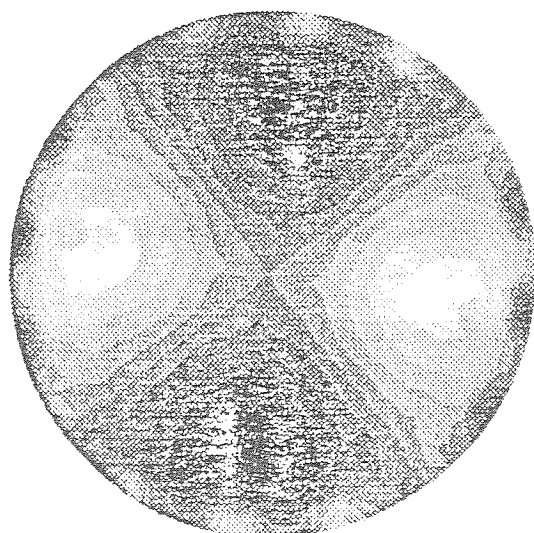
26



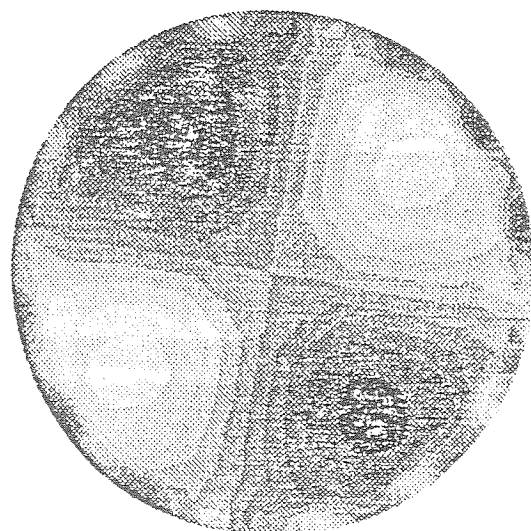
27



28

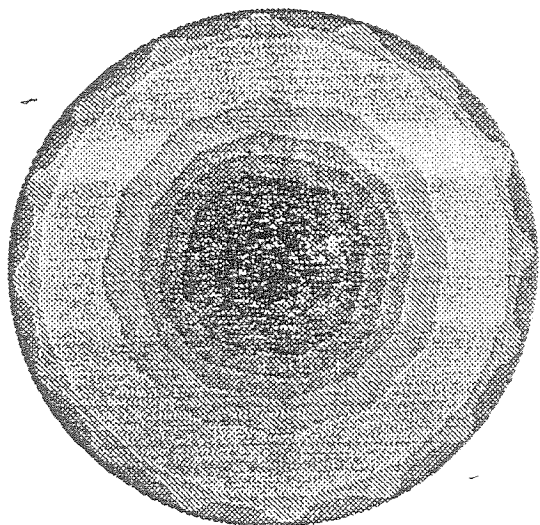


29

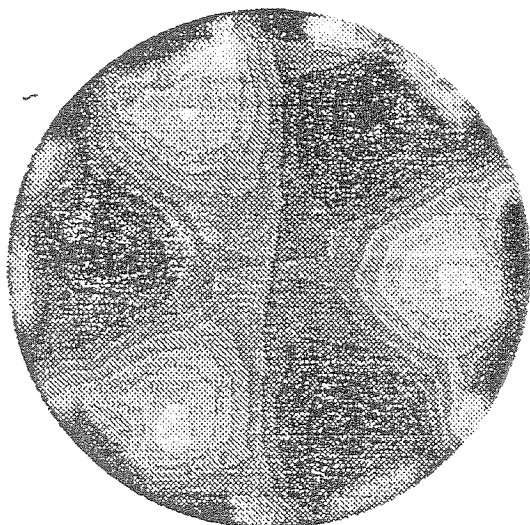


30

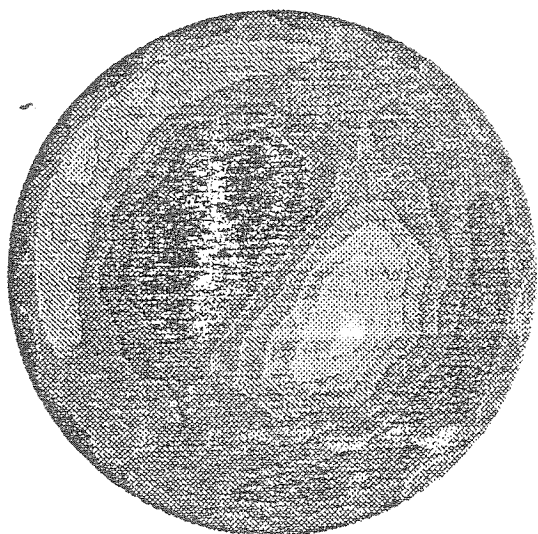
Figure 5.4: *A selection of later singular functions.*



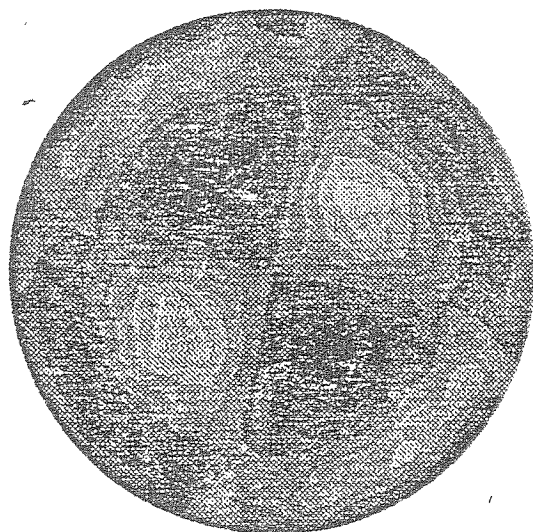
33



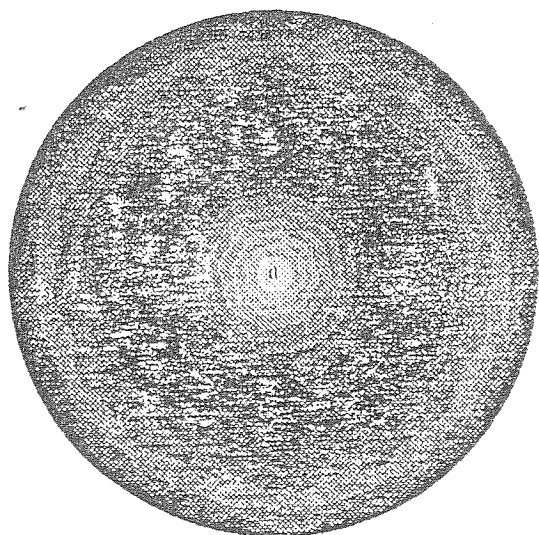
35



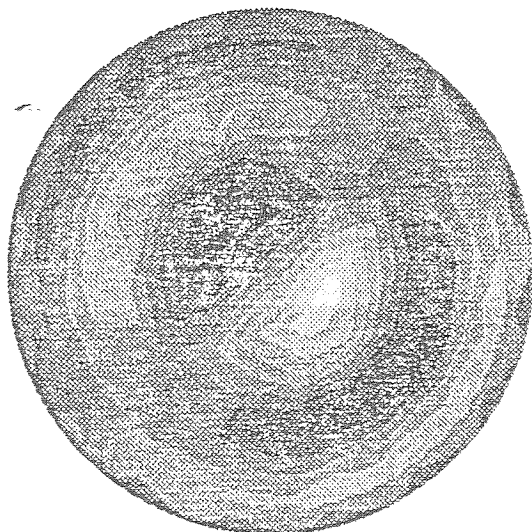
40



49



50



60

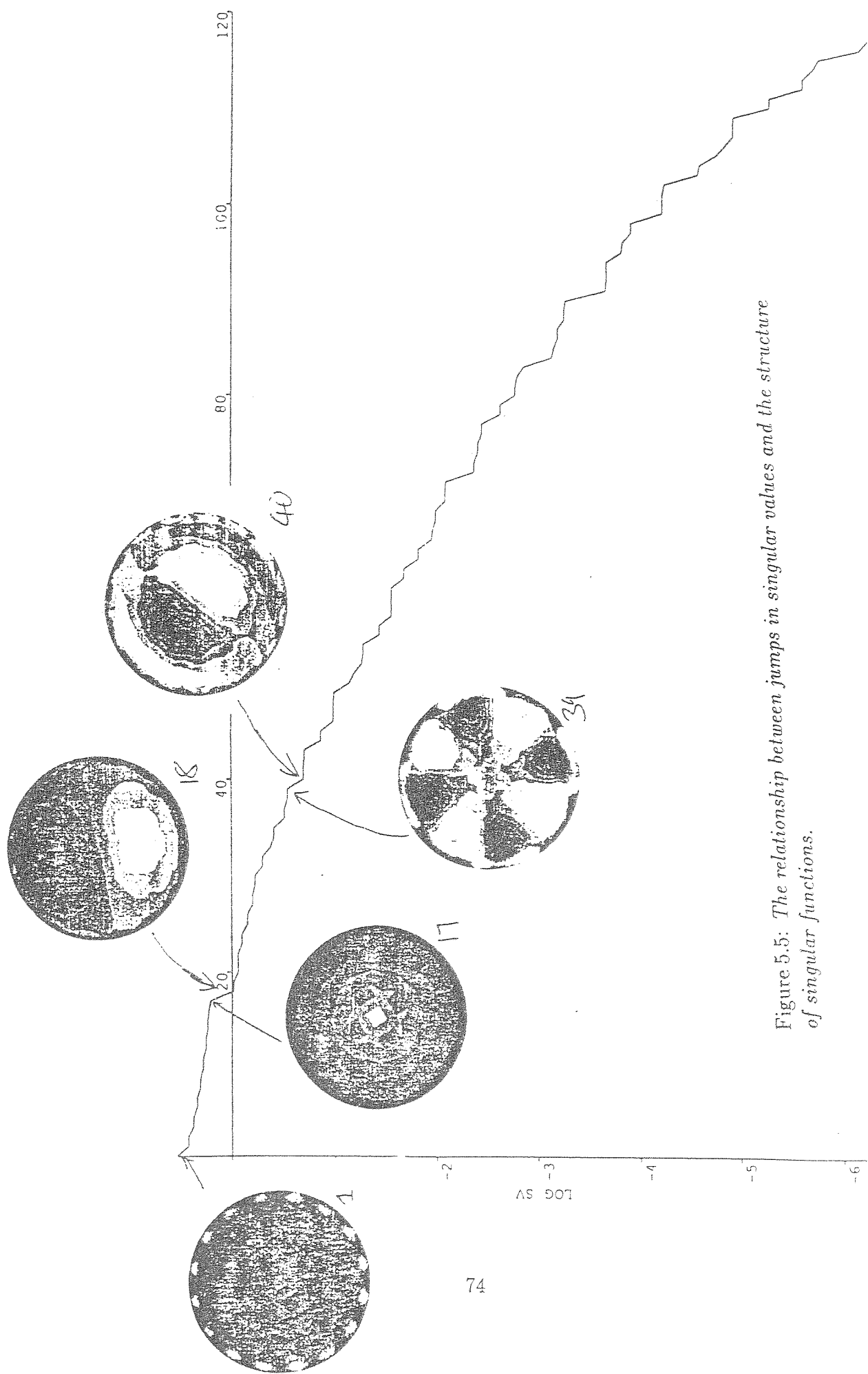


Figure 5.5: The relationship between jumps in singular values and the structure of singular functions.

5.10 Polar or Adjacent?

Nothing has been said yet of the influence of the current patterns used on the SVD. The main effect of varying the current patterns, provided a complete set of data is taken, is to alter the singular functions, the overall decay of the singular values persisting. The singular functions reflect the differing sensitivity of current patterns to conductivity changes where the field strength is greatest. One particular special case of interest is current patterns using pairs of electrodes, as these have been the most widely used in practical EIT systems. Authors who advocate the use of pairs of drive electrodes differ as to which configuration is best. Barber and Brown [5] advocate driving adjacent electrodes, approximating a dipole current source, whereas Tarassenko [89] advocates driving a polar configuration, in which opposite electrodes are driven.

Certainly one factor in choosing which measurements to make is the number of linearly independent measurements possible with a particular measurement scheme. A useful way to look at this is to study the rank of the derivative matrix A . In real life, with measurement error, we are not concerned so much with the exact rank of the matrix, but the effective rank for a given precision of measurement. That is, if two rows of the matrix are linearly dependent *to within measurement precision* the measurements corresponding to those rows contain no different information³. The *effective rank* of A to precision ϵ may be defined as the number of singular values greater than ϵ .

Let us study first the exact rank of A . For clarity we will consider Ω to be the disk in R^2 , although this is not essential. We envisage a system with N equally spaced electrodes, numbered 0 through to $N - 1$. Let us suppose that current is driven through electrode k to electrode $k + p$, for $0 \leq k \leq N - 1$. We call p the offset. We can then measure the voltage difference V_{kj} between electrodes $k + j$ and $k + j + 1$, for $0 \leq j \leq N - 1$ (electrode numbers will be assumed to be reduced modulo N). These N^2 voltage measurements are not linearly independent: two types of dependency exist in all cases except the special case of $p = N/2$ where N is even which is considered below. Now taking v_{kj} to be the absolute voltage at electrode $k + j$, with respect to some fixed reference potential, then $V_{kj} = v_{kj} - v_{k,j+1}$ and we have

$$\sum_{j=0}^{N-1} V_{kj} = v_{k0} - v_{k1} + v_{k2} - v_{k3} + \cdots + v_{k,N-1} - v_{k0} = 0 \quad (5.23)$$

for $0 \leq k \leq N - 1$. This gives N independent linear equations relating the voltage measurements.

³The concept of linearly dependent to within measurement precision must be treated with extreme caution as it is not an equivalence relation

The second type of linear relation comes from what in the Electrical Engineering literature is called the Reciprocity Theorem [32], and in mathematics is the fact that R_γ is self adjoint. In electrical terms assume that a current I_1 is passed between a pair of electrodes A_1 and B_1 and the voltage V_1 is measured between an identical pair of electrodes A_2 and B_2 . A current I_2 is then applied to A_2 and B_2 the voltage V_2 between A_1 and B_1 satisfies

$$V_1 I_2 = V_2 I_1. \quad (5.24)$$

It follows that if we assume that we have identical electrodes, and that a constant current source is used across the drive electrodes we have

$$\sum_{i=0}^p V_{k,j+i} - V_{j,k+p} = 0. \quad (5.25)$$

This is because the voltage between electrode j and $j + p$ is the sum of the voltage differences between intermediate adjacent electrode pairs. For $0 \leq k < j \leq N - 1$, the equations in 5.25 are independent of each other giving $N(N - 1)/2$ equations.

The N^2 possible measurements satisfy, therefore, $N + N(N - 1)/2$ relations leaving

$$N^2 - N - N(N - 1)/2 = N(N - 1)/2 \quad (5.26)$$

independent measurements, in particular 120 for the case of $N = 16$

An interesting exception to this is the particular case where $N = 2M$ is an even number and one is driving opposite electrodes — that is $p = M$. In this case we have the symmetry relation $V_{kj} = -V_{k+M,j+M}$ for each k and j . Now amongst the N^2 variables we have $N^2/2$ relations of symmetry, M sum relations from Equation 5.23 and only $M(M - 1)/2$ reciprocity relations from Equation 5.25 which are independent of each other. This is because Equation 5.25 is equivalent to the same equation with k and j replaced by $k + M$ and $j + M$. The total number of independent measurements is

$$N^2 - N^2/2 - M - M(M - 1)/2 = M(3M - 1)/2. \quad (5.27)$$

For the example $N = 16$ this works out at 92. The number of independent voltage measurements, the row rank of A , will be the rank of A provided the discretisation of the conductivity allows sufficient variation.

This consideration, the reduction in the number of independent measurements for the polar drive configuration led Tarassenko [89] to choose a ‘near-polar’ arrangement in which current is driven between electrodes which are one electrode away from being opposite. This certainly increases the row rank of A by 28 but there is cause for suspicion. If polar electrodes were slightly

misplaced would this result in a dramatic increase in the number of independent measurements ? Intuition would indicate not. Barber and Brown made a different choice. Given that polar electrodes give less information go to the other extreme and use adjacent drive pairs. Tarassenko argued that near polar drives result in higher current densities in the centre of the body where the sensitivity is lowest but the interest greatest. Barber and Brown argue adjacent drives give better reconstructions with their algorithm.

The SVD reveals something of resolution to this conflict. A plot of the logarithmic singular values of A for all possible values of the offset p was made (see Figure 5.6). It was found that the singular values were nearly indistinguishable on the graph for offsets between 1 and 6. The graph indicates a super linear decay followed by a dramatic decrease at 120, the expected exact rank. In contrast, the polar drive configuration with offset 7 resulted in a sharp decrease in the singular values at 92, again the expected exact rank. However the 'knee' in the curve at 92 is already below a relative error level of 10^{-4} , which is beyond any practically achievable accuracy. We conclude that the additional relations between the measurements for the polar drive configuration are irrelevant in a practical system.

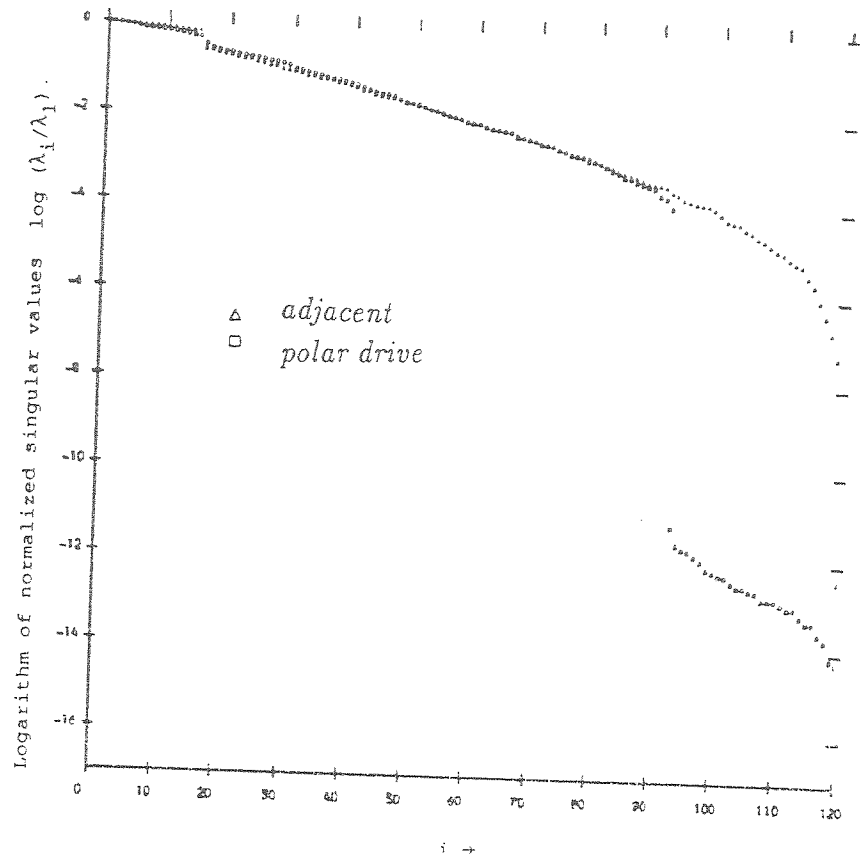


Figure 5.6: Normalized logarithmic singular values for adjacent and polar drive configurations. Singular values for intermediate offsets were indistinguishable from the adjacent case on this scale.

Chapter 6

Reconstruction Algorithms

...the Game was so far developed that it was capable of expressing mathematical processes by special symbols and abbreviations. The players, mutually elaborating these processes, threw these abstract formulas at one another, displaying the sequences and possibilities of their science.

Hermann Hesse, *The Glass Bead Game*

6.1 Introduction

In this chapter algorithms are studied for solving the EIT reconstruction problem numerically. The non-linear problem is ill-posed and its linearisation is ill-posed. It is necessary to regularise to produce a well posed problem. The only thing one can do with an ill-posed problem is to ask a different, well-posed, question. Just as Joshu by replying ‘Mu’ rather than ‘Yes’ or ‘No’ caused his student to see the question of a dog having or not having a Buddha-nature in a completely different light. Typically our search for a more appropriate question will involve returning to the source of the problem. For example the person who asked the ill-posed question. In EIT the original ill-posed problem is ‘Find the conductivity distribution corresponding to these measurements’: a problem which might have been asked by a medical physicist. On further questioning, they may find the well posed problem ‘Find a conductivity distribution with a specified smoothness agreeing with this data to within measurement precision.’ to be closer to what they had in mind. Ultimately this could have come from a clinician whose question was ‘Does this data lead us to suspect a haemorrhage in this premature infant?’.

The approach taken to ill-posedness in this chapter will be *regularisation* in the sense used by Tikhonov [91]. This corresponds to ‘asking a different question’ by changing our criteria for what constitutes a solution. Given that the data we use has measurement error and that there will be errors in our model, we will accept a solution which fits the data as closely as possible, but not exactly, and amongst the possible candidates for such a solution we shall select one which satisfies a given smoothness criterion.

In addition to the ill-posedness of the reconstruction problem, another difficulty we need to overcome is its non-linearity. In the absence of direct methods for this type of non-linear problem we will inevitably have to use an affine approximation at some stage in the algorithm. As we have seen in Chapter 5 both the non-linear problem and its affine approximation using the derivative, are ill-posed. One has a choice then at what stage to regularise. The regularisation can be performed before or after linearisation. The regularisation itself can then be performed in many possible ways. This results in an enormous number of possible alternative strategies only a selection of which will be covered in this thesis. Some have been considered in some depth and I have been able to implement and test them. Others are methods which have been implemented by other authors which I attempt to put in a more unified setting. Others still are pointers towards future work.

6.2 Non-linear Data Fitting

6.2.1 How many minima?

The EIT reconstruction problem can be regarded as a non-linear data fitting or parameter identification problem. The data being the set of voltage measurements taken and the parameters to be fitted the free conductivities. We seek to minimise a ‘cost’ function which measures the discrepancy between the measured voltage data and the voltage data derived from our model. One must distinguish carefully between minimization problems where the objective is to find *any* minimum or near-minimum of an objective function and those in which the solution sought is *the* solution which best explains the data. Examples of the former are a least cost solution to a design problem (such as the placement of transistors on a VLSI¹ chip), or a model which will be used for prediction but not explanation (such as tide tables). In these cases uniqueness is not a problem and a local minimum of the objective function will often suffice, the cost of finding a global minimum may be prohibitive and it may be unnecessary to do so. In the first category uniqueness of the minimum is not a

¹Very Large Scale Integration

problem, any solution which gives a low value of the cost function will do. In the second category we are more interested in the solution than the cost. The EIT problem falls into the second category, we seek a conductivity distribution which subject to certain *a priori* hypotheses best explains the observed data. Fortunately numerical evidence suggests that the problem is not one in which the surface of the objective function is pitted with many distinct local minima for the unwary minimisation algorithm to fall into. Rather it is like a bent, (i.e. non-quadratic) deep sided valley with a very flat floor where the lowest point is hard to find. Plots were made which show contours of the objective function when two degrees of freedom of the conductivity were varied². They vary in shape depending on the two variable conductivity modes. All exhibit a single minimum, the contour shapes varying from elliptical to banana-shaped. See Figure 6.1. This numerical evidence is encouraging but not conclusive. It would be more satisfactory to have better analytical results.

6.2.2 Newton's Method for a critical point

Let us first define some notation. In this chapter the assumption is that the problem is in discrete form. The conductivity is represented by a finite dimensional vector γ of k free conductivity parameters (such as nodal conductivities in a finite element representation). A finite set of currents $\mathbf{j}_1, \dots, \mathbf{j}_n$ is applied to $\partial\Omega$ each represented as an m -vector of currents at each of the m electrodes, together they form the matrix \mathbf{J} . A set of m true measurements of voltage \mathbf{v} is made from the apparatus (the number of current injecting electrodes and voltage measurements need not be the same but in practice this is usually the case). The discrete forward mapping is represented by $\mathbf{F}(\mathbf{J}, \gamma)$. We seek a γ so as to get agreement

$$\mathbf{v} = \mathbf{F}(\mathbf{J}, \gamma). \quad (6.1)$$

However in the presence of error from measurement and modelling we will not get exact agreement. Instead we seek to minimise the error $\|\mathbf{d}\|$ where

$$\mathbf{d} = \mathbf{F}(\mathbf{J}, \gamma) - \mathbf{v}. \quad (6.2)$$

The measurements made be made in many different patterns, for example between adjacent electrodes or between one fixed electrode and the rest. Alternatively the measurements could be scaled or indeed formed by linear combinations of other measurements. The choice of measurement data set will be assumed to have been built in to the forward mapping \mathbf{F} . There will therefore be no loss in generality in measuring the quality of our approximation by the

²These were produced using a transputer implementation of `fwprob` written by Kevin Paulson.

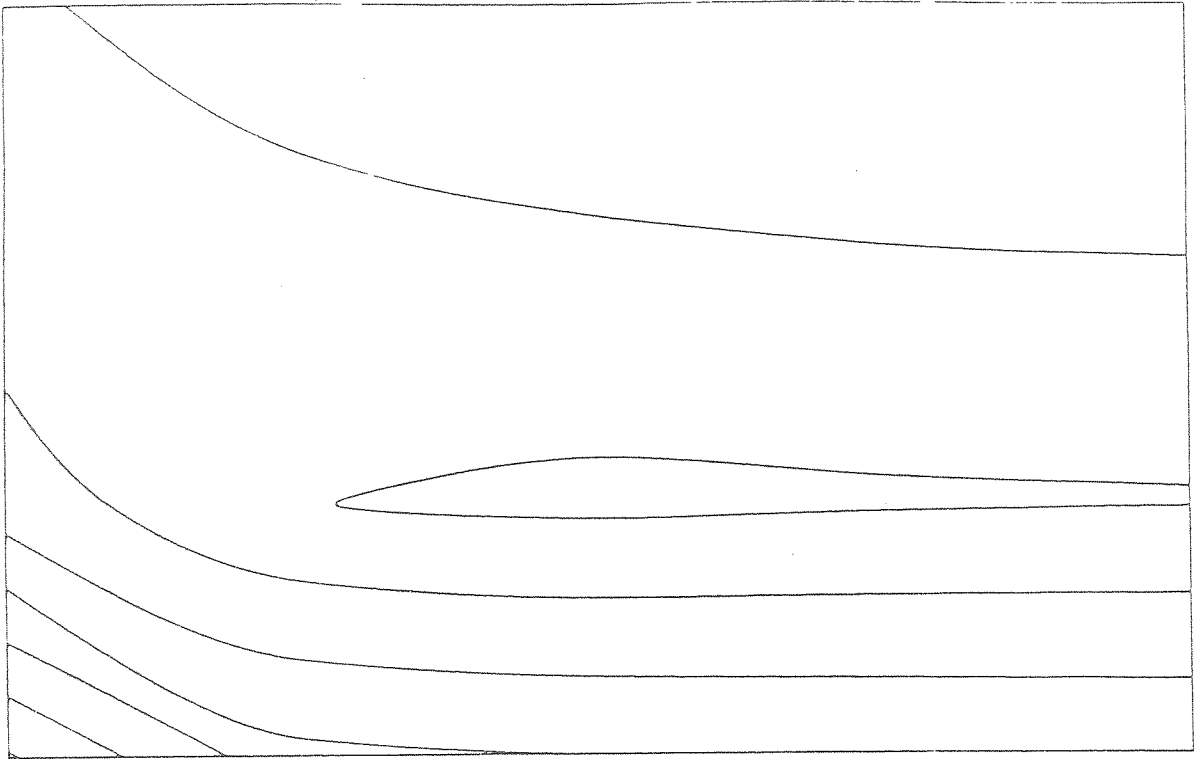


Figure 6.1: A 73 node finite element model of the two dimensional disk was taken. Two nodes were selected, one near the edge and the other in the centre. An initial conductivity was taken to be 100 at each of these nodes and 1 else where. The horizontal and vertical axes represent variations in these two nodal conductivities and the contours the norm of the difference in voltage readings.

objective function

$$e(\gamma) = \frac{1}{2} \|d\|^2. \quad (6.3)$$

At a critical point of e we have $De(\alpha) = 0$ for every $\alpha \in R^k$. The global minimum we seek is in particular a critical point so we may seek it by making the Newton-Kantarovich step [50] from our starting point γ . The step is $\delta\gamma$ where

$$De(\alpha) = -D^2e(\alpha, \delta\gamma) \quad (6.4)$$

for all α . More explicitly

$$-\langle d, DF(\alpha) \rangle = \langle d, D^2F(\alpha, \delta\gamma) \rangle + \langle DF(\alpha), DF(\delta\gamma) \rangle \quad (6.5)$$

where $\langle \cdot, \cdot \rangle$ is the standard inner product in R^{mn} . Using the notation A for the matrix $\partial F_i / \partial \gamma_j$ and H_l for $\partial^2 F_l / \partial \gamma_i \partial \gamma_j$ this can be re-expressed as

$$(A^T A + \sum_{l=1}^{nm} d_l H_l) \delta\gamma = -A^T d. \quad (6.6)$$

If the second derivative were non-singular then this would uniquely define $\delta\gamma$. Even if this were not the case any $\delta\gamma$ satisfying Equation 6.6 is a minimizer of the second degree Taylor polynomial for e . In that sense Equation 6.6 can be thought of as a quadratic approximation to the original problem.

6.2.3 Calculation of the second derivative

Calculation of the second derivative matrices is relatively expensive. From the Taylor series in Chapter 3 we see that

$$D^2F_l(\alpha, \alpha) = (\overline{GL}\alpha)^2 u_l \quad (6.7)$$

where u_l is the potential when current j_l is applied. To convert this quadratic form into a bilinear form we use the polarization identity

$$D^2F_l(\alpha, \beta) = \frac{D^2F_l(\alpha + \beta, \alpha + \beta) - D^2F_l(\alpha - \beta, \alpha - \beta)}{4}. \quad (6.8)$$

To calculate $\partial^2 F_l / \partial \gamma_i \partial \gamma_j$ explicitly let η_i be the i th basis function in the space of discretized conductivities and define $\xi_{ij}^+ = \eta_i + \eta_j$ and $\xi_{ij}^- = \eta_i - \eta_j$ then

$$\partial^2 F_l / \partial \gamma_i \partial \gamma_j = \frac{D^2F_l(\xi_{ij}^+, \xi_{ij}^+) - D^2F_l(\xi_{ij}^-, \xi_{ij}^-)}{4} \quad (6.9)$$

or more explicitly

$$\partial^2 F_l / \partial \gamma_i \partial \gamma_j = \frac{(\overline{GL}\xi_{ij}^+)^2 u_l - (\overline{GL}\xi_{ij}^-)^2 u_l}{4}. \quad (6.10)$$

This formulation requires the matrices corresponding to the Greens operators to be calculated, in discrete terms this is equivalent to the calculation of the inverse of the system stiffness matrix. If the Choleski factorization of the matrix were known this would amount to performing k forward and backward substitutions this would take $O(kp^2)$ floating point operations where p (which is at least k) is the number of degrees of freedom in the finite element mesh. The product $GL\xi_{ij}^+$ would require the assembly of a system stiffness matrix for the conductivity ξ_{ij}^+ and a matrix multiplication of order $O(p^3)$. The resulting matrix must be squared, another operation of order $O(p^3)$. As this must be performed for each i and j this results in $O(k^2p^3)$ operations. For a first iteration from known conductivity these calculations could be done in advance, but for subsequent iterations this 5th order process should be avoided if possible.

Perhaps the simplest approach to the problem of calculating the second derivatives would simply be to ignore them and solve

$$A^T A \delta \gamma = -A^T d \quad (6.11)$$

which is equivalent to taking the affine approximation to \mathbf{F} and minimising $\|d - A\delta\gamma\|^2$. However we have seen that this is illposed as A is the matrix of a compact mapping and has rapidly decreasing singular values. Indeed there is a sense in which the linear problem may be worse than the non-linear in that even if v may be in the range of \mathbf{F} there is no guarantee that d will be in the range of the derivative $D\mathbf{F}$.

At this point the problem can be looked at from two points of view. One is to concentrate on this ill-posed linear problem, and apply the standard regularisation techniques the other is to go back to the (still ill-posed) quadratic problem (Equation 6.6) and look at ways of modifying the second derivative terms. Yet another alternative is to go back to initial non-linear problem and create a well posed non-linear problem which can then be tackled by one of the standard methods for solving non-linear problems, some of which will be discussed in the next section.

6.3 Regularised Newton Methods

6.3.1 The Levenberg-Marquardt Method

As a linear step in an iterative method, Levenberg [59] and Marquardt[64] suggest one solves the Tikhonov regularised linear approximation

$$(A^T A + \mu I)s = -A^T d \quad (6.12)$$

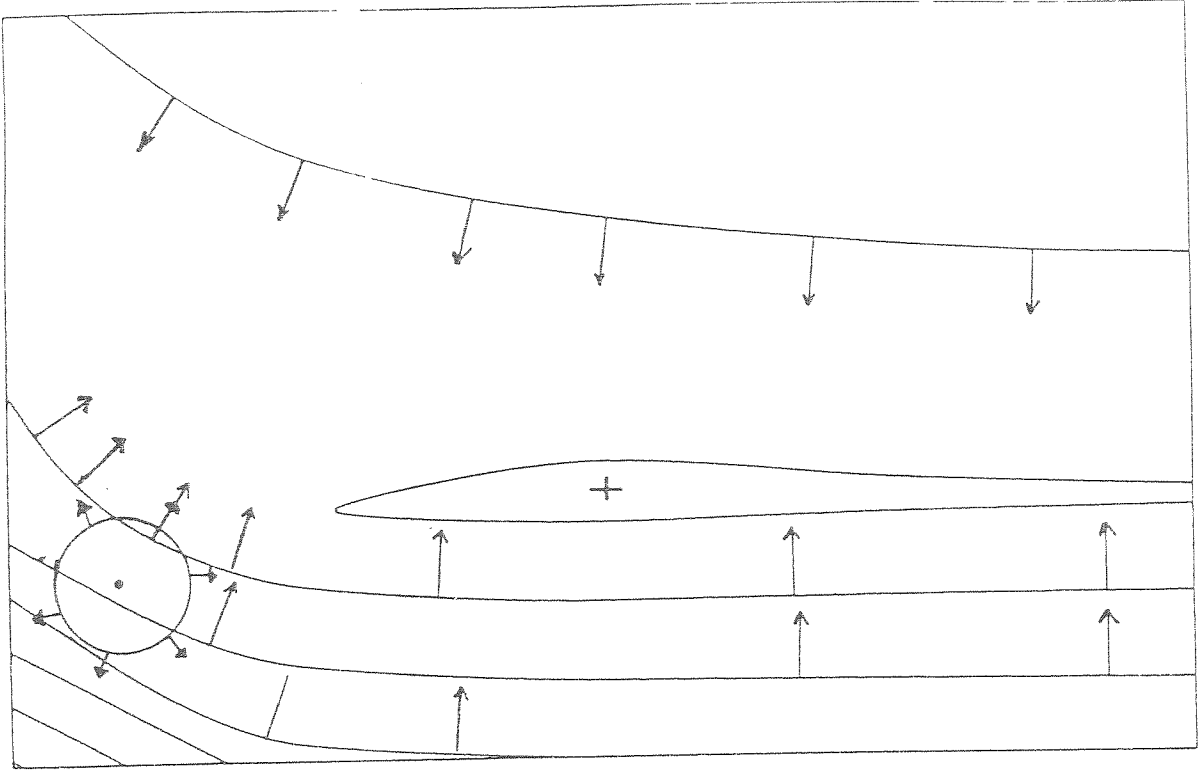


Figure 6.2: Contours of the objective function (see Figure 6.1) showing a hypothetical approximate solution γ_0 and circular trust region. The long arrows indicate the steepest descent direction, the short arrows the gradient of $\|s\|^2$.

which is the minimiser of $\|As + d\|^2 + \mu^2\|s\|^2$. This is equivalent to the constrained minimisation problem: minimise $\|As + d\|^2$ subject to $\|s\| \leq \rho$. Here μ is the Lagrange multiplier when the constraint is active. The ball of radius ρ centred at the current approximation γ_0 is called the ‘trust region’ (see Figure 6.2) as it represents a region in which we can trust the affine approximation to F .

For $\mu = 0$ the method becomes simply Newton-Kantorovich but for large μ the direction of the update vector $\delta\gamma$ tends to the direction of the steepest descent direction $-A^T d$. In this way the Levenburg-Marquardt update can be thought of as an interpolation between the slow but sure steepest descent method and the rapid but unreliable Newton-Kantorovich. Marquardt [64] himself puts it as follows:

... a *maximum neighbourhood* method is developed which, in effect, performs an optimal interpolation between the Taylor series method and the gradient method, the interpolation being based on the maximum neighbourhood in which the truncated Taylor series gives an adequate representation of the nonlinear model.

It remains to say how the regularisation parameter μ is chosen. Marquardt [64] gives a scheme based on decreasing μ by a constant factor provided this improves the residual error compared with the previous iteration, or exceptionally increasing μ by the same factor if no improvement in the error can be made otherwise. Gradient methods (that is steepest descent) are characterised by good initial progress followed by extremely slow convergence, whereas Newton-Kantorovich may at some times actually diverge but when close to the solution converges superlinearly. As μ is reduced the method becomes closer to Newton-Kantorovich and thus converges faster. More recent developments of the Levenberg-Marquardt algorithm are given by Moré [65]. These include, in particular, more sophisticated schemes for the choice of the regularisation parameter such as that due to Hebden [42].

However for ill-posed problems where $A^T A$ may not be invertible (its continuous counterpart being an operator with an unbounded inverse). This means that the regularisation parameter must not be reduced to zero. If an estimate for the error in the data is available, one strategy is to not reduce μ below a level where the radius of the trust region is equal to that error.

6.3.2 Using the SVD.

The system of linear equations in the Levenburg-Marquardt method can be solved by any method suitable for solving a symmetric system of linear equations. The first iteration, assuming that an initial approximation to the conductivity (eg $\gamma = 1$) is known, can be performed quickly by storing a pre-computed inverse or decomposition of the matrix. The singular value decomposition has several advantageous properties for this purpose. Tikhonov regularisation can be represented as a filtered singular value decomposition (see [35]). This means that the stored SVD can be used to solve the first LM step for any value of μ . Also the components of the right hand side corresponding to right singular functions with small singular values can be ignored. That is a projection can be taken onto the space spanned by the first n singular functions where $\lambda_i < \epsilon$ for $i \leq n$ and ϵ an estimate of the L^2 measurement error. This reduces the amount of arithmetic required for the first step as well as serving as a regularisation technique. Extending this idea further a stored SVD allows other regularisation techniques to be used other than Tikhonov or

truncated SVD.

While a stored SVD seems ideal for the first step in the LM algorithm, it is computationally expensive (see Dongarra [26]) and therefore unsuitable for subsequent steps. The numerical experiments performed by the author to date have either used Choleski factorisation or recalculated the SVD so that the structure of the derivative matrix could be followed at each iteration. To establish the principle, speed of execution was not considered important. Reports on these experiments were first published in Breckon and Pidcock [15, 16], and more details are given later in this Chapter.

6.3.3 Iterative Methods

Linear solution methods can be classified broadly as iterative or direct although compromises between the two also exist. Direct methods usually provide answers to within machine precision for well conditioned problems in a known number of operations. Iterative methods however approach such accuracy asymptotically, the number of steps required depending on the quality of the initial guess. The advantage of using iterative techniques in this context is that one need not require the linear equation to be solved to within full machine accuracy. The data collected would typically only be measured to 12 bits, and is unlikely to be that accurate. Double precision accuracy using 8 byte real numbers typically includes about 50 bits for the mantissa. This means that a much smaller number of iterations need be done for each linear step. A slightly more sophisticated approach to the same argument is to treat a truncated iterative method itself as a regularisation procedure. One can then apply the Morozov stopping criterion (Morozov [66]) of terminating an iterative solution method when it has converged to within the accuracy of the data.

Three well known iterative techniques with interesting regularisation properties are successive approximation, steepest descent and conjugate gradient. Here they are stated in terms of solving the system of linear equations $Ax = b$. In each of these methods, a sequence x_i of successive approximations to the generalised solution is calculated. Define the error r_i by

$$r_i = A^T A x_i - A^T b. \quad (6.13)$$

The iteration scheme for successive approximation (also called Landweber's Method [3]) is then

$$x_{i+1} = x_i + \tau r_i \quad (6.14)$$

where τ is a fixed relaxation parameter. Steepest descent is given by a similar formula

$$x_{i+1} = x_i + \tau_i r_i \quad (6.15)$$

but in this case τ_i is given by

$$\tau_i = \|r_i\|^2 / \|Ar_i\|^2. \quad (6.16)$$

The conjugate gradient method is slightly more sophisticated, the iteration being given by

$$x_{i+1} = x_i + \zeta_i p_i \quad (6.17)$$

where

$$p_0 = r_0 = -A^T b, \quad p_i = r_i + \nu_{i-1} p_{i-1} \quad (6.18)$$

and

$$\zeta_{i-1} = \langle r_i, p_i \rangle / \|Ap_i\|^2, \nu_{i-1} = -\langle Ar_i, Ap_{i-1} \rangle / \|Ap_{i-1}\|^2 \quad (6.19)$$

The regularisation properties of these iteration schemes can be understood in terms of singular values [88]. For example the i -th term in the successive approximation is given by

$$x_i = \sum_k s(\lambda_k) \langle f, \phi_k \rangle_X \psi_k \quad (6.20)$$

as in Section 5.5. The filter in this case being

$$s_i(t) = \frac{1 - (1 - \sqrt{\tau} t^2)^i}{t}. \quad (6.21)$$

it can be seen from the form of the filter that for large i the iteration converges to the Moore-Penrose solution A^\dagger provided $\sqrt{\tau} \lambda_1 < 1$. For λ_k large, convergence is rapid, but for small singular values it is slow. Contributions from the large singular values appear early in the iteration process, while those from smaller singular values appear later. Stopping after a finite number of steps is, therefore, a regularisation method. Taking the iteration too far contaminates the solution with errors present in the contributions of the small singular values. This semi-convergent behaviour for iterative methods applied to ill-posed problems is quite general.

Marquardt proves that for his Tikhonov regularised step the angle θ between the steepest descent direction $A^T b$ and the vector $s_\mu = (A^T A + \mu I)^{-1} b$ is a continuous monotone decreasing function of μ such that $\theta \rightarrow 0$ as $\mu \rightarrow \infty$. As $\mu \rightarrow 0$ the vector s_μ approaches the Moore-Penrose solution $A^\dagger b$. All the above iterative solution methods share a similar property. The initial guess is always in the steepest descent direction and for well posed problems they converge to $A^\dagger b$. This shows that a truncated iteration method, that is an iterative linear technique stopped short of convergence to machine accuracy, can replace Tikhonov regularisation in the linear step of the Levenberg-Marquardt method. The role of the regularisation parameter is taken by the number of

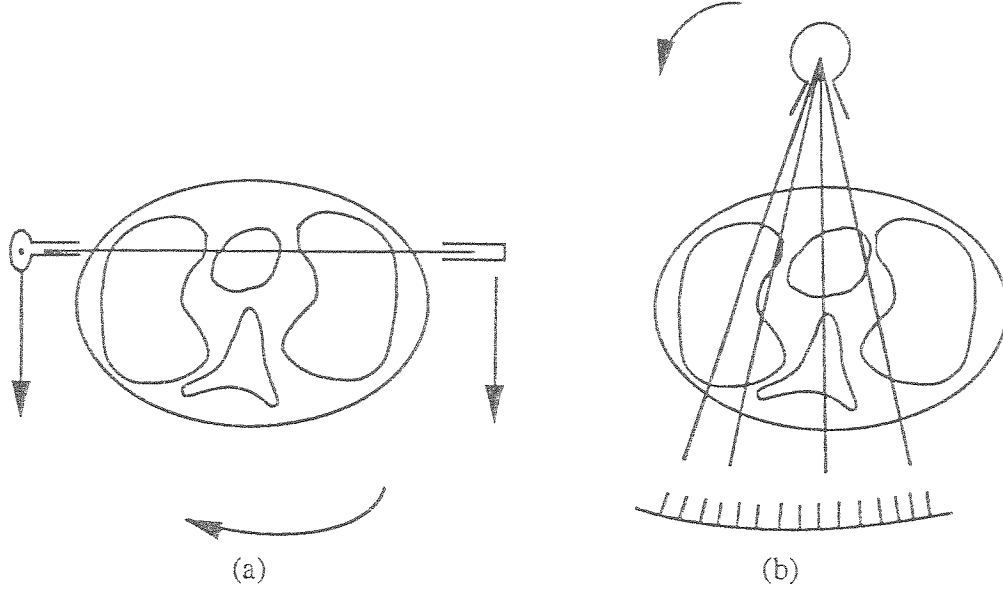


Figure 6.3: *X-Ray CT scanning geometries: (a) Parallel, (b) Fan Beam.*

iterations performed the trust region radius being the size of the linear residual $\|r_i\|$. A potential advantage of the use of an iterative linear solution technique is the reduced computational cost of the linear step of a non-linear algorithm. In the early stages of the iteration, or in the presence of high noise, only a few linear iterations will be required. The best regularised linear solution technique to use will depend on a number of factors used including the architecture of the computer and the accuracy of solution required. The use of these methods on Multiple Instruction Multiple Data machines is the subject of an investigation by Paulson, and is reported in [17]

6.4 X-Ray CT

Most of the algorithms devised by medical physicists have been inspired by methods used in X-Ray CT. It is necessary, therefore, to review X-Ray CT reconstruction algorithms. In X-Ray CT collimated beams of X-Ray radiation are passed through the patient and measured by a detector on the opposite side. The X-ray source and detector are mounted on a device which allows them to be moved in such a way as to scan a two dimensional section of the patient. This is done either by a combination of rotary and lateral motion (parallel geometry) or by two rotary motions about different centres (fan-beam geometry). See Figure 6.3.

If $\theta \in S^1$, is on the unit circle, and $p \in R$ then the Radon transform of a function f defined on a two dimensional domain is

$$(\mathcal{R}f)(\theta, p) = \int_{x \cdot \theta = p} f(x) dx. \quad (6.22)$$

This generalises easily to the case of $\theta \in S^{n-1}$ and $p \in R$, and still further to integrals over families of submanifolds [9,31].

There are two principal classes of methods for X-Ray CT reconstruction (For details see [72]). The first, Radon transform inversion, is also called filtered back-projection. This relies on the special structure of the Radon transform which has an explicit inversion formula. The second major class uses a discretized matrix representation of the Radon transform. The resulting system of linear equations is solved using iterative matrix inversion techniques. The most widely known examples of this are Algebraic Reconstruction Technique (ART) [40] and Simultaneous Iterative Reconstruction Technique (SIRT) [33].

The adjoint operator $\mathcal{R}^* : L^2(R^2) \longrightarrow L^2(S^1 \times R)$ is given by the formula

$$(\mathcal{R}^*g)(x) = \int_{S^1} g(\theta, x \cdot \theta) d\theta. \quad (6.23)$$

This is known as the back-projection operator as the measurement of X-ray intensity $g(\theta, x \cdot \theta)$ is projected back along the line $x \cdot \theta = p$. Back-projection itself does not constitute a solution method, however by taking Fourier transforms it can be shown [72] that

$$(\mathcal{R}\mathcal{R}^* + \mu I)^{-1}g = h * g \quad (6.24)$$

where h is the filter function

$$h(t) = \frac{1}{2}(2\pi)^{-\frac{3}{2}}|t|S_\mu(t) \quad (6.25)$$

where

$$S_\mu(t) = \frac{1}{1 + \frac{1}{2}(2\pi)^{-3/2}\mu|t|^{-1}}. \quad (6.26)$$

The convolution is taken with respect to the second variable. Filtered back-projection is therefore a method of reconstructing an X-Ray CT image implicitly using Tikhonov regularisation.

In iterative methods a discrete approximation to the Radon transform is used, usually using a pixel basis [92]. This matrix will be denoted by A . Following the convention adopted in the above discussion of iterative algorithms, the data vector will be denoted by b and the image to be reconstructed by

x . The data is naturally partitioned into blocks, We will think of the LM equations represented by $Ax = b$ as being partitioned into M blocks each of size L . The j -th row of A will be denoted by a_j . If we denote $i \pmod{M} + 1$ by m_i then a generalisation for the ART iteration formula (following [45]) is

$$x_{i+1} = x_i + \tau \sum_{\ell=1}^L \frac{b_j - a_j x_i}{\|a_j\|^2} a_j \quad (6.27)$$

with $j = (m_i - 1)L + \ell$. Here τ is a relaxation parameter.

For the particular case of a block size of $L = 1$ the algorithm is that of Kaczmarz [92,49] which is the same as the original ART proposed by Gordon [40]. If the number of blocks M is 1 then the method is the classical method of Cimmino [25], which is very close to the SIRT method of Gilbert *et al* [33]. Another obvious choice for the block size is the number of measurements taken for each projection angle of the CT scan.

It will be useful to introduce the more general framework of Eggermont [27]. The matrix A will be partitioned into blocks A_m each of L rows of A , and b partitioned similarly

$$A_m = \begin{pmatrix} a_{(m-1)L+1} \\ \vdots \\ a_{mL} \end{pmatrix} \quad (6.28)$$

$$B_m = \begin{pmatrix} b_{(m-1)L+1} \\ \vdots \\ b_{mL} \end{pmatrix}. \quad (6.29)$$

The iterative step given by Eggermont is then

$$x_{i+1} = x_i + A_{m_i}^T \Sigma_{m_i} (B_{m_i} - A_{m_i} x_i) \quad (6.30)$$

where the Σ_m are referred to as relaxation matrices. Eggermont proves that, provided the equations are consistent and the starting value x_0 is in the range of A^T , the algorithm converges to the Moore-Penrose solution

$$\lim_{i \rightarrow \infty} x_i = A^\dagger b \quad (6.31)$$

as long as the relaxation matrices satisfy

$$\|A_m^\dagger (I_L - A_m A_m^T \Sigma_m) A_m\|_2 < 1. \quad (6.32)$$

The algorithm described by Equation 6.27 is a particular case with

$$\Sigma = \text{diag}(\tau \|a_{(m-1)L+1}\|^{-2}, \dots, \tau \|a_{mL}\|^{-2}). \quad (6.33)$$

Convergence is guaranteed by choosing τ small enough.

There are a very large variety of methods for solving systems of linear equations and it is worth noting the particular features of the reconstruction problem for X-Ray CT have lead to this particular class of methods being employed. The important properties (see [92]) are that the matrix A is:

- *Large.* The matrix is very large — possibly with 10^{10} entries. Typically more rows than columns (an over determined system).
- *Sparse.* Most of the entries in the matrix are zero — typically more than 99%.
- *Singular.* The rank of A will be less than the number of columns.
- *Unstructured.* The distribution of non-zero entries has no regular structure.
- *Ill-conditioned.* The singular values decay with order $O(k^{-1/2})$

Methods like ART are *row* action methods — a single iterative step requiring access to only one row of the matrix. This cuts down the memory requirements of such an algorithm. Also the original matrix, and hence its sparse structure is preserved. As discussed above, iterative methods in general have a regularising effect when stopped short of convergence. This can be used to solve the problem caused by the decay of the singular values. It must be noted, however, that these methods are in some respects not ideal, despite their popularity (see [92]). For example the conjugate gradient method discussed above has superior convergence properties at an only slightly increased storage cost.

6.5 Radon Transform Inversion

6.5.1 Not a Radon Transform

The early work of Tasto and Schomberg [90] assumed that the forward problem in EIT could be represented as a generalised Radon transform. However, there is a simple example which shows that the forward problem for EIT is not a generalised Radon transform. The demonstration of this elaborates on an idea of Bates *et al* [4].

First it is necessary to define a generalised Radon transform (GRT). The strict definition given by Gel'fand *et al* [31] is as a double fibration, however

since the counter example is rather simple, it is unnecessary to introduce the full sophistication of the definition. The counter example has as its domain the unit disk D in R^2 , so we shall give a definition of a GRT for this case.

Let I be an interval in R , then for any point α on the circle S^1 and any number $\beta \in I$, let $L_{\alpha,\beta}$ be a curve in D . These curves will be considered as the ‘rays’ along which integrals are taken. In the case of the standard Radon transform the rays would be $L_{\alpha,\beta} = \{x \in D | x \cdot \alpha = \beta\}$. The rays $L_{\alpha,\beta}$ will be assumed to depend smoothly on α and β . Moreover, we require that for a fixed α the $L_{\alpha,\beta}$ do not intersect (strictly for each α the rays form a foliation of D). For a function $f \in L^1(D)$ we define the GRT of f as

$$(\mathcal{R}f)(\alpha, \beta) = \int_{L_{\alpha,\beta}} f. \quad (6.34)$$

Each ray will be assumed to have a measure depending smoothly on α, β . This can be thought of as a (positive) weighting function $w_{\alpha,\beta}$ so that

$$(\mathcal{R}f)(\alpha, \beta) = \int_{x \in L_{\alpha,\beta}} f(x) w_{\alpha,\beta}(x) dx. \quad (6.35)$$

The GRT is a linear operator, so at best we could hope for agreement with the linearised forward problem for EIT. Consider the one-parameter family of conductivities

$$\gamma_t(r, \theta) = \begin{cases} 1+t, & \text{for } r < \rho \\ 1, & \text{for } \rho \leq r \leq 1 \end{cases} \quad (6.36)$$

Given some one-parameter family of current patterns j_α , and some parameterisation $\theta(\beta)$ of a subset of ∂D , let us assume that to first order the potential on the boundary $u_\alpha(\theta)$ is a GRT of the conductivity. That is

$$\left. \frac{\partial u_\alpha(\theta(\beta))}{\partial t} \right|_{t=0} = \mathcal{R} \left. \frac{\partial \gamma_t}{\partial t} \right|_{t=0}. \quad (6.37)$$

However, suppose that for some α_0 , $j_{\alpha_0}(\theta) = \cos \theta$. In this case

$$\left. \frac{\partial u_{\alpha_0}(\theta(\beta))}{\partial t} \right|_{t=0} = -2\rho^2 \cos \theta \left. \frac{\partial \mu}{\partial t} \right|_{t=0} \quad (6.38)$$

$$= -\rho^2 \cos \theta \quad (6.39)$$

from Equation 2.10. Notice that this vanishes only when $\cos \theta = 0$, hence the integrals along all but possibly two rays are non-zero. However the variation in the conductivity is only non-zero inside a circle of radius ρ . This means that all rays pass through that circle. Taking the limit as ρ goes to zero, we conclude that all rays $L_{\alpha_0,\beta}$ pass through the centre of the disk. This contradicts the non-intersection property of the rays.

6.5.2 Consistent Updates

Despite the counter example of the previous section, one successful EIT reconstruction algorithm uses the idea of a generalised Radon transform — that of Barber and Brown [5]. To motivate their method consider first the problem of finding a conductivity consistent with only one set of Neumann and Dirichlet data. Clearly this problem is under specified and the solution will not be unique. However a unique solution can be found under the initial hypothesis that the gradients of γ and u are parallel (we will assume also that $u, \gamma \in C^1(\Omega)$):

$$d\gamma \wedge du = 0. \quad (6.40)$$

For the definition of d and \wedge in this context see [1]. In coordinates Equation 6.40 becomes

$$\frac{\partial \gamma}{\partial x_i} \frac{\partial u}{\partial x_j} = \frac{\partial \gamma}{\partial x_j} \frac{\partial u}{\partial x_i} \quad (6.41)$$

for all i and j (this argument will work in dimensions two or three so the generality will be maintained). In this particular case, which was treated by Cannon and Halton [23], a unique solution can be found. Notice that Equation 6.40 implies that $d(\gamma du) = 0$ which means that γdu is a closed form (equivalently $\gamma \nabla u$ is an irrotational vector). For a simply connected domain Ω this implies the existence of a function h with

$$\gamma \nabla u = \nabla h. \quad (6.42)$$

Now h satisfies Laplace's equation $\nabla^2 h = 0$ and the same Neumann data as u :

$$-\gamma \nabla_{\mathbf{n}} u = -\nabla_{\mathbf{n}} h = j. \quad (6.43)$$

A straight-forward method for finding both γ and u is now apparent. First h is found by solving the Neumann problem for Laplace's equation. Given an point $x \in \Omega$ look at the intersection of the level-set of h which passes through x with the boundary: $\{y \in \Omega | h(y) = h(x)\} \cap \partial\Omega$. In two-dimensions this set will consist of one or two points, in three-dimensions it will be a curve. The gradients of γ , u , and h are parallel, so the level-sets of all three functions coincide. The value of $\gamma(x)$ is therefore the same as its value on the whole of the level-set of h through x . In particular its value on the boundary can be determined. Let \mathbf{t} be a vector tangential to $\partial\Omega$ then $\gamma \nabla_{\mathbf{t}} u = \nabla_{\mathbf{t}} h$ and

$$\gamma(y) = \frac{\nabla_{\mathbf{t}} h(y)}{\nabla_{\mathbf{t}} u(y)}. \quad (6.44)$$

A solution to the problem will exist exactly when the expression given in Equation 6.44 is constant on $\{y \in \Omega | h(y) = h(x)\} \cap \partial\Omega$ for all $x \in \Omega$.

The method of Barber and Brown [5] is to take a weighted average of conductivities parallel with the equipotentials and consistent with each Neumann-Dirichlet data pair. This is only possible where such conductivities exist which means, as shown above, that the right-hand side of Equation 6.44 must be constant on the intersection of the equipotential lines with the boundary. In three-dimensions this is unlikely to be the case. In two-dimensions agreement of this function at two points is also unlikely. Barber and Brown consider the case where Ω is the two-dimensional disk. However the particular choice of current patterns used by Barber and Brown make the method possible. They take the current pattern to be dipole source:

$$j_\alpha(\theta) = -\pi \frac{\partial \delta_\alpha}{\partial \theta}. \quad (6.45)$$

The equipotential lines, assuming $\gamma \equiv 0$, are semi-circles passing through the point α where the dipole is centred. Voltage data is not known at this point, which is a singularity of u , hence the value of $\nabla_{\mathbf{t}} h(y)/\nabla_{\mathbf{t}} u(y)$ at the other end of the equipotential is used.

Details of the weighting function used in the averaging process, and the connection with Radon transform inversion, can be found in [80].

6.6 Application of ART to EIT

As shown in the previous section, the linearised forward problem in EIT is not a generalised Radon transform. One consequence of this is that the derivative matrix A , in contrast to a discrete Radon Transform, is not sparse. Indeed zeros of the kernel function $K_{kl} = \gamma \nabla u_k \cdot \nabla v_l$ occur only on a set of measure zero. This means that one good reason for using row-action methods which applied to Radon transform inversion, is not valid for EIT. However this does not invalidate the use of ART-like methods which are valid for any matrix.

An application of an ART-like algorithm to EIT can be found in the work of Kim [53]. Both Kim and Yorkey [96] use the term back-projection in a way which may lead to confusion. When treating the Radon transform, back-projection, as defined by Equation 6.23 is equivalent to multiplication by A^T . However in EIT the geometrical interpretation of the adjoint operator is not as clear. It is ‘back-projection’ in the sense that the data on the boundary is cast into the interior. However each point in the interior of the region receives a contribution from all the measurements made, rather than just those made at the boundary points of the current stream-line or equipotential on which it lies.

Kim’s formulation differs slightly from the approach taken in this thesis in

that it treats relative, that is percentage, changes in the resistivity and the relative changes they produce in the currents. For comparison this presentation will use the equivalent absolute measurements. Recall that Kim uses constant voltage sources and measures the changes in current. An appropriate forward mapping is

$$j = F(\rho) \quad (6.46)$$

where $\rho = 1/\gamma$ is resistivity. For the k -th voltage pattern the i -th current measurement will be denoted j_{ki} and its deviation from the current predicted by the finite element model δj_{ki} . The perturbation matrix used by Kim is an approximation to the matrix A of partial derivatives $a_{kil} = \partial j_{ki} / \partial \rho_l$ where ρ_l is the resistivity of the l -th element in the finite element model. Kim's algorithm updates the n -th estimate of resistivity $\rho^{(n)}$ as follows:

$$\rho_l^{(n+1)} = \rho_l^{(n)} + \tau \frac{\sum_i a_{kil} \delta j_{ki}}{\sum_i |a_{kil}|}. \quad (6.47)$$

This iteration is repeated for each voltage distribution k cyclically. The iteration given in Equation 6.47 is very close to the general ART method of Equation 6.27 differing only by the use of the 1-norm in the denominator. For a suitable choice of τ this clearly falls within the more general framework of Eggermont (Equation 6.28) and hence converges to the Moore-Penrose solution (By the equivalence of norms on finite dimensional normed spaces, the 1-norm can be estimated in terms of the 2-norm. This changes the bounds on τ which guarantee convergence). Moreover stopping short of convergence will act as regularisation. This explains the success of the method when applied by Kim. The poor convergence found by Yorkey [96] is understandable when compared to direct methods.

The algorithm as it stands is simply a linear solution method. However it is amenable to natural non-linear extensions. The derivative, or perturbation, matrix can be recalculated at any stage in the iteration. If all measurements are cycled through using Equation 6.47 a number of times before recalculating the matrix, this amounts to another variation of the regularised Newton's method discussed above. The number of iterations is determined by the Morozov stopping criteria. Another alternative would be to recalculate the matrix at each iteration of Equation 6.47, or perhaps an intermediate stage after each voltage profile had been used K times since the last update of the matrix. The determination of optimal K would depend on the speed of solution of the forward problem solution relative to the inverse problem, and the advantage gained by the recalculation of the matrix. Although Kim did not implement this non-linear extension, Yorkey implemented a SIRT version of the Kim method, updating the matrix after each iteration. This was found to converge, but not as rapidly as Levenberg-Marquardt.

6.7 Implementation of Regularised Newton Methods

6.7.1 Reconstructing the Moat Object

Regularised Newton methods were implemented using truncated SVD as the regularisation technique. Initial results were reported in [15]. The mesh used was a square with 100 nodes and square bilinear elements. The sixteen electrodes were taken as edges of elements on the boundary. The current drive pattern used was the adjacent pair technique, as described in Chapter 5. Of the conductivity distributions tested, one of the more difficult to reconstruct was the ‘moat’ object illustrated in Figure 6.4. This consists of a ‘moat’ of conductivity 50 units against a background of 100. One can expect that this ring of higher conductivity will prevent current from penetrating into the ‘mound’ deeper into the region. This will result in problems in reconstructing the central area. This was found to be the case in practice. The first iteration produced a recognisable moat but a much diminished mound. The detail became evident as the iteration proceeded, the mound becoming more accurate only after the moat had taken shape. After five iterations little visible difference was detectable on the reconstructed image, although the algorithm was still able to make changes which reduced the residual 6.5.

6.7.2 The Effect of Data Errors

The effect of errors in the data was investigated by simulating three possible types of inaccuracy. The first was in the position of the electrodes, for both current application and voltage measurement purposes. This was achieved by interpolation using the finite element basis. Secondly the shape of the boundary was distorted by adding a small random vector to the coordinates of the boundary nodes. Finally, uniformly distributed random noise was added to the measurements.

Referring to the singular value decomposition of Chapter 5 it can be anticipated that the effect of even small amounts of noise in the data will be dramatic. This was indeed found to be the case as illustrated in Figure 6.5. The initial iteration, in which only the first few singular functions were used (typically about 30), resulted in a similar reduction in the residual even in the reconstructions with contaminated data. The images obtained in all the cases of contaminated data were all qualitatively similar to Figure 6.4(b). However with the contaminated data no further improvement can be made after the second iteration as the residual is already commensurate with the

error in the voltage data.

Figures 6.5 and 6.6 show that with fairly small amounts of *unsystematic* noise little improvement is made after the first iteration. This is true whether the error is voltage noise, electrode position error or random perturbations in boundary shape. These observations contrast with the experience of the Sheffield group using the Barber and Brown method which is remarkably stable with respect to variations in geometry. The technique has been used on half planes rather than cylinders without changes in the forward modelling! This phenomenon has yet to be explained satisfactorily. It is worth mentioning in this context that the Barber and Brown method is a linear algorithm and thus measures only conductivity changes, a feature they call *Dynamic Imaging*. From Figure 6.5 it can be seen that the first iteration, which is a linear method is less effected by the data errors than subsequent iterations.

The simulated errors used in this study were plausible but somewhat artificial. A more useful test is to apply the algorithms to real data, either from phantom studies or measurements taken on patients. It seems likely that systematic errors from inaccurate forward modelling may have a less dramatic effect on image quality. This exercise has not yet been performed by the author, although Yorkey [96] has tested a regularised Newton method on data derived from a phantom study. One purpose of constructing the OXPACT machine [69] has been to investigate the real types of error generated by a measurement system. Some analysis of accuracy of the system is given by Furner in [30]. It is already clear that imaging real data will require more accurate forward modelling, including detailed treatment of the electrodes, and much larger finite element meshes. A method for accurately locating the electrodes on the patient must also be found.

6.7.3 The Positivity Constraint

One problem which has not been addressed by previous authors is the fact that the conductivity must be bounded away from zero. As discussed in Chapter 4 the Choleski solver will fail if any conductivity falls below a certain level, as this will cause the system stiffness matrix to be no longer positive definite. Let us denote the smallest feasible nodal conductivity by c . The problem becomes a constrained minimization problem, with an inequality constraint. There are many ways to treat such a problem, one of the most common is the penalty function method in which a penalty function is added to the objective function which becomes very large as the solution approaches the constraint. Another approach, the barrier method, is simply to truncate the conductivity so that negative conductivities are replaced by the minimal value c . A third possibility is to scale the update vector $\delta\gamma$ by a constant $\alpha < 1$ so that $\gamma_i + \alpha \delta\gamma > c$

at each node. Both the barrier and scaling methods were employed and, in general the scaling method found to be superior.

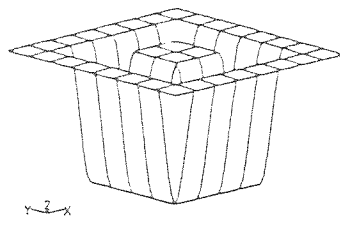
Another possible approach is to change the parameterisation of γ so that negative conductivities are not possible. The most obvious way of achieving this is to define a new parameter $\nu = \ln \gamma$. This method was tested and found to have no particular advantage. The new parameter ν can become very large and negative. When γ is calculated to solve the new forward problem the stiffness matrix may no-longer be numerically positive definite.

The positivity constraint usually becomes active on particularly complicated conductivity distributions. A typical situation is where there are two central maxima close together. The Levenberg-Marquardt update in such a case typically includes the two maxima but has a negative overshoot in the coll between them. If the background level is already low, this negative overshoot is reduced by the scaling method, at a cost of not increasing the maxima sufficiently. The maxima will be increased further at the next iteration of the algorithm.

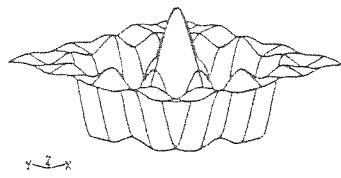
One additional method which in some cases prevents the problem of negative overshoot, is to scale the initial approximation to the conductivity before beginning the iterative technique. An average conductivity can be obtained from $\|j\|/\|v\|$, and the initial approximation to the conductivity scaled to agree with this. As no constant function appears amongst the singular functions, a constant background conductivity will not be found immediately by the method. If a negative overshoot was caused by an the initial conductivity being too low, this can be circumvented by an initial scaling.

It may seem, on first sight, that the use of the logarithmic $\ln \gamma$ may overcome the problem caused by the positivity constraint. In some ways, the use of logarithmic conductivity is more natural, however, it does not solve this particular problem. The bound on the logarithmic conductivity will be $\ln c$, which in practice results in the same problem.

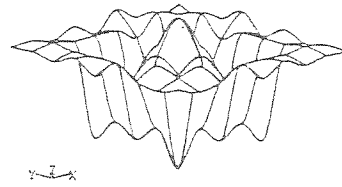
A particularly testing conductivity distribution, in respect of the positivity constraint, is the object funny which is an arbitrarily constructed distribution designed to be difficult. It is illustrated in Figure 6.7. It was reconstructed using 32 electrode trigonometric currents. Without the scaling treatment of the positivity constraint the algorithm was not able to proceed. However, using the scaling method it was successful as is illustrated in Figure 6.7.



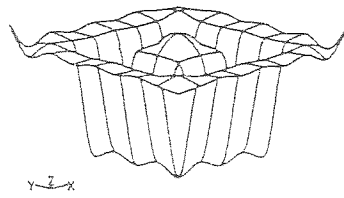
Actual conductivity



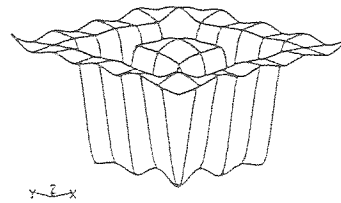
1



2



3



4

Figure 6.4: *The original 'moat' object and the first four iterations using exact data.*

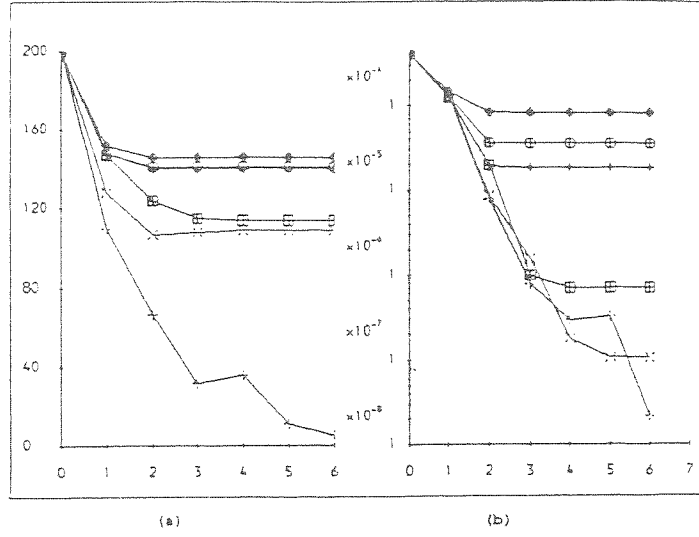


Figure 6.5: (a) The error in the conductivity $\|\gamma_i - \gamma\|$ plotted against iteration number i . (b) The error in the voltage $\|F(\gamma_i) - v\|$ on a logarithmic scale, against i . Exact data +. Error in electrode position: 0.1% of inter electrode distance \times , 0.5% \boxplus . Error in boundary shape 0.1% of inter electrode distance \ast . Noise added to voltage measurements 0.1% \bullet , 0.05% \oplus .

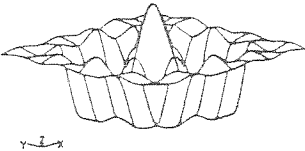
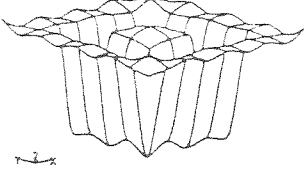
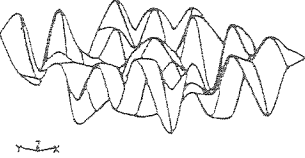
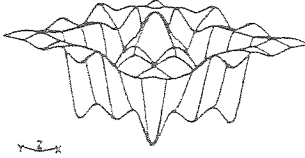
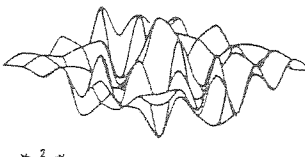

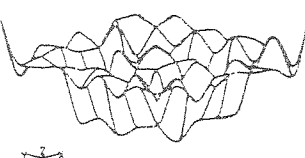
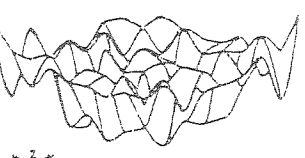
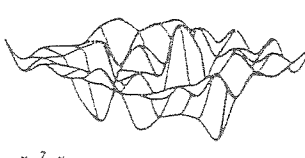
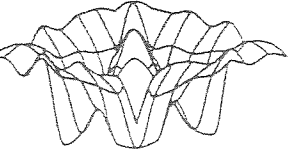


	Iteration 1	Iteration 5
+		
x		
⊕		
+		
⊕		
⊕		

Figure 6.6: The first and fifth iteration of reconstructions of the moat object with the errors used in Figure 6.5

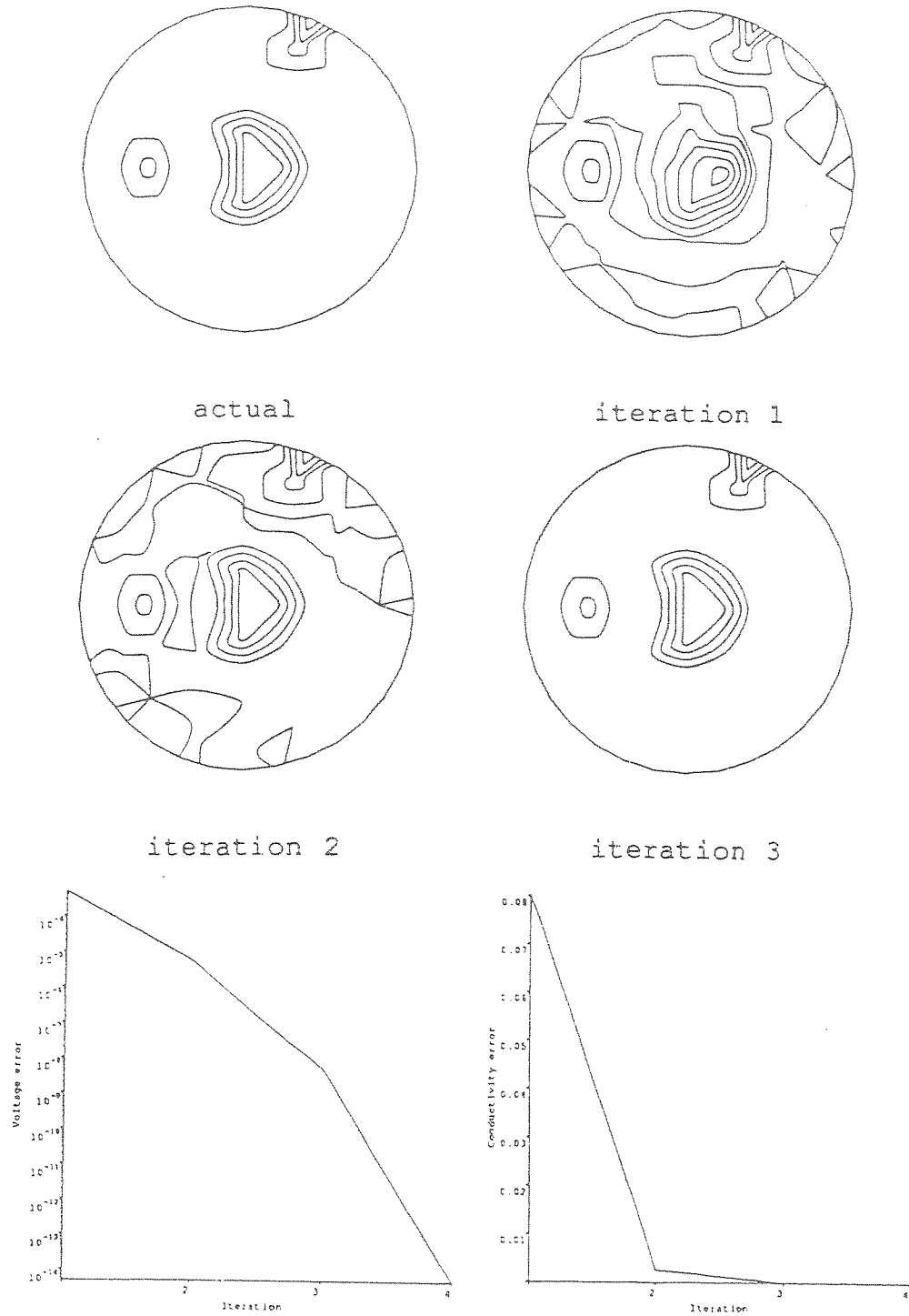


Figure 6.7: *The conductivity distribution funny. Contour plots of actual distribution and first three iterations of reconstruction. Graphs show voltage error (logarithmic scale) against iteration and conductivity error against iteration.*

Chapter 7

Adaptive Methods

Thirty spokes share one hub. Adapt the nothing therein to the purpose in hand and you will have the use of the cart.

Lao Tzu

7.1 Which Measurements to Make?

Thus far in the thesis it has been assumed that the data set has been given to us as a fixed collection of current patterns and voltage measurements. Suppose now that we are in a position to prescribe which current patterns are applied which should we choose? There are various alternatives. For ease of reconstruction we might attempt to apply the Calderón fields or if our best *a priori* guess for γ was not constant, the Sylvester-Uhlmann fields. This would be impossible practically as the current patterns are band limited by the finite number of electrodes.

Another approach would be to attempt to apply trigonometric current patterns. From the data derived from such an experiment one could synthesise an approximation to the data which would have resulted from applying a different current set of current patterns. One could also apply some other orthonormal set of currents, such as a basis of characteristic functions of intervals. Such data would be similar to the data gathered by the Sheffield apparatus [8].

A different approach still would be to optimise the signal to noise ratio of the measurements taken. One such approach, pioneered by Isaacson [47] is described in the next section.

7.2 Two-norm Optimal Currents

Isaacson [47] defines two conductivities γ_1 and γ_2 to be *distinguishable* by measurements of precision ϵ if there is a current density $j \in H^{-1/2}(\partial\Omega)$ for which

$$\delta(j) = \frac{\|R_{\gamma_1} - R_{\gamma_2}\|}{\|j\|} > \epsilon. \quad (7.1)$$

The number δ is called the *distinguishability*. The best currents in the sense of Isaacson are those which maximise $\delta(j)$. For simplicity, R_γ will be considered as the a map $R_\gamma : H^0 \rightarrow H^0$. The definition of $\delta(j)$ can be reformulated as follows:

$$\delta(j) = \sup_{\|j\|_{-1/2}=1} \|R_{\gamma_1} - R_{\gamma_2}\|_{1/2} \quad (7.2)$$

$$= \sup_{\|j\|_{-1/2}=1} \langle j, D^2 j \rangle \quad (7.3)$$

where $D = |R_{\gamma_1} - R_{\gamma_2}|$. (Note that $R_{\gamma_1} - R_{\gamma_2}$ may not be positive definite).¹

The map D is a compact, self adjoint pseudo-differential operator : $H^0 \rightarrow H^0$. It has a complete set of orthonormal eigenfunctions $j_k \in C^\infty(\partial\Omega)$, with eigenvalues $\lambda_1, \lambda_2, \dots$ with $\lambda_k \rightarrow 0$ as $k \rightarrow \infty$. From the min-max principle one can deduce that the largest distinguishability possible is λ_1 which is achieved when j is an eigenfunction with this eigenfunction. These currents are the optimal two-norm currents to apply in the sense that they produce the largest r.m.s. boundary voltage. If the measurement system has a fixed level of noise these currents will give the highest signal to noise ratio.

It is instructive at this point to illustrate this with an example. The concentric anomaly of Section 2.4.2 provides the simplest case. Take γ_1 and γ_2 to be concentric anomalies with conductivity contrasts σ_1 and σ_2 and radii ρ_1 and ρ_2 . The transfer impedance operator for any radially symmetric conductivity on the disc must have the Fourier basis as its eigenfunctions. These will also be the eigenfunctions of $R_{\gamma_1} - R_{\gamma_2}$ and hence of D . The eigenfunctions of D will be

$$\lambda_k = \frac{1}{k} \left| \frac{1 - \mu_1 \rho_1^{2k}}{1 + \mu_1 \rho_1^{2k}} - \frac{1 - \mu_2 \rho_2^{2k}}{1 + \mu_2 \rho_2^{2k}} \right| \quad (7.4)$$

where

$$\mu_\ell = \frac{\sigma_\ell - 1}{\sigma_\ell + 1}, \quad \ell = 1, 2. \quad (7.5)$$

¹Any compact self-adjoint operator A can be represented as $A = U^* \Lambda U$ where Λ is the diagonal operator of eigenvalues and U is a unitary operator. A power of such an operator can then be defined by $A^k = U^* \Lambda^k U$ and the absolute value by $|A| = U^* |\Lambda| U$

As $k \rightarrow \infty$ it is seen that

$$\lambda_k = O\left(\frac{1}{k}\rho^{2k}\right) \quad (7.6)$$

where $\rho = \max(\rho_1, \rho_2)$. The currents which maximise the distinguishability are therefore $\cos \theta$ and $\sin \theta$. From Equation 7.6 the distinguishability of the higher order trigonometric currents $\cos k\theta$ and $\sin k\theta$ decreases exponentially with k . The optimal currents are the best currents to tell γ_1 and γ_2 apart. The higher order trigonometric currents give decreasing information.

For more general conductivities on the disk the eigenfunctions will not necessarily be trigonometric functions but will reflect any asymmetry in the two distributions.

7.3 Algorithms for eigenfunctions

7.3.1 Power Method

Isaacson's algorithm for calculating this optimal current is based on the power method (see for example [48]). We will take γ_1 to be the (unknown) conductivity of the body and γ_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  $\tilde{v}(n) = R_{\gamma_2}j^{(n)}$ ;
    Set  $\lambda_n = \|v(n) - \tilde{v}(n)\|$ 
    Set  $j^{(n+1)} = (v(n) - \tilde{v}(n))/\lambda_n$ 
Until  $\|j^{(n+1)} - j^{(n)}\| < \epsilon$ .
```

Implicitly Isaacson is assuming that the voltage measurement apparatus has a fixed dynamic range, thus all measurements will be made to a fixed absolute accuracy. As an analogy one could think of a voltmeter, reading for example 10mV full scale, graduated in 1mV divisions. The relative accuracy to which smaller voltages are measured is therefore less than that for larger signals. Suppose that instead of using Isaacson's procedure we simply took an arbitrary orthogonal basis of currents, applied these and measured the voltages, resulting in a transfer impedance matrix. This matrix would have each entry accurate to within the same absolute precision, say 1mV. Any numerical technique could then be applied to find the first eigenfunction of this measured

matrix. However these functions may not be eigenfunctions of the actual transfer impedance operator, even to within measurement precision. The arithmetic operations required to compute the eigenfunction involved combining measurements give the possibility of increasing the error. In Isaacson's method the accuracy is maintained by performing a new measurement each time a voltage is required, rather than performing a matrix-vector product.

7.3.2 Higher eigenfunctions

The power method only produces the largest eigenfunction, which would give only one measurement with which to estimate γ_1 . It would be desirable to have a basis of currents at least spanning the same space as the eigenfunctions j_k with $\lambda_k < \epsilon$. For this reason a procedure is needed which computes all the eigen functions j_k with $\lambda_k > \epsilon$. The following algorithm has been found to work in computer simulations with pseudo random noise in the measurement stage.

```

Guess  $j_1^{(0)}, j_2^{(0)}, \dots, j_m^{(0)}$ 
      (an orthonormal basis with  $\int_{\partial\Omega} j_k = 0$ )
Repeat
  Measure  $v_k^{(n)} = R_{\gamma_1} j_k^{(n)}$ , compute  $\tilde{v}_k^{(n)} = R_{\gamma_2} j_k^{(n)}$  for all  $k$ 
  Compute  $r_{lk} = \langle j_k, v_k^{(n)} - \tilde{v}_k^{(n)} \rangle$ 
  Calculate the eigen system for  $R = [r_{lk}]$ ,  $RU = UL$ 
  Set  $j_k^{(n+1)} = SU_{kl} j_l$ 
Until  $\|R - L\| < \epsilon$ 

```

This method is very much in the spirit of Isaacson's. It repeatedly makes new measurements and is 'not satisfied' with the current patterns found until, to within measurement precision, they are eigenfunctions of the transfer impedance operator. However it makes maximum use of the measurements made at each stage by calculating the eigenvectors of the matrix numerically. Any stable numerical algorithm can be used to find the eigenvectors. An iterative technique such as the Jacobi method is particularly appropriate as the eigensystem need only be found to within measurement precision. In numerical experiments with white noise simulated by a pseudo random number generator, little was gained after two iterations of the algorithm. It has yet to be tested on a real system, in which the measurement errors may well have a different structure.

7.3.3 Critique of the Measurement Model

To understand the nature of the errors involved in making impedance measurements it is necessary to understand something of the electronics of an adaptive current impedance tomograph. Currents are generated by constant current sources, one for each electrode. The current level is set using a Digital to Analogue Converter (DAC), which converts a binary number generated by the computer into an analogue current or voltage level. The DAC takes a number with a specified number of bits, typically 12, and produces an analogue signal to within an accuracy of ± 1 bit. For the specific case of a 12-bit DAC the accuracy is one part in 4096. This digital limitation is a base line for the errors in the current generation stage. Ideally the other components of the system would be designed so that any errors they caused would be less than this digital error. However that is difficult to achieve in practice. The current source, for example, has a limited compliance. It will maintain the preset current level to within a certain accuracy only over a certain range of operating loads. If the load, in this case the body, has too high a resistance it will not be able to apply a sufficiently high voltage to achieve the desired current. Other components too will deviate from their specifications. Amplifiers in the system will have a slightly non-linear response, making the transfer impedance, as 'seen' from the computer appear to be a non-linear operator. Different components, supposedly identical, will have slightly different specifications resulting in differences in the response of each current channel.

The voltage measurement stage of an EIT system typically comprises of an instrumentation amplifier, which is a sensitive, stable amplifier, and an Analogue-to-Digital Converter (ADC). The ADC performs the reverse function of a DAC, converting an analogue voltage into a binary number. The accuracy of the ADC has a similar limitation to the DAC, however increased accuracy can be gained at the expense of speed can be obtained by taking the time average of a series of measurements. This can effectively generate far more bits of accurate measurement than the design of the ADC allows. This means that the discretization error is no longer the limiting factor in the accuracy of the system.

While the assumption of fixed absolute accuracy is a reasonable first assumption, such a limitation can be circumvented by more elaborate electronics. The OXPACT system [69], for example, has a programmable gain amplifier in the voltage measurement stage. This is an amplifier with an adjustable amplification factor controlled digitally from the host computer. The analogy with a simple volt meter would be a range selection knob. If the voltage to be measured were tenths of a mV, the range knob is turned so that the meter has a full scale deflection of only 1 mV rather than 10mV. The smaller sig-

nal can then be measured with the same relative accuracy as the larger. The same method could be applied to the current generation stage, so that, within certain limits, small currents could be set with the same relative accuracy as large. With this complication, it is not as clear how to design a measurement algorithm. Nevertheless, the fixed absolute accuracy model is still a useful base-line from which to work.

7.4 Reconstructing With Optimal Currents

Given that an EIT system used is capable of making optimal measurements in the sense of Isaacson, what algorithms should be used to reconstruct the conductivity? It would be possible to use the optimally measured data as the input to the type of regularised Newton's method described in the previous chapter. It is important to remember, however, that the current patterns chosen using Isaacson's criteria are optimal only for distinguishing between two specific conductivities.

Given an initial guess, γ_0 , for the conductivity, current patterns can be chosen to distinguish between γ_0 and the real conductivity γ . Notice at this point that, if there is a known absolute error level ϵ , only a small number of current patterns will be used, those with $\lambda_n > \epsilon$. This would result in an apparently smaller data set than using all possible current patterns with a different configuration. However the situation is similar to the controversy between those advocating adjacent rather than polar drive pairs. The apparently lost information was not actually present at given the level of error in the measurements. Indeed, since these patterns are optimal, current patterns which have components outside the span of the eigenfunctions with eigenvalues $\lambda_n > \epsilon$ contain misleading information.

Now suppose that a linear step has been taken and our conductivity estimate up-dated to γ_1 . This estimate will typically include the larger features of the image. Suppose for example that the real conductivity has a large central maximum and the initial guess is a uniform distribution. The optimal currents will then concentrate current in the central region as far as possible. However these currents will not be ideal for distinguishing between γ_1 and γ . Optimal currents for this purpose will concentrate of the regions of disagreement such as the periphery of the central object which, in γ_1 may not yet be the right shape. This suggests the following adaptive Maquardt method.

```

Given initial guess  $\gamma_0$ 
Repeat
    Calculate optimal currents

```

```

Measure resulting voltages
Calculate derivative matrix  $A$ 
Solve  $(A^T A + \mu I)\delta\gamma = -A^T \delta v$ 
Update:  $\gamma_{i+1} := \gamma_i + \delta\gamma$ 
Until  $\|\delta v\| < \epsilon$ 

```

The parameter μ here is chosen using the same criterion as the non-adaptive methods discussed in Chapter 6.

Although apparently fewer measurements are used at each iteration in this method, the total number of independent current patterns used is much larger as the optimal currents will, in general, be completely different at each iteration. The use of a smaller number of measurements reduces the number of rows of the matrix A . This means that $A^T A$ is rank deficient in exact terms as well as in terms of measurements of a limited precision. However it also indicates that a row-action iterative solution method, may be more efficient. An iterative method could be particularly appropriate as a solution is only required to within measurement precision.

7.5 Numerical Results

Experiments conducted with this algorithm reported in [16] were extremely encouraging. Figure 7.1(a) shows the original conductivity to be reconstructed. A central ‘spike’ was chosen deliberately to emphasis the difference between optimal currents and adjacent pair current patterns. The first iteration with adjacent current patterns produced the updated conductivity shown in Figure 7.1(b), where as the first iteration with optimal currents produced that shown in Figure 7.1(c). After 10 iterations with the adjacent drives the ‘spike’ was recovered to a similar accuracy, whereas the first iteration with optimal currents produced almost as good an image as was possible given that the problem is ill-posed. Further iterations produced little qualitative improvement.

7.6 Point Optimal Currents

Given that Isaacson’s criterion for optimality is not indisputable, what other criteria could be used? One idea is, rather than optimising the norm of the voltages, one could seek a current pattern which optimises one individual measurement of voltage. This would result in a smaller data set than Isaacson’s

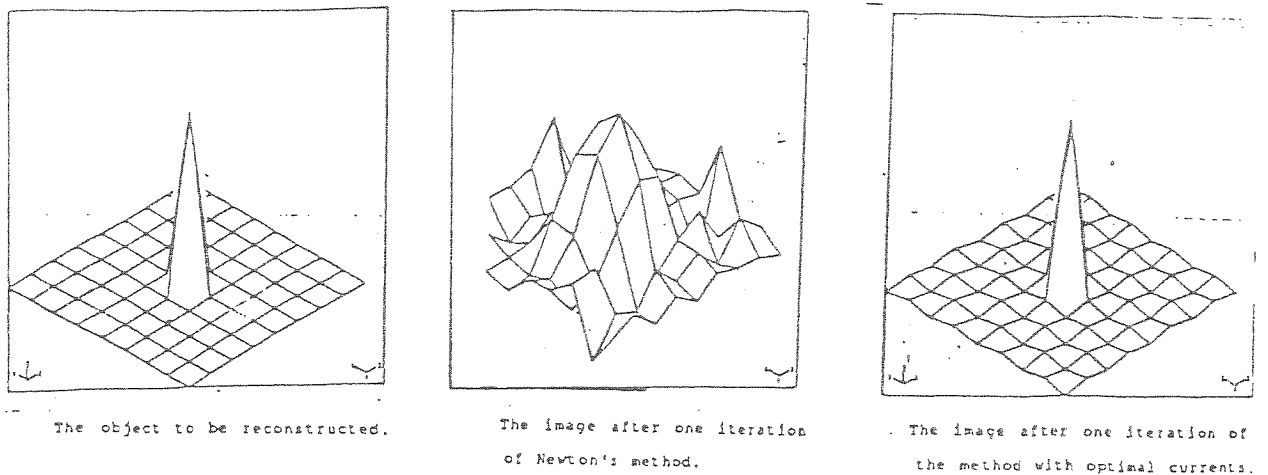


Figure 7.1:

criterion, but as stated in Section 7.4, this need not matter if the currents used are different at each stage in the iteration.

No numerical experiments have yet performed using these *point optimal currents* (which were introduced in [18]) as any advantage they may have would only be apparent with real data. Simulations would only be useful if the structure of the voltage errors were known. Here, a method for their determination is presented.

As in Isaacson's treatment γ_1 and γ_2 are conductivities between which we wish to distinguish. Let $p \in \partial\Omega$ be the point at which we make the measurement $v_1(p) - v_2(p) = (R_{\gamma_1}j)(p) - (R_{\gamma_2}j)(p)$. We seek a j such that

$$\tau(j) = |v_1(p) - v_2(p)| \quad (7.7)$$

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

The current pattern can be expressed in terms of the eigenfunctions as $j = \sum a_k j_k$. The optimisation problem is then

$$\begin{aligned} &\text{maximise} \quad \sum a_k j_k(p) \\ &\text{subject to} \quad \sum a_k^2 = 1. \end{aligned} \quad (7.8)$$

Using Lagrangian procedure we have

$$a_k = C \lambda_k j_k(p) \quad (7.9)$$

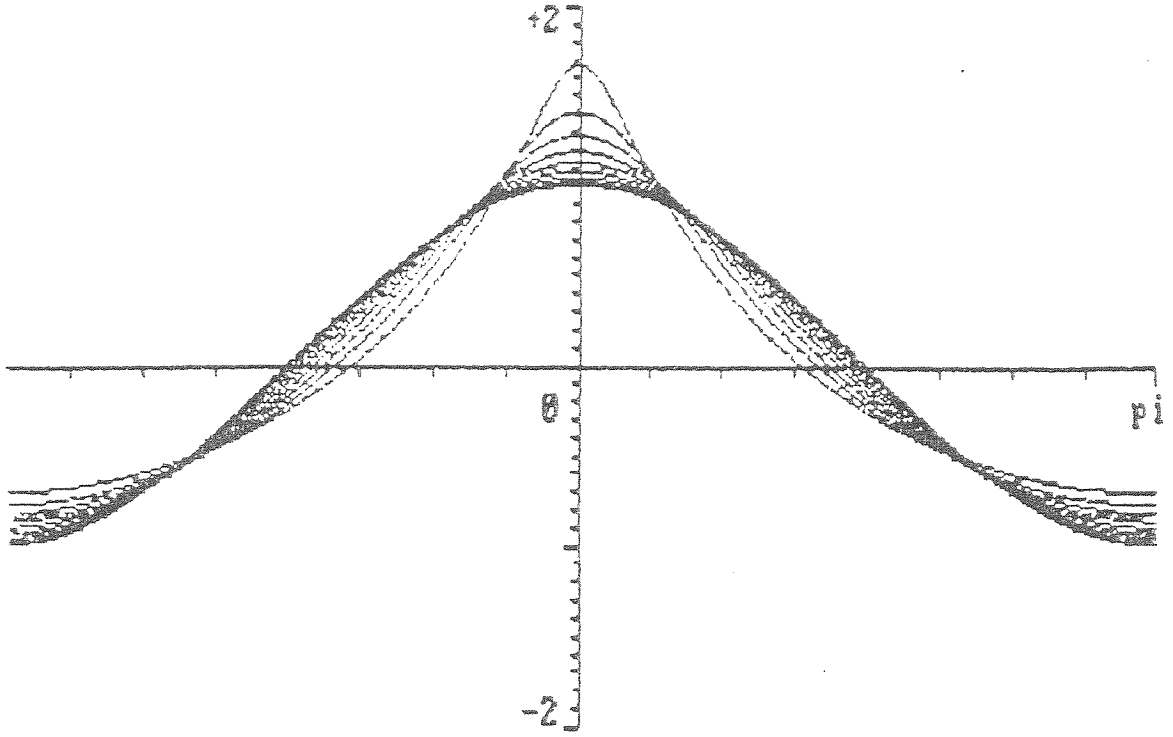


Figure 7.2: Point optimal current patterns for $\sigma = 0.3$ and $r = 0.1$ to 0.9 .

where

$$C = 1/(\sum (\lambda_k j_k(p))^2). \quad (7.10)$$

The measurement procedure would be first to calculate the eigenfunctions as detailed above, and then make the measurement at p when the current

$$j = \sum_k C \lambda_k j_k(p) j_k \quad (7.11)$$

is applied.

In the simple case of a concentric anomaly, these functions can be calculated explicitly as

$$\lambda_k = -\frac{2\mu\rho^{2k}}{k(1 + \mu\rho^{2k})}. \quad (7.12)$$

A plot of point optimal currents for this case in Figure 7.2. It can be seen that for a small object in the centre these currents differ least from the two-norm optimal case.

Chapter 8

Conclusions and Further Work

Work is love made visible.

Kahlil Gibran

8.1 Summary of this Work

The main focus of this thesis has been to show why the Reconstruction problem for EIT is inherently difficult and how the difficulties can be overcome. In Chapter 3 a Taylor series for the forward mapping was derived. This emphasised that the problem is non-linear but also showed how it can be approximated by a linear problem. Chapter 4 presented a technique for modelling the forward problem using finite elements. In Chapter 5 it was demonstrated how both the non-linear and linear problems are ill-posed. The ill-posedness of the linear problem was characterised in terms of a singular value decomposition. In Chapter 6 iterative algorithms were presented which include regularisation to combat the ill-posedness. In Chapter 7 strategies were presented to collect optimal data and it was shown how these techniques may be integrated with iterative reconstruction algorithms.

8.2 The Future of EIT

During the period from October 1985 to January 1990, over which this work was done and this thesis written, enormous progress has been made in EIT. Many new workers have joined the field from a wide spread of disciplines. Mathematicians have been working on many aspects of EIT, medical physicists

have built prototype impedance tomographs, and clinicians have tested the technique for medical diagnosis. However the subject is still young and there is still much work to be done before EIT will be used routinely as a tool for medical diagnosis and research.

The principal areas where further work is vital for progress to be made are as follows.

- An accurate model is needed for the electric field around an electrode on the skin, including contact impedance.
- A method must be found of accurately placing a large number of electrodes on the body which is easy to apply.

The first point must be addressed by Mathematicians and Physicists, the second by Medical Engineers.

In a more theoretical vein, there are still interesting problems to solve which have important implications for EIT. An analytical characterisation of the singular value decomposition of the derivative of the forward problem would greatly increase our understanding of the problem. Also if a direct method were found to calculate the left singular functions and singular values, the solution of the linear step in an iterative reconstruction algorithm could be speeded up significantly.

Another theoretical problem is to find to what extent errors caused by inaccurate knowledge of the boundary shape or electrode position can be separated from voltage differences caused by genuine changes in impedance. A useful result in this area would be as follows. Suppose that the data is changed by a boundary perturbation or electrode position change which is bounded in size by δ . Is there a distance $r(\delta)$ such that no conductivity change within the region $r(\delta)$ from the boundary can explain the change in the data? If the anatomy were known near the surface of the body but an anomaly near the centre was sought, the data which could be explained as conductivity variations near the surface could be used to fit a more accurate model for the boundary shape and electrode positions.

8.3 Some Leads on the Problems

8.3.1 Electrode Modelling

In this work a naïve model of the boundary conditions has been used - that one can specify Neumann conditions using electrodes. This is not exactly the case.

Firstly electrodes have finite area and finite gaps between them. Secondly the electrodes are conductors - which means that the voltage is constant on them. Thirdly there is effectively a highly resistive layer between the electrode and the skin.

This last effect, the contact impedance, has been the cause of some controversy in EIT. It has been argued that the contact impedance cannot be modelled adequately so that measurements of voltage cannot be made at current carrying electrodes. This argument has lead to the adoption of the 'four electrode' measurement technique where voltage measurements are made between electrodes other than those carrying current. Advocates of this position [8,61] argue that contact impedance can be represented as an unknown, non-interacting resistor in series with the electrode. If current passes through this resistor there will be an unknown voltage drop and the measurement of voltage will be useless.

The other school of thought, originating at Rennselaer Polytechnic Institute (see [38] and [24]), proposes the following model for contact impedance. The ℓ -th electrode E_ℓ has an 'infinitesimally thin' layer with *effective contact impedance* ζ_ℓ (whose units are Ωm^{-2}). The voltage V_ℓ on the electrode is therefore the voltage u minus the voltage drop across the contact impedance layer

$$V_\ell = u + \zeta_\ell \gamma \nabla_n u \quad \text{on } E_\ell \quad (8.1)$$

As the electrode is conducting V_ℓ is constant. Equation 8.1 together with the the knowledge of the total current flowing through each electrode,

$$I_\ell = \int_{E_\ell} \gamma \nabla_n u, \quad (8.2)$$

gives a set of boundary conditions for u . It is not clear if these boundary conditions result in a well posed boundary value problem as they do not fall conveniently into the usual framework. However for the case of the uniform disk they can be solved by expressing u as a series in the usual way:

$$u = \sum_{k=0}^{\infty} r^k (A_k \cos k\theta + B_k \sin k\theta). \quad (8.3)$$

Equations 8.2 and 8.1 then provide linear equations relating the V_ℓ , A_k and B_k to the known I_ℓ which can be solved numerically. Cheng [24] compared data measured from a two dimensional, circular, saline solution phantom tank with this model. He found an excellent correlation after fitting one parameter, the effective contact impedance, which was assumed to be the same for each electrode. Other models, including the non-interacting resistor model, did not fit as well. This would indicate, at least for saline solution phantoms, that the model is correct.

8.3.2 Electrode Placement and Boundary Shape Errors

The problem of accurately placing large numbers of electrodes must be overcome by ingenious engineering. One possibility is to design a belt which fits the body with equally spaced electrodes. This has the disadvantage that the shape of the belt is not known. Another approach would be to have a rigid array of electrodes connected electrically to the patient via a compressible conducting material of similar impedance to the tissues. This layer between patient and electrodes would be imaged just as the air surrounding the patient is imaged in X-Ray CT. This configuration would allow a very large number of electrodes, which would offset the effort wasted in reconstructing the surrounding medium. This would also alleviate the the problem of an unknown boundary shape.

An alternative solution would be to determine the position of the electrodes and the body shape using additional sensors. This could be done using mechanical position transducers or optical range-finding, for example.

8.4 Final Remarks

Progress continues in EIT and an increasing number of research groups are working in the area but formidable challenges still remain. The original programme of work for this thesis has been in some sense over-ambitious. Its scope has ranged from the abstractions of Sobolev spaces to the practicalities of circuits and saline solution. The author's original aim was to understand impedance tomography mathematically and translate this understanding into a practical system. Needless to say neither of these have been achieved. Indeed there will be many more doctoral theses written on the subject long before these aims are met. The author hopes that other workers in the field will find the work presented in this thesis useful in their own investigations.

Appendix A

Design of an Adaptive Current Tomograph

A.1 Introduction

This appendix describes the design, construction and initial testing of the Oxford Polytechnic Adaptive Current Tomograph (OXPACT). The system was initially designed by Dr D. Murphy in 1987 in conjunction with the current author. The system was developed and built by Dr C. McLeod, Dr F.J. Lidgley, Mr S. Owen, Mr T. Davey-Winter and Mr P. Furner, again, together with the current author.

The idea of an Adaptive Current Tomograph stems from Isaacson, Newell and Gisser [47,?] at Rensselaer Polytechnic Institute in the USA. This method employs a single-ended current source on each electrode and allows for measurement of voltage also at that electrode. The current pattern can be adjusted under software control to optimise the distinguishability of the conductivity distribution of the object.

Following successful reconstructions of computer simulated data using an adaptive current algorithm (see Chapter 7) it was decided to build a prototype system to test these methods in practice. In particular it was hoped that greater insight would be gained into the type of errors inherent in measurements from an adaptive current tomograph, and the accuracy of the forward modelling techniques used could be tested and, if necessary, improved.

The specification of the system was that currents should be set, and voltages measured to within an accuracy of 0.1%. The system was to be designed to meet medical safety standards. This dictated the limits of current and voltage levels which could be applied, and the frequency of the applied current.

It also meant that the instrument needed to incorporate isolation barriers to prevent dangerous voltages being applied to the patient in the event of a malfunction. Although it was not anticipated that this version of the system would be used in a clinical setting, it was necessary to design the system to these safety standards so that an accurate assessment of errors could be obtained.

It was decided at the design stage that the instrument would be capable of measuring both the in-phase and quadrature components of the signal. If the alternating voltage is represented by a complex value, this arrangement allows both real and imaginary components to be measured separately. This allows the future study of objects which have a non-zero reactive component in their impedance.

A.2 System Overview

The layout of the overall system is illustrated in Figure 9.1. The OXPACT system is housed in a single instrument case. The instrument is controlled by a host computer, which can set current levels and operate the measurement system under software control. The host computer is a standard PC-compatible fitted with a parallel interface card connected to the 22 digital lines used to drive the instrument. A 12-bit analogue-to-digital converter card is also fitted in the host computer which is connected to the instrument's analogue outputs.

The digital output lines are divided into address, data and control lines. These enable independent setting of each current source, effect measurement from any electrode pair, and set the programmable gain stage in the measurement system.

Internally the instrument comprises a motherboard and four electrode interface boards for a 32 electrode system (see Figure 9.2). The motherboard, which is directly controlled from the host, has the following features:

- Optical isolation on all digital input lines from the host.
- Address decoding for selection of each electrode-interface board.
- Selection of differential voltage signal from electrode-interface board.
- A Wien-bridge oscillator.
- Two phase-sensitive demodulators for in-phase and quadrature measurement.
- Programmable gain control for the voltage measurement.

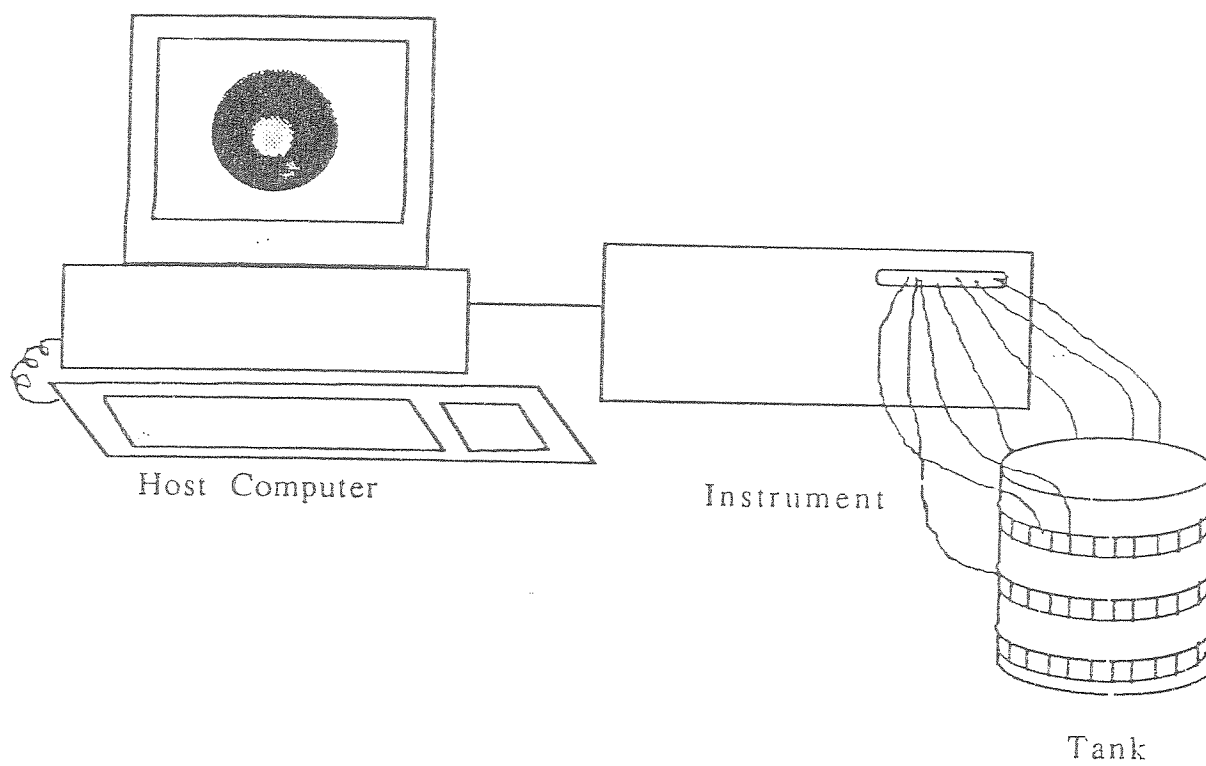


Figure A.1: Overall layout of OXPACT system

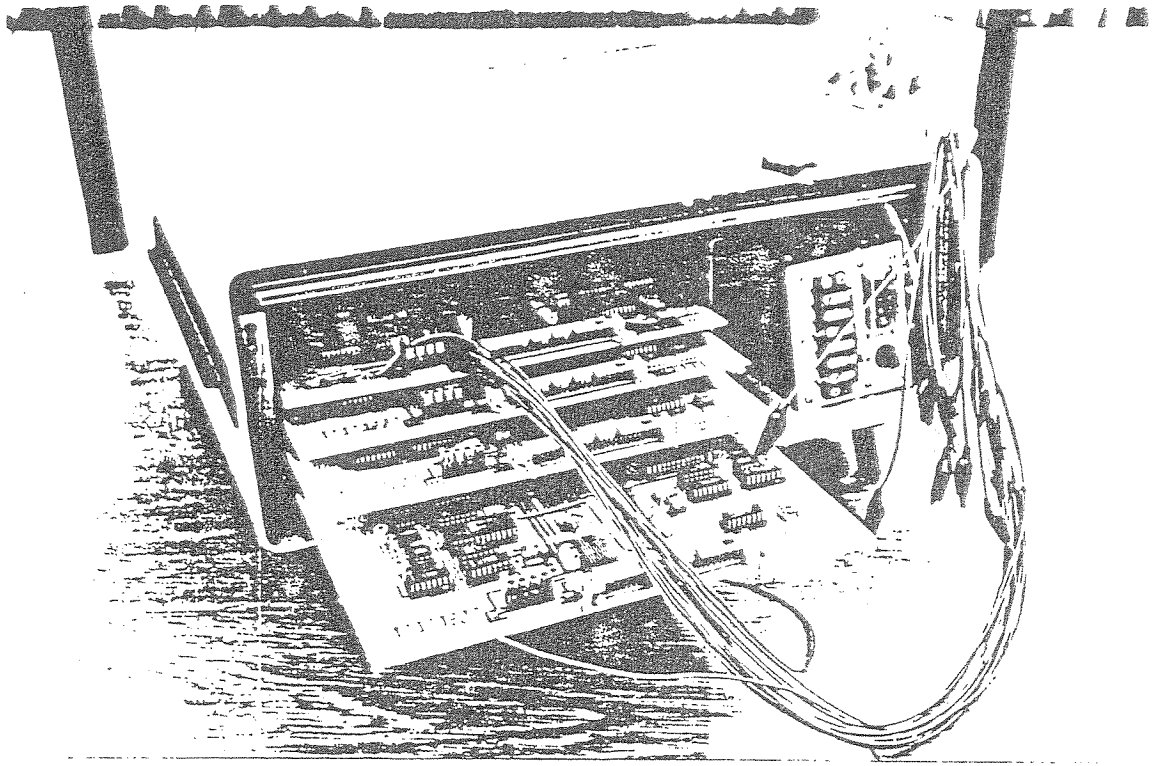


Figure A.2: The OXPACT system with front panel removed showing four electrode interface board, each with eight screened leads terminated with crocodile clips

Each electrode interface board consists of the following:

- A programmable current source for each of eight electrode lines.
- Voltage measurement buffer and screen driver on each electrode line.
- Multiplexing for selecting voltage measurement from any one or two electrodes.

The number of electrodes may be increased in increments of eight, up to a limit of 96, by adding electrode-interface boards.

A.3 Circuit Details

A.3.1 Motherboard

The circuit diagram for the motherboard is shown in Figure 9.3. The digital lines from the host computer are passed via a 26-way IDC connector. The usage of each line is detailed in Table 9.1. Each digital line is optically

isolated to provide isolation between the host computer and the instrument. This is achieved via IC's 1-6 which are quad opto-isolators. A 4-to-16 line decoder, IC 7, provides a mechanism for selecting one of twelve electrode interface boards or selecting the multiplexer latch, via address lines A0-A3. Communication with the electrode interface board is achieved with a 50-way IDC connector. This is used to pass the isolated bus signals and decoded board-select lines to each electrode interface board.

The motherboard is capable of making one differential measurement between any pair of channels. The measurement pair selected is held on IC 8 which is an 8-bit latch (74LS374). Four bits of the output go to each 16-1 multiplexer (IC 9 and 10). Each of the 16-channel multiplexers, IC's 9 and 10 (HI-506A) switches the signal from the selected electrode to one side of the differential instrumentation amplifier IC 11 (AD524).

Each of electrode-interface boards has its own pair of eight-channel multiplexers. Figure 9.4 illustrates the overall multiplexing scheme which is designed to:

- enable selection of any pair of electrodes, in either order,
- enable single-ended measurement from any electrode,
- be expandable by plugging in extra boards,
- minimise channel resistance and output capacitance, to reduce common-mode errors,
- minimise feedthrough and maximise isolation,

The instrumentation amplifier has a pin-programmable gain which can be set to 1, 10, 100 or 1000 using a DIL switch. The output of the amplifier, a single-ended high frequency voltage, is fed to an isolation transformer, T1. (This meets the safety standards set out in BS5724). The output of the transformer is connected to a cascade connection of two programmable gain amplifiers (IC 12 and 13 — AD526). The gain of each amplifier is set digitally, using the unisolated digital lines, to a value of 1,2,4,8, or 16, with a non-linearity of better than 0.05%. The total gain, therefore, is any value of the form 2^r , where $0 \leq r \leq 8$. This programmable gain selection is incorporated for two reasons: signal levels obtained from phantoms where not known *a priori* and the voltages measured will have a large dynamic range.

The output from the programmable gain amplifier passes to the phase-sensitive detector circuit. The AD630K is a switching type demodulator which is employed as IC 14 and 15. The carrier for the demodulation comes from

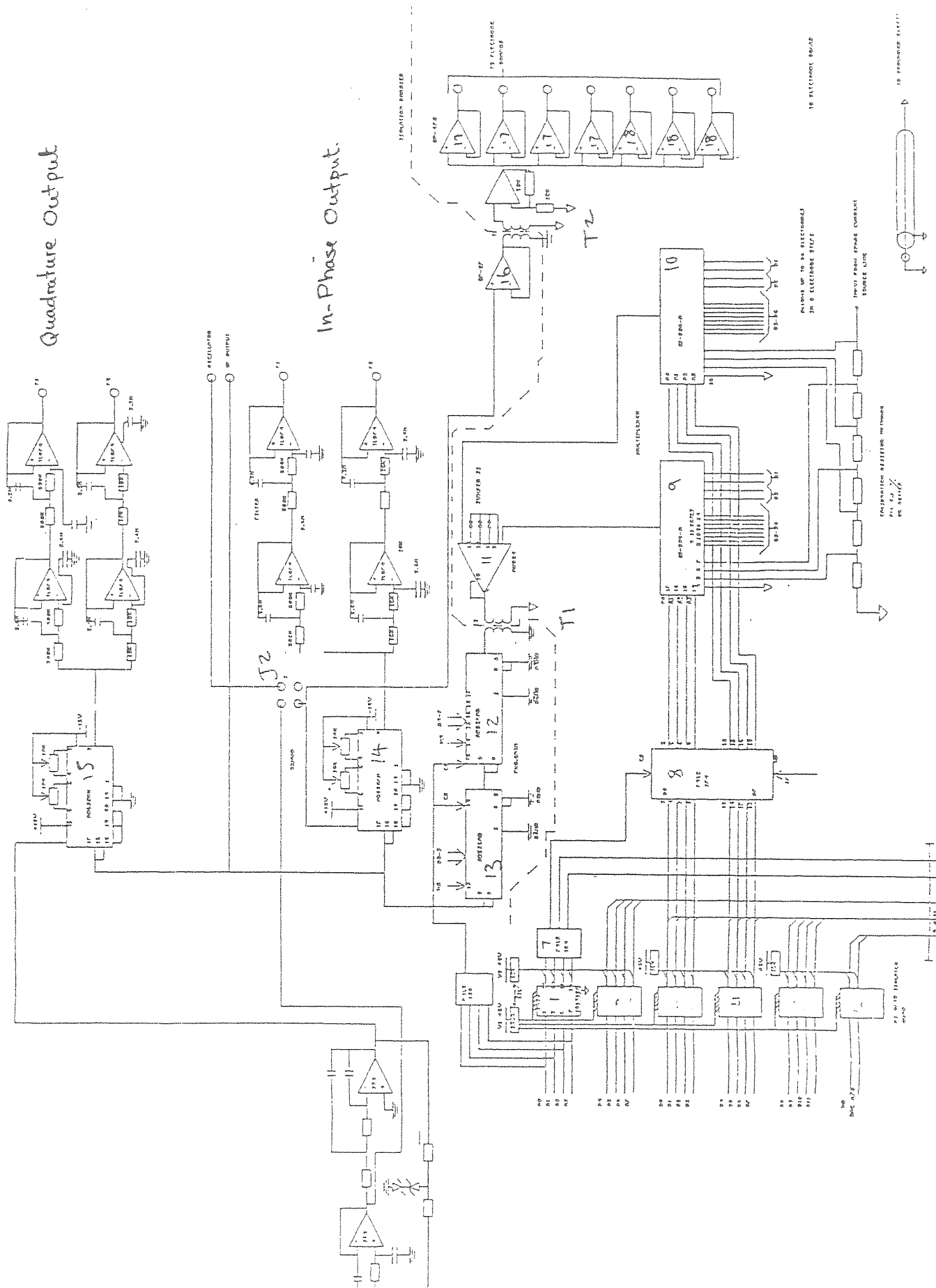
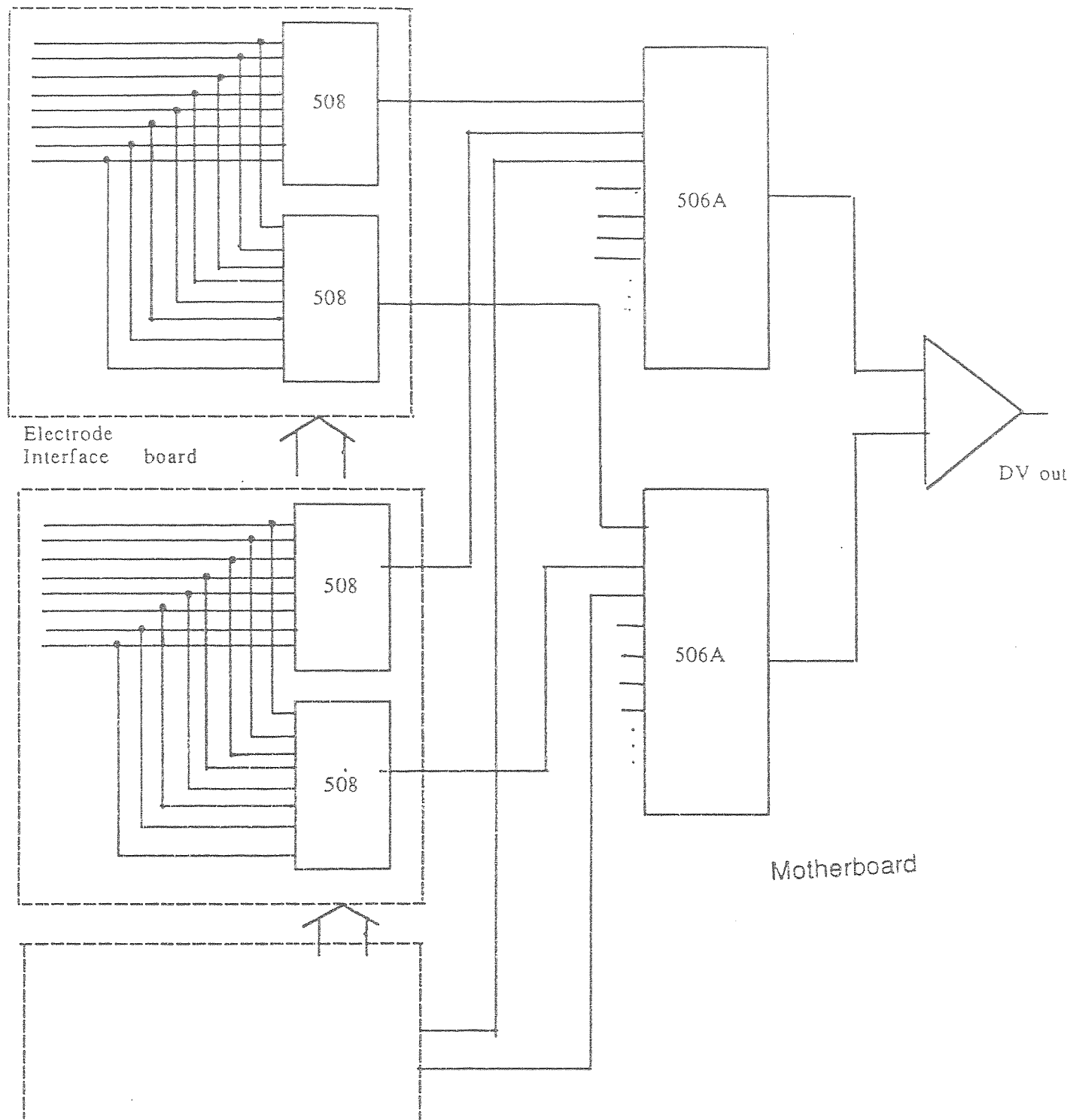


Figure A.3: Circuit diagram for motherboard

PW #	Description	PORT INTERFACE		FUNCTION		Mother-board		Electrode board		Mother-board
		Bit Signif	Byte/Port Assignment	Rule 1	Rule 2	Rule 3	Mother-board Assignment	Opto- Isolator P.W. IN/OUTS	Opto- Isolator P.W. IN/OUTS	
24	NC	LSB	A0				NC			
23			A1							
22			A2							
21	WRITE DATA	MSB	A3	WR			WR	15	20 (WR)	
20	DATA LINES	LSB	A4				D11	7	17 (D00)	3 (2EN)
19			A5				D10	6	16 (D01)	4 (1EN)
18			A6				D9	3	14 (D02)	7 (2AZ)
17		MSB	A7				D8	2	14 (D03)	8 (2A1)
		LSB	B0				D7	7	10	
16	DATA LINES		B1				D6	6	11	13 (2A0)
15			B2				D5	3	14	14 (1A2)
14			B3				D4	2	15	17 (1A1)
13		MSB	B3						10 (D07)	15 (2A2)
		LSB	B4							18 (1A0)
12	DATA LINES		B5						9 (D08)	17 (1A0)
11			B6						8 (D09)	14 (1A1)
10			B7						7 (D00)	15 (1A2)
9		MSB	B7						6 (D01)	17 (1A3)
		LSB	C0							
8	SELECT DAC A/B	LSB	C1	DN			A7	7	10	
7	ADDRESS LINES		C2				A6	6	11	
6			C3				A5	3	14	
5		MSB	C3				A4	2	15	
4		LSB	C4							
3			C5							
2	ADDRESS LINES		C6							
1		MSB	C7							
		LSB	C4				A3	7	10	
			C5				A2	6	11	
			C6				A1	3	14	
		MSB	C7				A0	2	15	

Table A.1: The usage of the digital lines which communicate between the host computer and motherboard (from Furner [30])



Four such boards with eight more possible

Figure A.4: The overall multiplexing scheme for voltage measurement.

the oscillator. The output of each demodulator is low-pass filtered to remove the high frequency harmonics produced by this type of demodulation.

The oscillator employed of the Wein-bridge type. The frequency was fixed at 6.5kHz. Both in phase and quadrature outputs are produced which are fed to the demodulators. The in phase output is buffered and fed to the electrode interface boards via screened cables. An external oscillator was also available and frequencies from 5 to 20 kHz were used for testing.

A.3.2 Electrode-Interface Board

Each electrode-interface board supports eight electrodes. The circuit for each pair of electrodes is shown in Figure 9.5. These blocks are connected together as shown in Figure 9.6. Each board has is connected to the $\pm 15V, 0V$ power supply, the 50-way ribbon cable to the motherboard, a screened oscillator input from the motherboard, two screened measurement lines connected to the motherboard and eight electrode lines terminated in alligator clips.

The control functions required of the electrode-interface board are to select a current source and set its output level and to select one or two electrodes on the board for voltage measurement. Each electrode-interface board is identical apart from a jumper set to designate its board number (Figure 9.6). When the board-select line is activated from the mother-board, the 3-bit decoder (74LS138) is activated. This decodes the address lines A4–A6 for selection of devices on the board. Five of the eight possible outputs from this device are used. The first (000) is used to drive an 8-bit latch (74LS374). The others are used for the chip-select lines for each of the four dual DAC's used in the programmable current sources.

As described in Section 9.3.1 each electrode-interface board has its own pair of eight-channel multiplexers to select the measurement electrode pair. The 8-bit latch (74LS374) is used to latch the data lines for these multiplexers. The two least-significant bits are used to switch on and off each multiplexer. The two pairs of three bits remaining are used to control the channel selection on each multiplexer.

The programmable current source (Figure 9.5) can be thought of as two parts. The first is a Multiplying Digital-to-Analogue Converter (MDAC) capable of multiplying a high frequency signal by a signed 11-bit number. The second stage is a voltage to current converter (or transconductance amplifier) which translates a positive or negative voltage level into a proportional current.

The MDAC is realised using a standard 12-bit DAC, IC 1, (DAC 8212GP) which has a single ended output together with a standard summer network

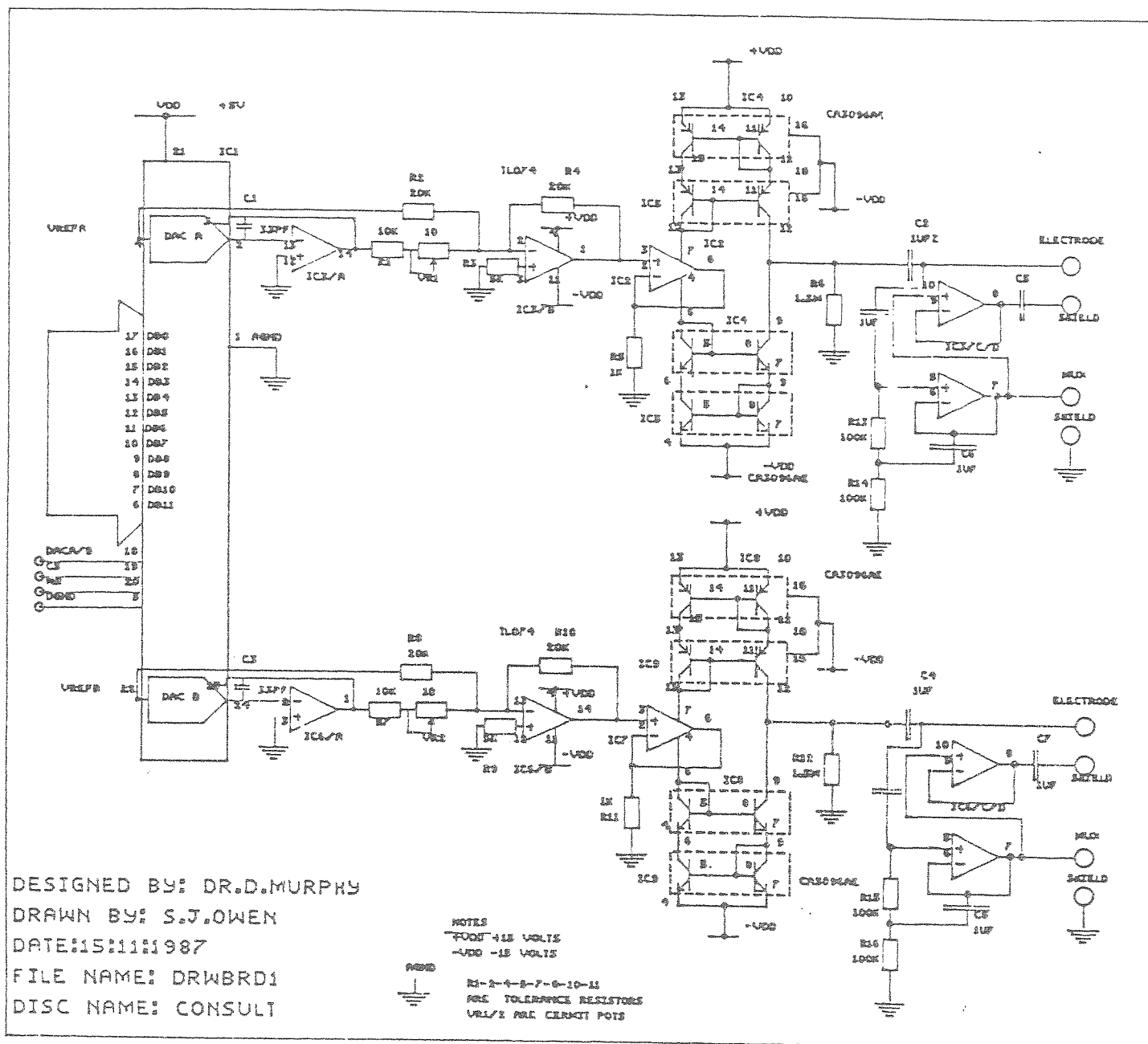


Figure A.5: The circuit for each electrode pair, including current source, measurement buffer and screen driver.

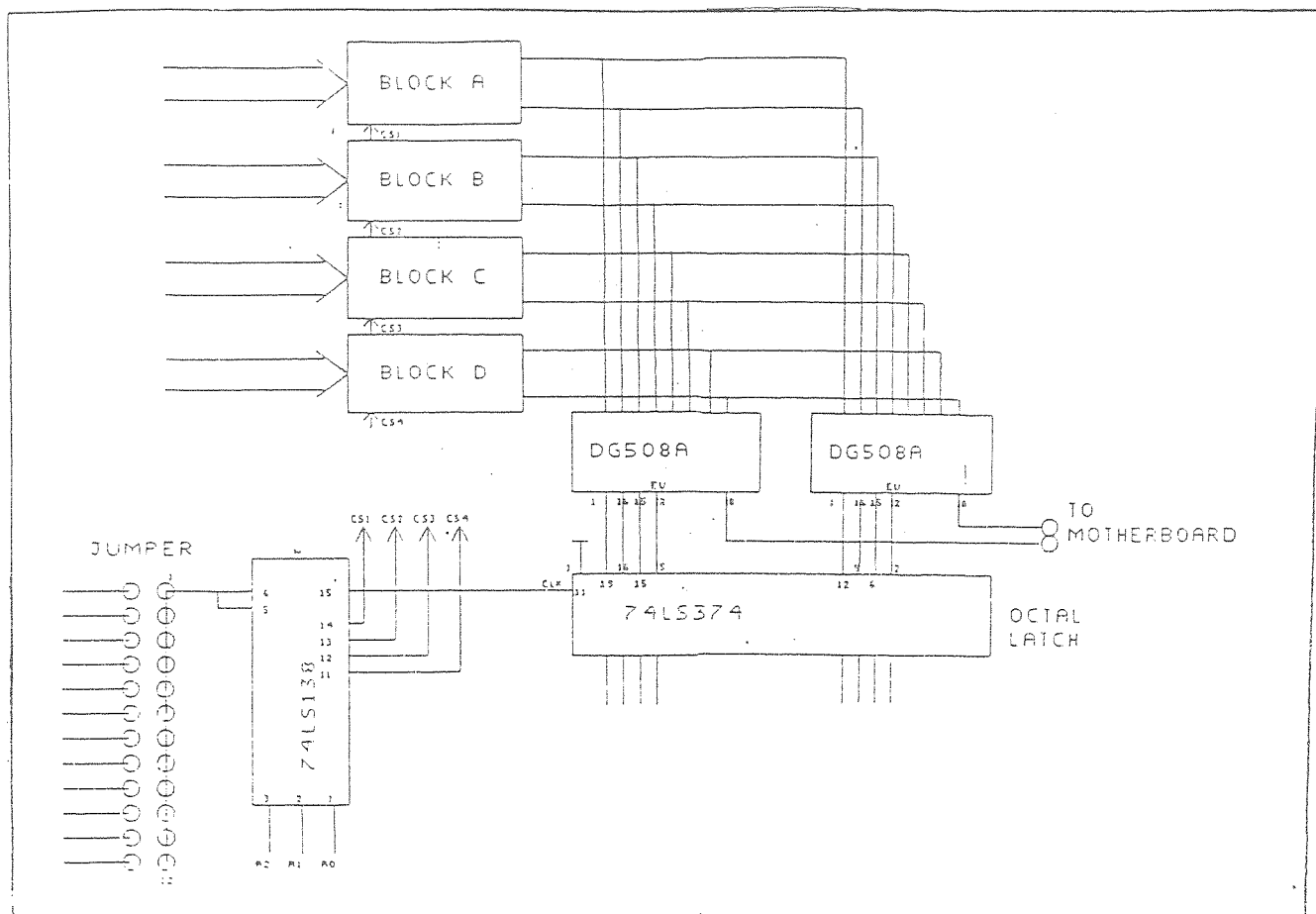


Figure A.6: The overall arrangement of the electrode-interface board

(IC 5 and 6) which allows for inversion of the output. The summer network uses precision matched resistors. Each IC 1 has a pair of DAC's each controlling a separate current source.

The current source circuit used was an ingenious design due to Lidgey [62]. It has the particular advantage of not requiring matched or accurate resistors. It would be particularly suitable for incorporation in a monolithic electrode-interface circuit at a later stage. The operation of the circuit is as follows (Figure 9.7). The operational amplifier, IC 2, is connected as a unity gain amplifier. If the voltage on the non-inverting terminal of the op-amp is V_{in} then a current of V_{in}/R_1 flows to ground through R_1 . This current must be drawn from the supply lines (pins 4 and 7). The novel element of this circuit is that the supply lines of the op-amp are connected to Wilson current mirrors (IC 4 and 5) which reflect the current drawn in the output. The polarity of V_{in} relative to ground will determine the direction of the current, and will determine which of the current mirrors is active and hence the direction of the output current. The resistor, R_2 , allows any DC component of the current to go to ground, while C_2 allows only the high frequency component to enter the patient. It was found that, if the output lines were unloaded while a non-zero current was set, the output transistors of the current mirrors would break down. This was prevented by fitting the Zener diodes shown, which had a reverse breakdown voltage just lower than the supply voltage used.

The final part of the circuit for each electrode is the voltage buffer and screen driver. The measurement buffer is an AC follower configuration giving maximum input impedance at high frequencies. A separate op-amp is used to drive the screen on the electrode line with a voltage following that carried by the central conductor, this minimises the effect of stray capacitance.

A.4 Software Drivers

Software drivers were written by the author in conjunction with Dr. C. McLeod. The drivers were written in Turbo Pascal, which provides a convenient development environment and easy access to low level facilities such as input-output ports. The interface was designed so that the low level mechanisms were hidden from the calling program. Three routines are used to interact with the tomograph: SetCurrent, MeasureVoltage and SetGain.

The electrodes are indexed via an integer subrange:

```
ElectrodeIndex : 0..95;
```

Current values are in the range MinI..MaxI, where MinI is the constant value

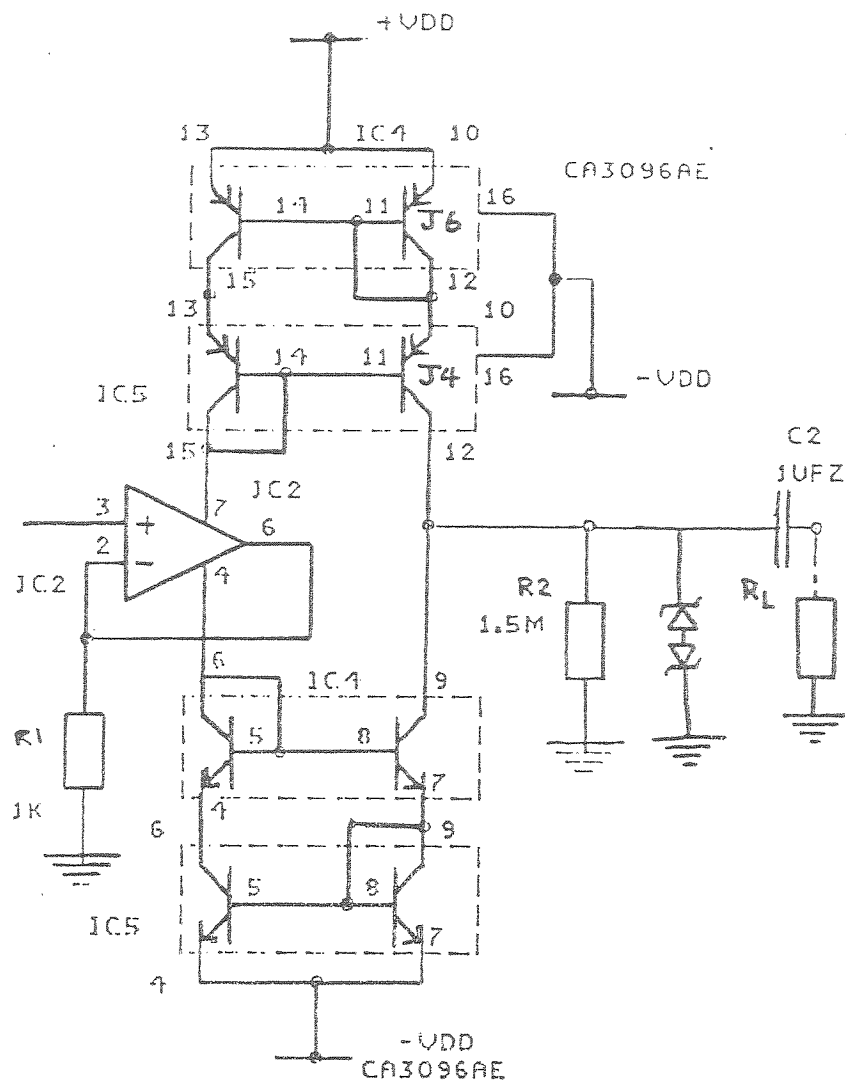


Figure A.7: The Lidgey current source

0 and MaxI is $4095 = 2^{12} - 1$. To set the current source E1 to the integer value Val one makes the procedure call

```
SetCurrent(E1,Val);
```

If all 32 channels are used, each connected to an electrode on the phantom tank, care must be taken that the sum of the currents set is zero. If this is not the case, and there is no additional earth connection on the tank, the operation of the current sources will break down and the additional current will be drawn from, or sunk into, the current sources despite their setting. If an extra earth electrode is used, any additional current will flow from or to earth via this route. If the current sources are properly matched there will be no current in the earth electrode. If the apparatus were connected to a patient care must be taken that part way through the setting of the full current pattern, an unacceptable current is not entering the patient via the earth electrode. The programmable gain amplifier can be set to gain 2^{r+s} by the statement:

```
SetGain(r,s);
```

A voltage measurement can now be made using the function call

```
v := MeasureVoltage(E1,E2);
```

which measures the voltage between electrodes E1 and E2. In fact it takes an average over ten measurements as this reduces the error in the measurement due to white noise.

Hidden in these routines are all the complexities of addressing the various devices in the instrument. No details will be give here but the complexities include allowing for the delays caused by the opto-isolators and the sending of write pulses to certain devices. Now these routines are written the programmer no longer has to allow for these complexities.

A.5 Problems with the Design

A very large number of problems have been encountered with the design, some were solved in this version, others will have to await the design of the next version. Some of the main problems will be outlined in this section. The first problem encountered in writing the driving software was the complexity of the addressing scheme. This complexity is due largely to the isolation barrier between the instrument and the computer. The instrument was connected

directly to the computer's bus this problem would be overcome — the digital devices in the instrument could be memory-mapped and addressing would be easy. One method of achieving this, without compromising patient safety, is to use a battery powered computer which is not connected to the mains. Such a computer has been purchased and its use is being investigated.

Another problem associated with connection to the computer is the presence of digital noise. To overcome this two completely separate power-supply units were used. One to power the parts of the instrument within the isolation barrier, the other to power those outside. This significantly improved the signals and eliminated the digital noise. It would, in any case, be necessary to have this dual power-supply arrangement for patient safety.

Apart from the modification mentioned in Section 9.3.2 of fitting Zener diodes, another problem was found with the programmable current source circuit. A phase-shift was found to occur in the op-amps IC 5/A and IC 6/A on the outputs of the DACs. This meant that the summer network did not work correctly as the reference signal (in R 2 and R 8) was not phase shifted. A suitable phase shifting network was introduced to compensate for this. It would be advantageous to eliminate the need for this summer network for this reason as well as its dependence on precision resistors. A alternative would be to use a DAC with a two-ended output.

A useful addition would be a programmable gain stage for the signal supplied to the current sources. At present the maximum current level is set with a potentiometer. This would also, at a later stage, avoid the problem mentioned above of a large current flowing through the earth electrode when only part of a current pattern is set. The DACs could be set and then the signal level raised to the desired level.

A final problem with the system is that, since its design, it has grown in a rather disorganised fashion as it was modified. Its appearance is rather baroque with flying leads and added-on circuit boards. It is now rather unmanageable and unreliable. The time to design and build OXPACT Mark 2 is long overdue.

A.6 Calibration and Testing

The voltage measurement stage of the instrument is common to all channels, this makes calibration of the measurement stage straightforward. It was found to be linear over the operating range to within the 12-bit measurement precision available. For the current sources calibration was more involved. Each current source was slightly mis-matched and a calibration procedure had to be devised.

A precision resistor was used as the load, each electrode lead was connected to this in turn. A testing program was written which applied a current ranging from MinI to MaxI in 32 steps. The voltage between the precision resistor and ground was measured by the measurement stage and linear regression was used to fit a straight line to the measurements. The intercept of the line indicated an error in setting the current zero (due, for example to, errors in resistor values in the summer network). The slope indicated the gain of the transconductance amplifier. Most channels had slopes agreeing to within 1%, however two clearly had a fault condition exhibiting an error of up to 10%. The 1% error is worse than the design performance, but once the calibration procedure has been done, that error can be removed in software. The fault on the other channels has not been found at the time of writing.

As not all channels were operating correctly, it was decided to make measurements on phantoms using only 16 channels. A tank of radius 14 cm and depth of 5 cm was constructed. Electrodes were made of stainless steel strip. The gaps between the electrodes were chosen to be the same size as the electrodes. The rectangular shape of the electrodes gives the tank symmetry in the z-direction and can thus be treated as a two-dimensional problem. The tank was filled with a solution made with Sodium Chloride and tap water. Sixteen working channels from the tomograph were connected to every other electrode. Trigonometric current patterns were applied to the tank and the voltages measured. The results are shown in Figure 9.8.

Measurements were also taken when cylindrical insulating targets were placed in the tank. The measurements showed that a voltage change was detectable for a 3 cm object in the centre of the tank.

The objects have not yet been successfully reconstructed.

A.7 Future Work

Work on reconstructing images from the measured data is in progress. It is hampered by the unreliability of the instrument but success is anticipated soon. The experimental program for the tomograph includes verification of the forward model by an analysis of the spectrum of the measured transfer-impedance matrix. The algorithm for finding the eigen-system of the transfer impedance matrix suggested in Chapter 8 will also be implemented and tested.

Image reconstruction algorithms have so far been implemented on a Sun 386i workstation, and on a transputer board fitted in the workstation. To avoid interfacing problems the host PC computer will be fitted with an ethernet card to facilitate easy data interchange. Measurement software can then run on the

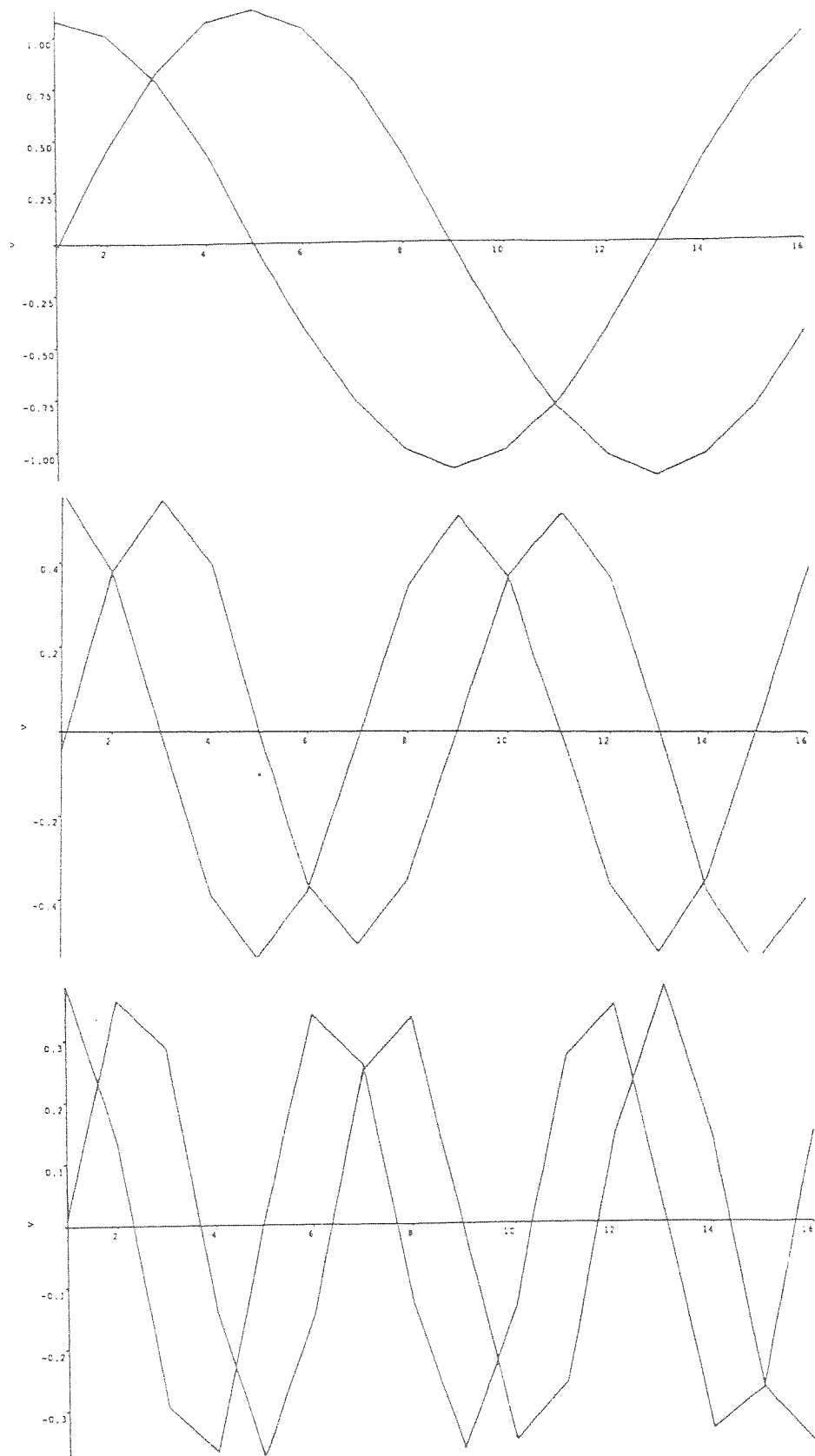


Figure A.8: The voltages measured when the trigonometric currents $\cos k\theta$ and $\sin k\theta$ were applied to the saline filled tank

host computer while the workstation reconstructs images.

The design and construction of OXPACT Mark 2 will commence shortly incorporating the lessons learnt from the construction of this prototype. It will have at least 96 channels so that fully three-dimensional problems can be investigated.

References

- [1] R. Abraham, J.E. Marsden and T. Ratau, Manifolds, Tensor Analysis, and Applications, Addison-Wesley, 1983.
- [2] G. Allesandrini, Stable determination of conductivity by boundar measurements, *Applicable Analysis*, 27, pp153–172, 1988
- [3] H.H. Barrett, Fundamentals of the Radon Transform, in *Mathematics and Computer Science in Medical Imaging*, eds M.A. Viergever and A Todd-Pokropek, NATO ASI Series, Vol F39, Springer-Verlag Berlin Heidelberg, pp105–141, 1988.
- [4] R.H.T. Bates, G.C. Mckinnon, and A.D. Seger, A Limitation on Systems for Imaging Electrical Conductivity Distributions, *IEEE Trans Biomed Eng*, BME-27 ,7, pp418–420, 1980.
- [5] D.C. Barber and B.H. Brown, Rescent developments in Applied Potential Tomography — APT, in *Information Processing in Medical Imaging*, ed. S.L. Bacharach, Martinus Nijhoff, pp106–121, 1986
- [6] D.C. Barber and B.H. Brown , *Applied Potential Tomography*, J.Phys. E: Sci Instrum, Vol 17, pp723–7333, 1984.
- [7] B.H. Brown, D.C. Barber and A.D. Seagar, Applied potential tomography: possible clinical applications, *Clin Phys Physiol Meas*, Vol 6, No2, pp109–121,1985.
- [8] B.H. Brown and A.D. Seagar, The Sheffield Data Collection System, *Clin. Phys. Physiol. Meas.*, 8 Suppl A, 91–98, 1987.
- [9] G. Beylkin, The Inversion Problem and Applications of the Generalised Radon Transform, *Comm Pure Appl Math*, V37, pp579–599,1984.
- [10] C. Bischof, Computing the Singular Value Decomposition on a Distributed System of Vector Processors, Tech. Rept. 87–869, Dept. Computer Sci., Cornell Univ., NY, 1987.

- [11] D.C.Barber and B.H.Brown, Applied potential tomography, J.Physics E, 17, pp723-733, 1984.
- [12] W.R. Breckon, Impedance Imaging: a preliminary report on the reconstruction problem, Oxford Polytechnic Department of Computing and Mathematical Sciences Departmental Research Report, No 4, 1986.
- [13] W.R. Breckon and M.K. Pidcock, Mathematical aspects of impedance imaging, Clin. Phys. Physiol. Meas., 8 Suppl A , pp77-84 , 1987.
- [14] W.R. Breckon and M.K. Pidcock, Some mathematical aspects of electrical impedance tomography, in Mathematics and Computer Science in Medical Imaging, eds M.A. Viergever and A Todd-Pokropek, NATO ASI Series, Vol F39, Springer-Verlag Berlin Heidelberg, pp351-362, 1988.
- [15] W. Breckon and M. Pidcock, Illposedness and non-linearity in Electrical Impedance Tomography, in Information Processing in Medical Imaging, eds C.N. de Graaf and M.A. Viergever, Plenum, New York, pp235-244, 1988.
- [16] W. Breckon and M. Pidcock, Data errors and reconstruction algorithms in EIT, Clin Phys. Physiol. Meas., Vol 9, Suppl. A, pp105-109, 1988.
- [17] W. Breckon, K. Paulson and M. Pidcock, Parallelism in EIT, in Proc of 11th IPMI (ed. J. Llacer), 1989.
- [18] W.R. Breckon, Measurement and Reconstruction in Electrical Impedance Tomography, Inverse Problems and Imaging, ed G. Roach, Longmans, 1990.
- [19] B. Brown, D. Barber and L. Tarassenko, eds., Proc. of 1st EC COMAC-BME workshop on Electrical Impedance Tomography — Applied Potential Tomograph, 2-4 July 1986, Clin. Phys. Physiol. Meas., 8 Suppl A, 1987.
- [20] B.H. Brown, Tissue Impedance Methods, Imaging with Non-ionising Radiations ed D.F. Jackson, Surrey University Press, pp85-110, 1983.
- [21] B. Brown, D. Barber and J. Jossinet, eds., Proc. 2nd EEC workshop on EIT, Clin. Phys. Physiol. Meas., 9 Suppl A , 1988.
- [22] A.P. Calderón, On an Inverse Boundary Value Problem, Seminar on Numerical Analysis and its Application to Continuum Mechanics, (Rio de Janeiro:Sociedade Brasileira de Matematica), pp65-73, 1980.

- [23] J.R. Cannon and J.H. Halton, The irrotational solution of an elliptic differential equation with an unknown coefficient, *Proc. Camb. Philos. Soc.*, Vol 59, pp680-682, 1963.
- [24] K.S. Cheng, Comparing Current Patterns used in Electric Current Computed Tomography, MS Thesis, Rensselaer Polytechnic Institute, NY, 1989.
- [25] G. Cimmino, Calcolo approssimato per le soluzioni dei sistemi di equazioni lineari, *La Ricerca Scientifica (Roma)* XVI, Ser II. Anno IX 1, pp326-333, 1938.
- [26] J.J. Dongarra, C.B. Moler, J.R. Bunch and G.W. Stewart, *Linpack user's guide*, SIAM, Philadelphia, 1979.
- [27] P.P.B. Eggermont, G.T. Herman, and A. Lent, Iterative algorithms for large partitioned systems with application to image reconstruction, *Lin Alg Appl*, 40, pp37-67, 1981.
- [28] M.J. Field , *Differential Calculus and Its Applications*, Van Nostrand Reinhold, N.Y., 1976.
- [29] G. B. Folland *Introduction to Partial Differential Equations*, Princeton University Press, Princeton, 1976.
- [30] P. Furner, Analysis and Development of an Electrical Impedance Tomograph, BSc dissertation, School of Engineering, Oxford Polytechnic, 1990.
- [31] I.M. Gel'fand, M.I. Graev, Z.Ya. Shapiro, Differential Forms and Integral Geometry, *Functional Anal Appl*, 3 p24-40, 1969.
- [32] D.B. Geselowitz, An application of electrocardiographic lead theory to impedance plethsmography, *IEEE Trans. Biomed Eng*, BME-18, pp38-41, 1971.
- [33] P.F.C Gilbert, Iterative methods for three dimensional reconstruction of an object from projections, *J Thor Biol*, 36, pp105-117, 1972.
- [34] G.H. Golub and W. Kahan, Calculating the singular values and pseudo-inverse of a matrix, *SIAM Numer. Anal.*, 2, pp202-224, 1965.
- [35] G.H. Golub and C.F. van Loan , *Matrix Computations*, North Oxford Academic, 1983.
- [36] D. Gilbarg and N.S. Trudinger, *Elliptic Partial Differential Equations of the Second Order*, Springer-Verlag, Berlin, 1977.

- [50] L.V. Kantorovich, The method of successive approximations for functional equations, *Acta Math*, 71, pp63–97, 1939.
- [51] Y. Kim, W.J. Tomkins and J.G. Webster, Medical body imaging using electrical impedance and non-linear reconstruction, *Proc. 10th Annual Northeast Bioengineering Conference*, NH, USA, 1982.
- [52] Y. Kim, J.G. Webster and W.J. Tomkins, Electrical Impedance Imaging of the thorax, *J. Microwave Power*, vol 18, pp245–257, 1983.
- [53] Y. Kim and H.W. Woo, A prototype system and reconstruction algorithms for electrical impedance technique in medical body imaging, *Clin. Phys. Physiol. Meas.*, 8 Suppl A, pp63–70 1987.
- [54] R.V. Kohn and M. Vogelius, Identification of an unknown conductivity by means of measurements at the boundary, in ‘Inverse Problems’, D. McLaughlin, ed., *SIAM-AMS Proc. No 14*, Amer. Math. Soc., Providence, pp113–123, 1984.
- [55] R. Kohn, M. Vogelius, Determining the conductivity by boundary measurements. *Comm. Pure Appl. Math.*, 38, pp643–667. 1985.
- [56] C.J. Kotre, A sensitivity coefficient method for the reconstruction of electrical impedance tomograms, *Clin. Phys. Physiol. Meas.*, Vol 10, No3, pp275–281, 1989
- [57] R.E. Langer, On an inverse problem in differential equations, *Bull Am Math Soc* , 39, pp814–820, 1933.
- [58] R.E. Langer, On determination of earth conductivity from observed surface potentials, *Bull Am Math Soc*, 10, pp747–754, 1936.
- [59] K. Levenberg, A method for the solution of certain non-linear problems in least squares, *Q J Appl Math*, 2, pp164–168, 1944.
- [60] P. Lorrain and D. Corson, *Electromagnetic Fields and Waves*, Freeman, San Francisco, 1970.
- [61] J. Rosell, D. Murphy, R. Pallas and P. Rolfe, Analysis and assessment of errors in a parallel data acquisition system for electrical impedance tomography, *Clin. Phys. Physiol. Meas.*, Vol 9, Suppl. A, 93–100, 1988.
- [62] C. Toumazou, F.J. Lidgley, C.A. Makris, Extending voltage-mode op amps to current-mode performance, *IEE Proc.*, Vol 137, Pt. G, No 2, pp116–130, 1990.

- [63] J.L. Lions and E. Magenes, Non-homogeneous boundary value problems, Springer-Verlag, 1972.
- [64] D. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, SIAM J Appl Math, Vol 11, No 2, pp431-441 1963.
- [65] J.J. Moré, The Levenberg-Marquardt algorithm: implementation and Theory, in 'Numerical Analysis', ed G.A. Watson, Lecture Notes in Mathematics Vol 630, Springer, pp105-116, 1978.
- [66] V.A. Morozov, Methods for solving incorrectly posed problems, Springer, Berlin, 1984.
- [67] T. Murai and Y. Kagawa, Electrical Impedance Computed Tomography based on a finite element model, IEEE Trans Biomed Eng, vol BME-32, pp177-184, 1985.
- [68] D. Murphy, Development of Electrical Impedance Methods with Application to Plethsmograph and Tomograph, DPhil Thesis, Bio-medical Engineering Centre, Univ. Oxford, 1987.
- [69] D. Murphy, The Oxford Polytechnic Adaptive Current Tomograph, Department of Computing and Mathematical Sciences Research Report No 14, Oxford Polytechnic, 1988.
- [70] D. Murphy, F.J. Lidgey, W.R. Breckon, C.N. Mcleod, T. Davey-Winter, An adaptive current impedance tomograph, Proc. 2nd IFBME Pan Pacific Symposium, Melbourne, 1989.
- [71] F. Natterer, A Sobolev Space Approach to picture reconstruction. SIAM J Appl Math 39, pp402-411, 1980.
- [72] F. Natterer, The Mathematics of Computerized Tomography, Wiley, Chichester, 1986.
- [73] F. Natterer, The finite element method for ill-posed problems, RAIRO Anal Numerique, 11, 271-278, 1977.
- [74] J. C. Newell, D. G. Gisser and D. Isaacson, An Electric Current Tomograph, IEEE Trans. Bio-med. Eng., Vol 35, No 10, 1988.
- [75] L.R. Price, Imaging of the electrical conductivity and permittivity inside a patient: a new computed tomography (CT) technique, SPIE Vol 206, Recent and future developments in Medical Imaging II. pp115-119, 1979.

- [76] L.R. Price, Electrical Impedance Computed Tomography (ICT): a new CT imaging technique, IEEE Trans Nucl Sci, No 26, 2, pp2736-2739, 1979.
- [77] A.G. Ramm, A simple proof of the uniqueness theorem in impedance tomography, Appl. Math. Lett, 1 , 1988.
- [78] W. Rudin, Functional Analysis, Tata McGraw-Hill, New Delhi, 1982.
- [79] K. Sakamoto and H. Kanai, A fundamental study of an electrical impedance CT algorithm, Proc., Vth Int. Conf. Elect. Bio-Impedance, Zadar, Yugoslavia, pp349-352, 1983.
- [80] F. Santosa and M. Vogelius, A backprojection algorithm for Electrical Impedance Tomography, Institute for Physical Science and Technology, University of Maryland, Technical Note BN-1081, 1988.
- [81] H. Schomberg, Reconstruction of spatial resistivity distributions of conducting objects from external resistance measurements, Philips GmbH Forschungslaboratorium, Hamburg, MS-H, 1908V/78, 1978.
- [82] H. P. Swann, Alternating current spectroscopy of biological substances. Proc IRE, 47, pp1941-1945, 1959.
- [83] A.D. Seagar, Probing with low frequency electric current, PhD Thesis, University of Canterbury, Christchurch, NZ .1983.
- [84] L.B. Slichter, The interpretation of the resistivity prospecting method for horizontal structures, J Appl Phys, v4, pp307-322, 1933.
- [85] G. Strang and G.J. Fix, An Analysis of the Finite Element Method. Prentice-Hall, Englewood Cliffs, N.J., 1973.
- [86] E. Suli, Convergence and non-linear stability of the Lagrange-Galerkin method for the Navier-Stokes equations, Numer. Math., Vol 53, pp459-483, 1988.
- [87] J. Sylvester and G. Uhlmann, A global uniqueness theorem for an inverse boundary value problem, Ann. Math 125, pp153-169, 1987.
- [88] G. Talenti, Inverse Problems, Springer Lecture Notes in Mathematics 1225, 1986.
- [89] L. Tarassenko, Electrical Impedance Techniques for the Study of the Cerebral Circulation and Cranial Imaging in the Newborn, DPhil Thesis, University of Oxford, 1985.

- [90] M. Tasto and H. Schomburg, Object Reconstruction from projection and non-linear extensions, in C.H. Chen (Ed.), Pattern recognition and signal processing, Alphen aan den Rijn, The Netherlands: Sijthoff and Noordhoff, pp485-503, 1978.
- [91] A.N. Tikhonov, Solution of incorrectly formulated problems and the regularisation method, Sov Math Dokl, 4, pp1035-1038, 1963.
- [92] M.A. Viergever, Introduction to discrete reconstruction methods in medical imaging, Mathematics and Computer Science in Medical Imaging, eds M.A. Viergever and A Todd-Pokropek, NATO ASI Series, Vol F39, Springer-Verlag Berlin Heidelberg, pp43-65, 1988.
- [93] J. Wloka, Partial Differential Equations, Cambridge U.P., 1987.
- [94] Y. Yamashita and T. Takahashi, Method and Feasibiliy of Estimating Impedance Distribution in the Human Torso, Proceedings of Vth ICEBI, Tokyo, pp87-90, 1981.
- [95] K. Nakayama, W. Yagi and S. Yagi, Fundamental Study on Electrical Impedance Tomography Utilizing Sensitivity Theorem on Impedance Plethysmography, Proceedings of Vth ICEBI, Tokyo, pp99-102, 1981.
- [96] T. Yorkey, Comparing reconstruction methods for electrical impedance tomography, Phd Dissertation, Dep., Elec. Comput. Eng., Univ. Wisconsin, Madison, Aug 1986.
- [97] T. J. Yorkey and J.G. Webster, A comparison of impedance tomographic reconstruction algorithms, Clin. Phys. Physiol Meas., Vol 9, Suppl. A, 55-62 , 1988.
- [98] T. J. Yorkey, J.G. Webster and W.J. Tompkins, Comparing Reconstruction Algorithms for Electrical Impedance Tomography, IEEE Trans. Biomed. Eng., BME-34, No. 11, 1987
- [99] O.C. Zienkiewicz and R.L. Taylor, The Finite Element Method, 4th Edn., McGraw-Hill, London, 1989.

