Reconstruction of Grounded Objects from Cauchy Data on a Plane

van Berkel, Cornelis

2008

MIMS EPrint: **2008.76**

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

Reconstruction of Grounded Objects from Cauchy Data on a Plane

A mathematical study on shape detection for interactive displays

A thesis submitted to the University of Manchester for the degree of
Doctor of Philosophy
in the
Faculty of Engineering and Physical Sciences

2008

Cornelis van Berkel

Philips Research Laboratories, Redhill

&

School of Mathematics
The University of Manchester

CONTENTS

ABSTRACT

This thesis explores the use of capacitance measurements made between electrodes embedded in or around a display surface, to detect the position, orientation and shape of hands and fingers. This is of interest for unobtrusive 3D gesture input for interactive displays, so called touch-less interaction. The hand is assumed to be grounded.

The forward problem is solved using Green's theorem and an appropriate Green's function. This leads to an operator factorisation for the forward Dirichlet to Neumann map $\Lambda_D : L^2(\partial H) \rightarrow L^2(\partial H)$. The foward map is demonstrated to be compact, injective and depends uniquely on the object. An alternative factorisation based on double layer potentials and involving a Fredholm equation of the second kind is also presented. These operator expressions are used in numerical calculations in two and three space dimensions using the Boundary Element Method for discretization.

Four methods are presented for the solution of the inverse problem of recovering the object from a measured forward map. The first uses modified Gauss-Newton optimization. The method is successful if the degrees of freedom are limited to object position, size and orientation, but is unpractical for shape reconstruction.

The second method recovers the zero potential contour of a solution to Laplace's equation from Cauchy data on part of the boundary of a domain. An algorithm is used where at each iteration there is an approximation ∂D_k to ∂D on which approximate Cauchy data are calculated by solving a Tikhonov regularised linear system. This data is used to modify ∂D_k by extrapolation towards the zero-surface giving the next approximation ∂D_{k+1}.

In the third method the problem is solved with the so-called Factorisation Method. A test function g_z is used to characterise points $z \in D \iff g_z \in \mathcal{R}(\Lambda_D^{1/2})$. Implicit regularisation due to the finite aperture of the measurement electrode results in a level set $P(z)$ that is finite and differentiable everywhere. The level representing the object ∂D is found through minimization of the cost function.

The fourth method uses a monotonicity property of the forward map to test if a probe object is contained within the unknown object. For an infinitesimal probe object and finite aperture measurements the method is shown to be identical to the factorisation method.

The thesis closes with conclusions on the relative merits of these methods.

DECLARATION

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning

COPYRIGHT STATEMENT

PREFACE

"Is it possible, using capacitance measurements with electrodes around the edge of display, to detect the location, orientation and shape of one or more hands or fingers?" Put simply; "Can we give displays 'eyes' with which to see, in particular see in 3D?"

The question arose as part of my work on 3D displays and interactive displays. At Philips Research we had developed some empirical methods for simple object detection, but it was important to know if mathematics in general, and the magic of inverse problem theory in particular, could provide better answers.

Bill Lionheart at Manchester combined expertise and enthusiasm in this area and initially we tried to set up a collaboration as part of a European project, but this unfortunately failed. We then thought that a good way to work together was under the banner of an external PhD. Rather than wait for this or that grant, we could start immediately and there were obvious benefits to both organisations. There also was a personal motivation. Both my physics education at the university of Utrecht and my work at Philips Research had involved substantial amounts of mathematics, but here was a chance to understand, use and 'do' mathematics at a much more sophisticated level.

The maths, the research, the papers and also the writing of this thesis itself have been fantastic experiences. Not least because we have come up with answers, but also because at every stage the standard of the intellectual challenge was raised. I am grateful to Bill for guiding and coaching me through this process so skilfully.

I am indebted for the support and encouragement of a large number of people who have helped me in getting this thesis written. There is however one person who deserves special mention and that is my wife Fran. She has not only rewarded me with her support and admiration, but she has also put up with a distracted husband and taken the strain of running a large family while I was shut away during many evenings and weekends.

PUBLICATIONS

Numbers 31, 32, 34 and 35 of this list relate to the subject matter of this thesis.

1. 'Photovoltaic Detection of Magnetic Resonance in a Si:H Solar cells', K P Homewood, B C Cavenett, C van Berkel, W E Spear and P G LeComber in Optical Effects in Amorphous Semiconductors (Snowbird, Utah 1984) AIP Conference Proceedings 120 eds, P C Taylor and S G Bishop.

2. 'The Photosensitivity of Amorphous Silicon Thin Film Transistors', C van Berkel and M J Powell; J of Non Cryst Solids 77&78 1393 (1985)

3. 'Evidence for Metastable Defects in Amorphous Silicon Thin Film Transistors', A R Hepburn, J M marshal, C Main, M J Powell and C van Berkel; J of Non Cryst Solids 77&78 1409 (1985)

4. 'Photo field Effect in Amorphous Silicon Thin Film Transistors', C van Berkel and M J Powell; J Appl Phys 60 1521 (1986)

5. 'Metastable Defects in Amorphous Silicon Thin Film Transistors', A R Hepburn, J M Marshall, C Main, M J Powell and C van Berkel; Phys Rev Lett 56 2215 (1986)

6. 'Dynamic Characteristics of Amorphous Silicon Thin Film Transistors', C van Berkel J R Hughes and M J Powell; MRS Symp Proc 95 445 (1987)

7. 'Resolution of Amorphous Silicon Thin Film Transistor Instability Mechanisms using Ambipolar Transistors', C van Berkel and M J Powell; Appl Phys Lett 51 1094 (1987)

8. 'Bias Dependence of Instability Mechanisms in Amorphous Silicon Thin Film Transistors', M J Powell, C van Berkel, I D French and D H Nicholls; Appl Phys Lett 51 1242 (1987)

9. ‘The Resolution of A Si TFT Instability Mechanisms’, M J Powell, C van Berkel and I D French; J Non cryst Solids 97&98 321 (1987)

10. ‘Charge Trapping Effects in Amorphous Silicon / Silicon Nitride Thin Film Transistors’, A R Hepburn, C Main, J M Marshall, C van Berkel and M J Powell; J Non Cryst Solids 97&98 903 (1987)

11. ‘Deep Trapping Controlled Switching Characteristics in Amorphous Silicon Thin Film Transistors’, C van Berkel, J R Hughes and M J Powell; J Appl Phys 66 4488 (1989)

12. ‘Time and Temperature Dependence of Instability Mechanisms in Amorphous Silicon Thin Film Transistors’, M J Powell M J, C van Berkel and J R Hughes; Appl Phys Lett 54 1323 (1989)

13. ‘Analysis of Distributed Threshold Voltage Transistors’ M Matsumura, C van Berkel and Y Uchida; IEEE Trans on Elec Dev 36 2965 (1989)

14. ‘Characterization of Distributed Threshold Voltage Transistors with Amorphous Silicon’, C van Berkel and M Matsumura; Jap J Appl Phys 29 1649 (1990)

15. ‘Temperature Dependence of a Si:H NIP Diodes’, C van Berkel, M J Powell and I D French; MRS Symp Proc 219 369 (1991)

16. ‘Instability Mechanisms in Amorphous Silicon Thin Film Transistors and the Role of the Defect Pool’, M J Powell, C van Berkel and S C Deane; J Non Cryst Solids 137&138 1215 (1991)

17. ‘Amorphous Silicon Thin Film Transistors; Physics and Properties’ C van Berkel in Physics and Applications of Amorphous and micro crystalline semiconductor devices, ed J Kanicki, Artech House, New York (1992)

18. ‘Quality factor in a-Si:H nip and pin diodes’, C van Berkel M J Powell, A R Franklin and I D French; J Appl Phys 73 5264 (1993)

19. ‘2D image sensor arrays with nip diodes’, C van Berkel N C Bird, C J Curling and I D French; MRS Symp Proc 297 939 (1993)

20. ‘Physics of a-Si:H switching diodes’, C van Berkel, M J Powell and S C Deane; J Non-cryst Solids 164-166 653 (1993)

21. ‘Microlens arrays for 2D large area image sensors’, C van Berkel, B P McGarvey and J A Clarke; Pure Appl. Opt. 3 177 (1994)

22. ‘Reverse Current Mechanisms in a-Si:H nip diodes’, N Kramer and C van Berkel; Appl Phys Lett 64 pp1129-1131 (1994)

23. ‘a-Si:H Diodes for Active Matrix Addressing’, C van Berkel; Optoelectronics - Devices and Technolgies 9 323 (1994)

24. ‘Large-area Image Sensing Using Amorphous Silicon NIP Diodes’, N C Bird, C J Curling and C van Berkel; Sensors and Actuators A 46-47 444 (1995)

25. ‘Multiview 3D-LCD’ C van Berkel, D W Parker and A R Franklin; in Stereocopic Displays and Applications, Proc SPIE vol 2563 pp32-39 (1996)

26. ‘Design and Applications of Multiview 3D-LCD’, C van Berkel, A R Franklin and J R Mansell, Proc 16th Int Display Research Conf, EuroDisplay96 (SID), Birmingham Oct 1-3 pp109-112 (1996).

27. ‘Characterisation and Optimization of 3D-LCD module design’, C van Berkel and J A Clarke; Proc SPIE vol 3012 pp179-187 (1997).

28. ‘Image preparation for 3D-LCD’ C van Berkel, Proc SPIE vol 3639 pp84-91 (1999)

29. ‘Real-time auto-stereo visualisation of 3D medical images’, L Portoni, A Patak, P Noirard, J C Crossetie and C van Berkel. SPIE proc 3976 (2000).

30. ‘Multiuser 3D displays’, C van Berkel and J H A Schmitz Tile Conference, London 2000

31. ‘3D Touchless Display Interaction’ C van Berkel; SID Proc Int Symp, vol 33, number 2, pp1410-1413, May 19-24, 2002

32. ‘Electrostatic Object Reconstruction in a Half Space’ C van Berkel W R B Lionheart, Proc of the 5th Int Conf on In Prob in Engineering: Theory and Practice,

Cambridge, UK 11-15th July 2005, Chapter V01, Editor D Lesnic, Leeds University Press 2005

33. 'Pixel Design for Optimized 2D/3D Displays' C van Berkel, M Inoue, T Inada, H Louwsma, S T de Zwart and J R Hector; 12th International Display Workshops & Asia Display 2005 (IDW/AD05), Takamatsu, Japan, Dec 6th - Dec 9th, pp1759-1760 2005

34. 'An Iterative Method for Electrostatic Object Reconstruction in a Half Space' C van Berkel, W R B Lionheart; Meas Sci Tech, vol 18 pp41-48 (2007)

35. 'Object reconstruction with an indicator function' C van Berkel, W R B Lionheart; Inverse Problems in Science and Engineering vol 15 pp585-600 (2007)

1. INTERACTIVE DISPLAYS AND CROSS CAPACITANCE SENSING

1.1 Introduction

Electronic information displays are everywhere. Even if once the word 'display' was synonymous with 'TV', the word is now used in the context of phones, computers and even clothes[1], and this spectrum is certain to spread. Displays will appear in every aspect of our lives. In almost every case we want to interact with the content shown on the screen; we'd like to run our thumb over the display of our mobile phone, we will want to draw characters with our fingers in the space in front of the screen. In the kitchen, or bathroom, we'd like to control the channels on the screen without getting dirty finger marks on the screen. The interaction paradigm will not be just one mouse pointer, but encompass two hands, ten fingers and gestures.

Existing interaction techniques such as keyboard, mouse, joystick, touch screen, data gloves or camera-based gesture recognition are too awkward, limited or expensive to be used by different people across a large range of applications. An attractive alternative is to use weak electric fields around the display to sense finger and hand movements. The central 'thesis' is that it is possible to extract the position, size, orientation and shape information from capacitance measurements made in the plane of the display. Section 1.2 details the contribution that this thesis makes toward achieving this goal.

It is important to place the possibilities of capacitance based interaction in the context of other interaction technologies that have been developed and section 1.3 of this introductory chapter presents a selection of these. The chapter then presents a superficial description of the capacitance measurements of interest in this thesis and discusses a number of hardware demonstrations that have been made or are under development. A mathematically rigorous description of the electrostatics is presented

later in chapter 3. This introductory chapter proceeds with a brief review of other applications in which capacitance or impedance measurements are used. It is interesting to note that our ambition to provide a display with electrostatic eyes to 'see', is a direct analogy to the way some types of fish use electric fields to 'see' in dark and muddy water[2, 3]. The last section of this chapter therefore presents a brief review of electroreception in biology.

1.2 The Contribution of this Work

This thesis presents a comprehensive analysis of the forward problem, four different methods for solving the inverse problem and a numerical implementation of the forward and inverse solutions. A superficial solution to the forward problem had been presented by others[4] and the basic electrostatics of a transmitter in a ground plane is undergraduate electrostatics[5, 6]. However, the derivation of an explicit expression for the capacitive coupling between electrodes is, if not novel, at least useful. Calculating the change in this coupling in response to the introduction of a grounded object; the forward map and the operator factorisation of the forward map, had not been done before for the situation of an unbounded domain. The demonstration that this forward map is compact and has a monotonicity property (theorems 3.8 and 3.11) is new.

The thesis presents four approaches to solving the inverse problem of varying degrees of novelty. We show through numerical simulation experiments that the choice of optimum algorithm depends on the trade-off between required detail, the amount of measurement data, the noise in that data, and, of course, speed. The shape fitting approach described in chapter 5 does not provide new mathematical results but is worthwhile to present and evaluate because it represents a basic way of attacking any inverse problem. It is also practically useful for finding the initial guess and some determination of orientation. We show however that the method fails at meaningful shape reconstruction and this demonstrates the requirement for more sophisticated methods to solve the inverse problem. In the second and third inversion methods we apply and extend techniques developed for inverse acoustic and electromagnetic scattering to the situation here. Iterative reconstruction methods based on analytic continuation are common in inverse problems and we have developed one for the situation here in chapter 6. Methods like this are heavily dependent on regularisation and we study

the regularisation for the limited aperture and noise in this case. The Factorisation Method of chapter 7 is used here for the first time on an unbounded domain and a test function suitable for a Dirichlet to Neumann map rather than the scattering far field data or an electrostatic Neumann to Dirichlet map on a bounded domain. The chapter also makes a contribution to the understanding of the implicit regularisation in the Factorisation Method and the link with the Linear Sampling Method. The final method in chapter 8 presents a new adaptation of a method developed for Electrical Resistance Tomography by taking advantage of the monotonicity property of the forward map. A link is demonstrated between these independently developed methods by demonstrating that under certain circumstances the Factorisation Method and the Monotonicity Method can be regarded as equivalent.

1.3 Interaction Technologies

The functional advantage of an intuitive interaction system that responds to commands as well as humans do, was put succinctly by Bolt in the title of his paper 'Put-That-There Voice and Gesture at the Graphics Interface'[7]. The desire to create a system that combines contextual information (knowing what 'that' refers to) with gesture (where 'there' is) and an apparently unambiguous voice command, has remained strong in interactivity research directed at a wide variety to applications ranging from information kiosks, consumer electronics and medical systems. Separate research fields exists around the different components of this central question including voice recognition, context awareness and gesture recognition.

A number of technologies have been investigated for gesture recognition. Conventional touch screens, made of resistive or capacitive pads[8], are ubiquitous but are limited by the fact that they can only sense one finger position at the time. More interesting are systems that can detect multiple fingers or hands. This can be done for instance with an X-Y grid of wires and monitoring the capacitance at each cross over between the wires. Rekimoto[9] in particular has made some interesting demonstrations. Han[10] uses a glass plate in which total internal reflection is frustrated at a touch point to achieve multi finger touch input. At each touch point light is coupled out at the rear side and detected with a simple camera system. The demonstration movie

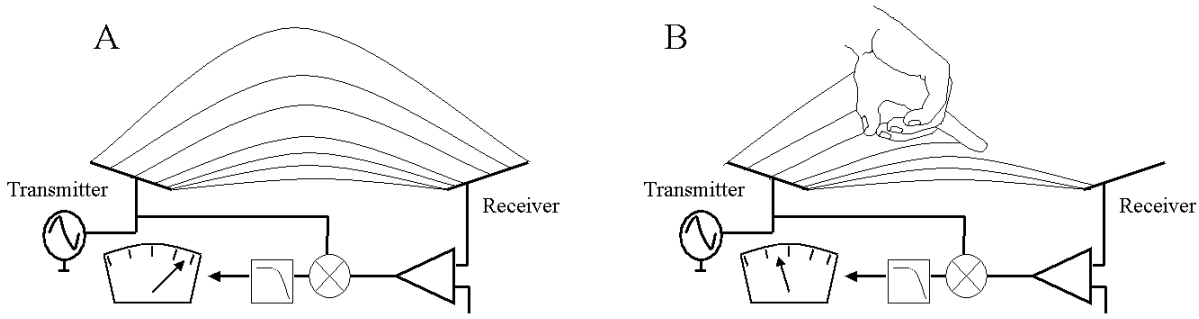

Fig. 1.1: Cross Capacitance Sensing

on the web is worth watching[1]. The detection of multiple hands and fingers extends to environments in which multiple users interact with a graphical display creating a social experience[9, 11, 12]. Although they are interesting platforms to demonstrate new methods to interact with computer graphics, these methods have not led to actual new products. This may be because the usefulness of multi-finger input is limited, but it may also be because the systems are physically too large, expensive and cumbersome for wide scale adoption.

Some gesture recognition technologies, particularly the electric field sensing technology that this thesis is concerned with, have been suggested to take the interaction away from the graphical user interface itself. An interesting example in this respect is a "Finger-Joint Gesture keypad"[13] in which it is proposed that the palm of the hand itself becomes the (imagined) interaction pad. The practical implementation of this simple, but challenging idea probably awaits substantial future improvements in both sensing technology and theory. Continuing down in scale, from displays to hands and then down to fingerprints, there have been suggestions to merge fingerprint sensing and interaction with some innovative technologies proposed, for instance NASA's capaciflextor[14] and different finger print sensor technologies[15, 16].

Because human beings have a strong visual bias, camera based gesture recognition has been investigated by many groups, in particular because, as web cams, camera capture systems have recently become low cost. A recent example is the visual Touchpad[17, 18] from the university of Toronto, which uses algorithms that go back to Segen[19]. It is worthwhile to note that the finger tip detection is made straightforward in this case by creating an artificially high contrast ratio by insisting on a uniform

[1] http://mrl.nyu.edu/ jhan/ftirtouch/multitouchreel.mpg

background. Indeed the point has been made[19] that the challenge in machine vision is not the recognition of the object itself, but its dissociation from the background. There has been little success in developing algorithms that are independent of assumptions about the background. One approach is to make the object highly distinct from that background, for instance through wearing special suits with reflective markers, as is common for motion capture in the film industry[20], or by placing reflecting dots on key joints[21]. These solutions are robust but have obvious limitations in a general purpose consumer setting. An alternative is to control the infra-red lighting of a scene or by using two or more cameras to create a confocal plane where hands are detected for interaction while the background is blurred out[22].

1.4 Cross Capacitance Sensing

The simplest form of capacitance sensing is one in which just one electrode is used, while the other half of the two plate capacitor is formed by the object that is being sensed. The electronics measures the amount of charge that is required to bring the single electrode to a fixed voltage. This capacitive load increases as the object comes in closer, thus providing a proximity sensor. This method is used for buttons[23], fingerprint sensors[24] and foot sensors[25]. It is available in cars to asess the size of a person occupying a seat[26, 27] to adjust the response of the airbags and there have been attempts at 3D profile sensing[28].

This thesis is concerned with a cross capacitance sensing (CCS) technique in which not the load capacitance of an electrode is measured, but the change in capacitance between two electrodes as the object approaches. Figure 1.1 shows a simple combination of two electrodes forming a capacitor, with a capacitive current flowing between them. A simplified description of the way cross capacitance works is that if a hand is placed near the electrodes, some of the field lines will be terminated on the hand and the current will decrease[29]. A measurable effect can be obtained in which the spatial range is roughly equivalent to the separation between the electrodes. The idea of sensing capacitance between electrodes is certainly not new, with some references going back to the work by Blumlein on capacitance sensing between two electrodes for aircraft altimeters during the second World War[30], described in more detail by McLeod *et al*[31]. Electric field sensing has also been used in works of art[32] and has

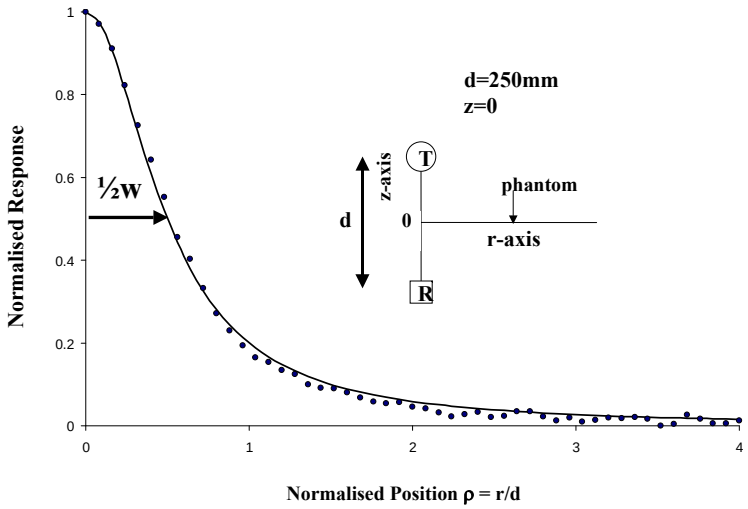

Fig. 1.2: Transmitter-Receiver Response

recently become available in integrated chip format[33].

To measure the response of a single transmitter-receiver pair, two electrodes were placed on a table at 25cm separation[34]. As a hand phantom an upright metal plate of 5cm (wide) x 10cm (height) connected to ground was used. The hand phantom was moved across the table on a line (the r-axis) perpendicular to the axis between the T/R pair (the z-axis), crossing the axis midway between the electrodes ($z = 0$). Figure 1.2 shows the normalized response curve. The horizontal axis is the normalized r position with respect to the transmitter-receiver separation. The response on the vertical axis is normalized with respect to minimum and maximum signal strengths. Measuring curves at different T-R separations or at different z-values yields similar curves if plotted in this way. A good fit for the normalized response curve for an electrode pair is given by a Lorentzian curve of the form

$$S(\rho) = \frac{w}{4\rho + w}, \tag{1.1}$$

in which w is the full width at half height. A typical value for w is 1 and this can be demonstrated to derive from a simple electrostatic model[34].

Although there have been some concerns about the effects of electric fields on humans[35], the relatively low voltage and frequencies (10V, 100kHz) used induce currents that are so small ($< 10^{-15}$A) that they are unlikely to cause problems.

Arrays of multiple electrodes offer rich possibilities for object and gesture recognition. For instance, n transmitters and m receivers arranged around the edge of a display will contain nm electrode pairs at many different separations. These measurements are performed at a frequency of $f = 100$kHz. At this frequency the impedance of the human body is relatively low, approximately 50kΩ, smaller than the capacitive coupling between hand and electrodes, estimated at $20aF$, or $1/2\pi$f C= 80MΩ. The body can be regarded as connected to ground through the capacitance of the shoes, which, using the permitivity of PVC, can be estimated at 310 nF, much larger than the above mentioned coupling between hand and electrodes and a much smaller impedance at 100kHz than the body impedance.

At 100kHz the wavelength in air is c/f =3km, much longer than the typical dimension of the transmitter-receiver distance or the display-user distance. The analysis can therefore be performed in the near field limit of an electrostatic formulation[29]. Moreover, commercial LCD screens typically carry a transparent indium tin oxide (ITO) electrode on the glass substrate closest to the observer which is either grounded or carries low frequency, low impedance signals. The display is therefore regarded as a ground plane with known voltage distribution.

The physics of the cross capacitance sensing allows us therefore to formulate the cross capacitance sensing as a boundary value problem for the Laplace equation in the (half) space surrounding the hand above the display. The boundary value on the plane (display) is given by the known potential (that is the amplitude of the 100kHz AC signal on the transmitter electrodes) and the boundary value on the hand is zero. The boundary value problem has to be solved to find the normal derivative of the potential, or field, on the ground plane which, on the ground plane, is directly proportional to the capacitance or current amplitude measured at the receiver electrodes. A formal definition of the problem and solution is given in chapter 3.

The unique opportunity provided by this sensing modality lies in the combination of these features: 1) The technique is compatible with active matrix technology. It can therefore be integrated into the glass of the display itself, thus providing a very cost-effective way of implementing arrays of many sensors. 2) The sensors can be arranged around the edge of the display and there is therefore no degradation of display performance as with many conventional interaction techniques. 3) The 3D sensitivity creates a 3D interaction space that extends beyond the 2D display and is capable of

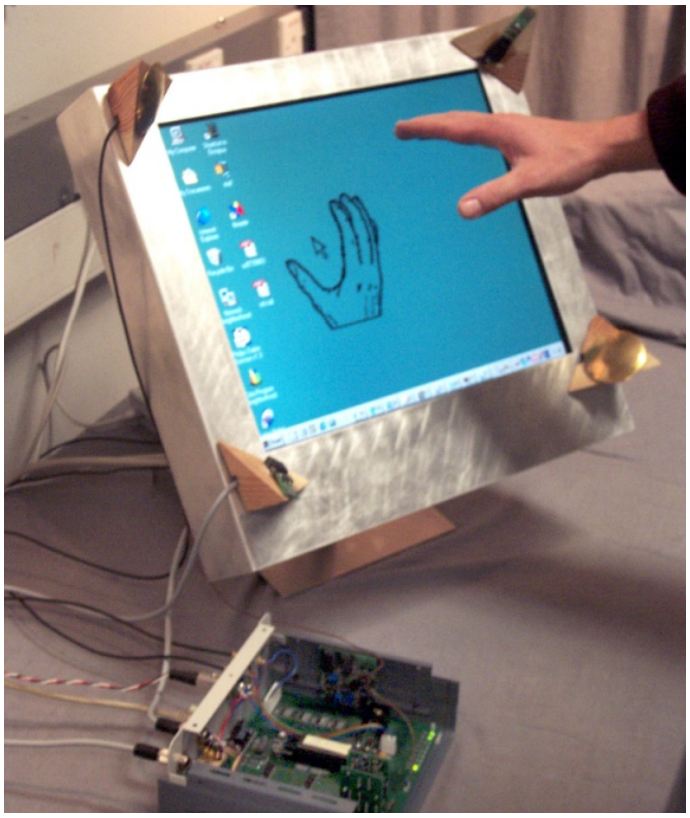

Fig. 1.3: Interactive display demonstration system

detecting multiple objects (hands, fingers) simultaneously.

1.5 Hand Mouse System

To test the capabilities of cross capacitance sensing we have used conventional electronics to build a demonstration system of two transmitter electrodes and two receiver electrodes arranged around the edge of an 18" LCD display. Figure 2 shows the system. The display is housed in a stainless steel box to provide additional shielding against electronic noise from the display and the electrodes are placed slightly proud of the display plane to maximize the signals. The transmit electrodes located in the top left and bottom right corners carry 3V pkpk, 100kHz signal. The receive electrodes, located in the top right and bottom left corners, carry current amplifiers connected to the data collection box in the foreground. This box contains signal generation circuitry, analogue phase sensitive lock-in amplifiers, and a Microchip PIC16C77 controller for data management and communication to the computer host.

The computer runs a background program to convert the sensor readings into x,y and z information and positions the cursor appropriately on the screen[34]. As an additional interface feature an image of a hand is drawn around the cursor to aid the intuitive understanding of the mouse control. The size of the hand image depends on the distance to the screen; the further the user's hand is from the screen, the larger the image, reflecting an increasingly imprecise relationship between hand position and screen coordinates as the hand moves away. At a critical distance of about 30cm, the hand image disappears altogether indicating that the touchless sensing is no longer active, and the cursor remains static.

During interaction, when the hand approaches an icon, the icon size changes inversely with the distance between hand and screen. Zooming screen icons like this has been shown to increase interaction efficiency[36]. Activation ('single' or 'double clicking') is achieved by holding the hand/cursor over an icon for 2 seconds and termination ('ESC', 'EXIT' or 'ALT+F4') through a quick swiping movement across the display. The command translation, i.e. whether for example there is something to activate and whether activation is a single or double click, is done on a context and application aware basis. This simple interface allows the user to perform tasks like opening and closing applications, Internet browsing and turning off the computer. More complex tasks are performed with a standard mouse, which functions alongside the touchless interaction.

The results of this demonstration system were very encouraging. The unique opportunity provided by this sensing modality lies in the combination of these features: 1) The technique is compatible with active matrix technology. It can therefore be integrated into the glass of the display itself, thus providing a very cost-effective way of implementing arrays of many sensors. 2) Because the sensors can be arranged around the edge of the display there is no degradation of display performance as there is with many conventional interaction techniques. 3) The 3D sensitivity creates a 3D-interaction space that extends beyond the 2D display and is capable of detecting multiple objects (hands, fingers) simultaneously. Touchless, 3D interaction becomes even more interesting when considering future autostereoscopic (no glasses) 3D displays[37]. With conventional touchscreens, objects like push buttons have to be touched at the display screen itself, even if the 3D display shows them floating in front of the display screen. Touchless 3D interaction modalities like the one discussed here would allow

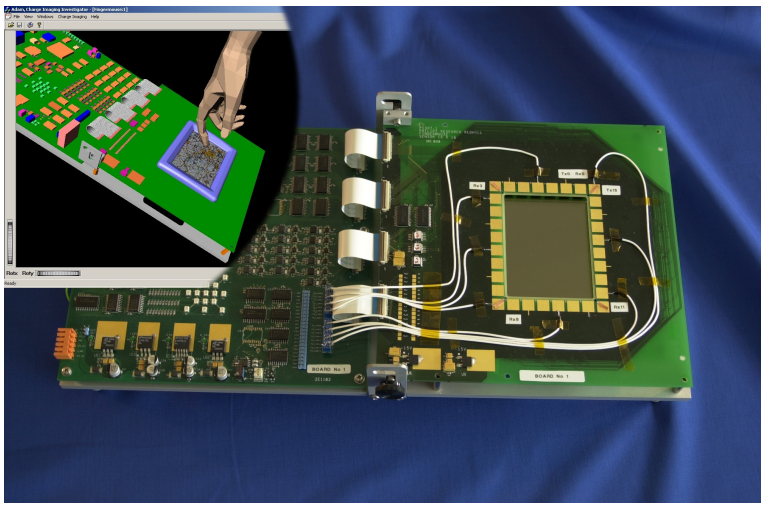

Fig. 1.4: Fingermouse experimental system for touchless interaction

user to see 3D and interact in 3D.

1.6 Multiple Electrode Systems

The capability of the hand mouse system mimicked the capabilities of the conventional mouse complemented by the z-axis detection for activation and zoom. To explore the future potential of the technology we are interested in arrays of multiple electrodes, which would offer possibilities for object and gesture recognition. For instance, n transmitters and n receivers arranged around the edge of a display will contain n^2 electrode pairs at many different separations. With sophisticated algorithms allowing the detection of individual fingers and gestures, these systems would create the promised interaction capability that goes beyond current mouse or touchscreen devices.

A demonstration system based on a digital design that is both more robust and suitable for integration in monolithic silicon or on glass has been made as the next step. The hardware platform is illustrated in figure 1.4. It consists of two parts, a system board, carrying FPGA, microprocessor, A/D converters, MUX and level shifters is located towards the left in figure 1.4. A sensor board on the right carries the electrode pads, 16 transmitters and 16 receivers together with the basic receiver circuits. A small 4” display is located inside the window of the sensor board.

A larger display system based on the same hardware platform is illustrated in figure 1.5. In this case there are 14 transmitters and 8 receivers hidden behind the edge face panel of the display. As demonstrated in the figure, the system is capable of object

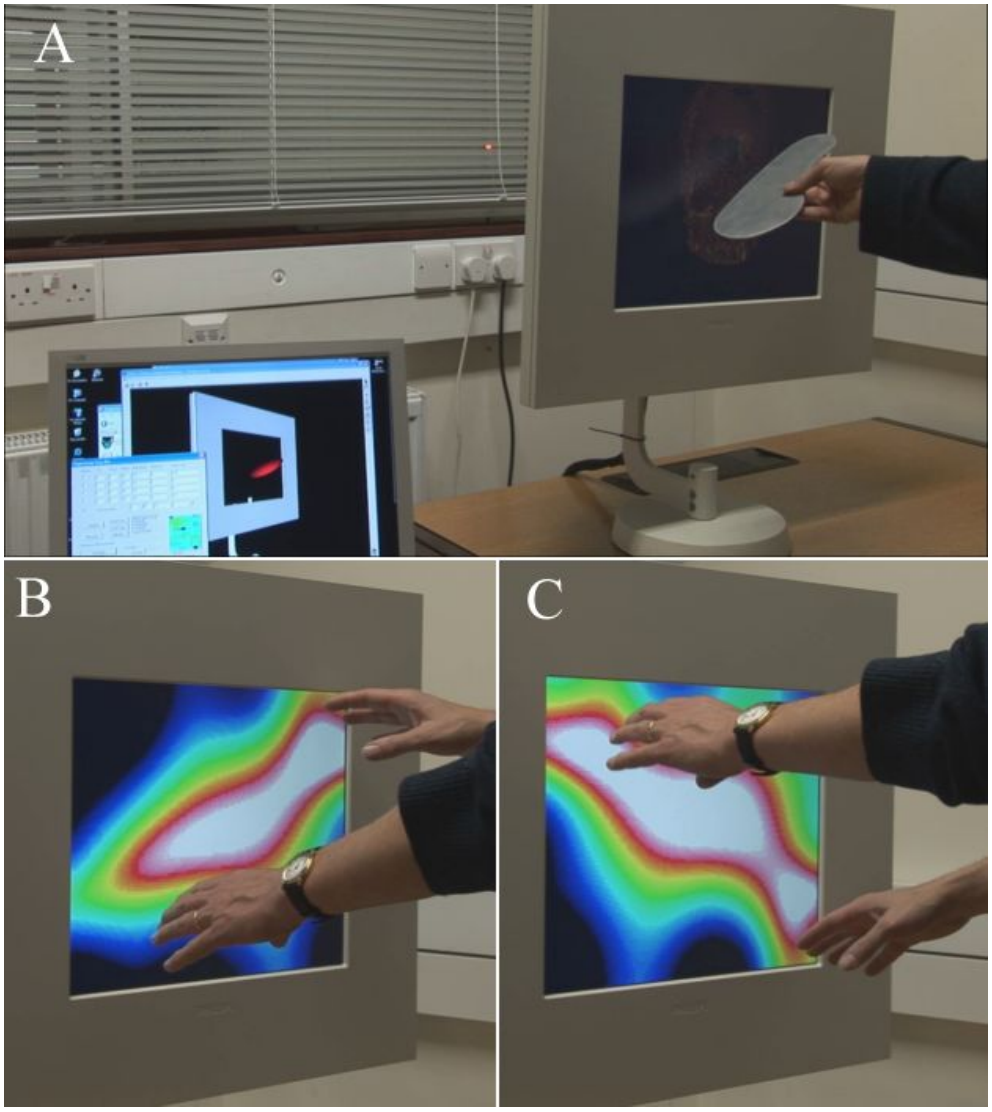

Fig. 1.5: 14 Transmitter and 8 Receiver system, demonstating A) Object orientation using the method of chapter 5 and B) and C) showing the Factorisation Method of chapter 7

orientation detection (figure 1.5A) and detection of multiple objects using the method discussed in chapter 7 (figure 1.5B and 1.5C).

1.7 Impedance Tomography

Whereas this thesis concentrates on electrical measurements for interactive displays, there are many different sensor applications in which resistance, capacitance or more general impedance measurements, are made. Electrical Impedance Tomography (EIT) aims to image organs inside the body[38, 39]. The potential between different electrodes is measured as an electric current is set up between a drive and sink electrode. Using equipotential lines calculated assuming a uniform body, local variations in resistivity can be detected and reconstructed[40]. The technique is used in a number of medical applications[41, 42, 43], particularly the monitoring of lung function[44, 45, 46], heart function and blood flow[47, 48], abdominal bleeding[49], breast cancer[50] and even deep vein thrombosis[51]. The mathematical formulation of the problem is due to Alberto Calderón[52], although it is very similar to the much older Electrical Resistance Tomography (ERT) which has been used in geophysics and archeology since the 1930s. A grid of electrodes is laid out or moved over the surface and the resistance between them measured[53]. Measurements further into the earth's crust are made by sinking bore holes in the ground and measuring currents between them. ERT is useful in mineral prospecting[54], ground water flow, detection of leaks in underground storage tanks[55], environmental cleaning[56] as well as industrial applications like corrosion detection[57]. It was in ERT that some of the early mathematical techniques for inversion of ill-posed problems found first application. In particular, A N Tikhonov the 'father of regularisation' applied his mathematical skill to the finding of copper deposits[58].

Electrical Capacitance Tomography[59] is a further variation on this theme used mainly in engineering process control. It measures the capacitance between points, but seeks to find domains of differing dielectric permittivity and is an inverse coefficient problem. The technique is used to measure flow (oil, grain etc) in pipes[60] as well as applications in the food industry[61].

Electrical measurement techniques are not only used at the macro scale of oil production of physiological function, but also at the micro scale of cell analysis. Increas-

Fig. 1.6: Gnathonemus petersii or elephant nose fish creates its own electric field. By sensing changes in the this field due to changes in the surrounding conductivity, the fish is able to navigate in dark and muddy waters.

ing interest in Lab on a Chip technologies has led to microfluidic systems for cell handling[62] and individual cell sorting based on impedance measurements of varying complexity[63, 64].

1.8 Electroreception in Biology

Electric fields are not among the human senses. We cannot feel our capacitance to metal objects or water around us and we do not see an electrostatic potential build up around a person walking across a polyester fibre carpet. We may get a shock when we then shake hands, but that is probably the closest we get to sensing electricity. Imagining ourselves to be sitting at the top of the evolutionary tree, we may be tempted to think that a sixth sense based on electric field is either impossible or not necessary, but we would be wrong. Electroreception, as biologists prefer to call it, appears in a variety of forms across different species. It is found among fishes, amphibians and even with the Platypus, an egg laying mammal[65]. This variety implies that it is not due to a single accidental mutation, but that electroreception has evolved independently several times[66, 67]. Nature has found in electroreception a solution to different problems.

For instance, a shark relies on jelly-filled canals inside its head to pick up on the tiny electrical charges that a potential next meal makes when it flexes its muscles, or

swims counter to the earth's magnetic fields.

Recent studies of neural crest cells in shark embryos[68] indicate that the common ancestor of all vertebrates detected electric fields. But as ancient vertebrates emerged from the sea, the cells that give rise to head and facial features, lost their electrosensing ability because electric field sensing is both harder, and perhaps less useful in air than it is in water.

Electrocytes, electroplaques or electroplaxes are modified muscle cells used by rays, electric eels and other electric fish for electrogenesis[69]. At rest, the cell maintains a relatively large potential of -90mV between the inside and the outside of the cell by pumping out both sodium and potassium ions. The cells are asymmetric, one side is rough, while the other side is smooth. Nerve cell synapses are attached only on the latter side. Upon excitation (postsynaptically), the nicotinic acetylcholine receptors open up ion channels only on the smooth side and the membrane potential there inverts to 50mV. Between the extracellular medium on both sides there is now a potential difference of 140mV. Knife fishes such as electric eels have thousands of these cells stacked both in series and in parallel and can produce 50A at up to 1000V[70].

Nature's best examples in electric field sensing and certainly most studied are the elephant fish or *Gnathonemus Petersii*[2, 3], illustrated in figure 1.6, *Kryptopterus bicirrhus* and *Eigenmannia virescens*. These fishes not only carry electroreceptors, but also generate their own electric field. The sensing appears sophisticated enough to allow for distance estimation[71] in a way similar to aerial perspective in the human visual system. Water is a conductive medium and the fishes sense changes in impedance rather than capacitance[72]. The sensing modality has more in common with Electrical Impedance Tomography (EIT) than with Electrical Capacitance Tomography (ECT).

Whether these fishes show us that the problem of electric field sensing is solvable in a practical way or not, is debatable. After all, our own ability to recognize objects, or interpret spoken commands, has not led to straightforward algorithms for computer vision or speech recognition. Yet, as the central 'thesis' is developed in the subsequent chapters, the example of the elephant nose fish provides at least some indication that the end goal may be possible.

2. INVERSE PROBLEMS

This chapter presents an introduction to inverse problems. It provides a broad background against which to set the more carefully argued work in subsequent chapters on the solution of the specific problem of interest. The first section presents a simple example of an ill-posed problem. It is the last section in this thesis that requires no specialist mathematical knowledge. Subsequent sections discuss a number of inverse problems that have similarities or are relevant to the inverse problem of shape reconstruction from capacitance measurements. In particular inverse boundary problems for the Laplace equation, as well as inverse problems for the Helmholtz equation are discussed.

2.1 A Simple Ill-Posed Problem

To illustrate some basic features of ill-posed inverse problems, consider the following very simple problem. Other illustrative examples of inverse problems can be found in [73]. Two sources and two detectors are separated by a lateral distance p (pitch) and a transverse distance l as illustrated in figure 2.1. If the sources have strength x_1 and x_2 and the registered intensity at distance r falls off with r^2, then we can write for the measured intensities y_1 and y_2

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = y = Ax = \begin{pmatrix} \frac{1}{l^2} & \frac{1}{l^2+p^2} \\ \frac{1}{l^2+p^2} & \frac{1}{l^2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \tag{2.1}$$

The inverse problem is to determine the original strength x_1 and x_2 from the observed, and potentially noisy, intensities. The practical justification of this problem might be a visually impaired couple arguing over whether the car in front is indicating right or left, or it might be the challenge of capacitively measuring a finger print at $100\mu m$ resolution through a $1mm$ thick dielectric layer. The problem is also illustrative of

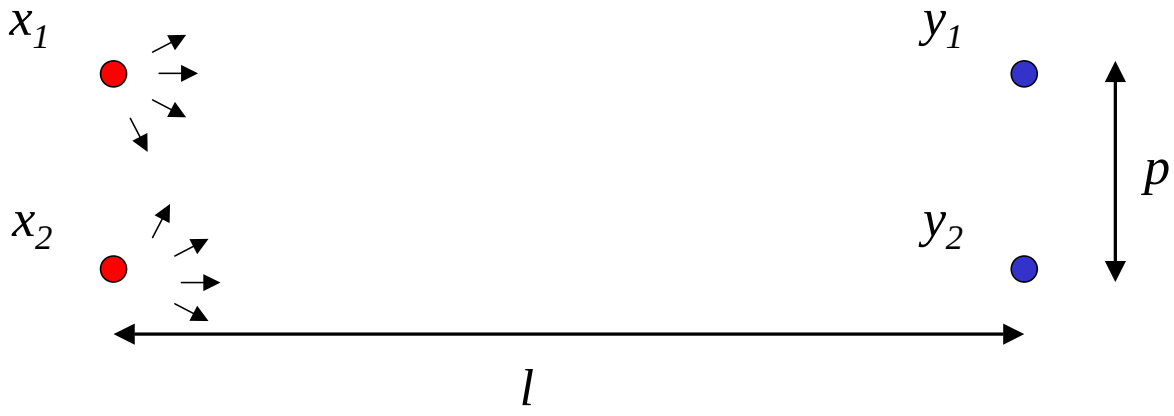

Fig. 2.1: Illustration of a simple inverse problem; from the intensity readings at y_1 and y_2, determine the source intensities x_1 and x_2.

deblurring in image processing[74].

Formally the problem is solved by $x = A^{-1}y$, in which

$$A^{-1} = \frac{l^2(l^2+p^2)}{p^4+2l^2p^2}\begin{pmatrix} l^2+p^2 & -l^2 \\ -l^2 & l^2+p^2 \end{pmatrix}. \tag{2.2}$$

If $l > p$, that is if the distance is greater than the pitch of the sensors, we see that the elements in A^{-1} differ in sign but become increasingly similar in magnitude as the distance l grows. The calculation $x = A^{-1}y$ relies on smaller and smaller differences between the elements of y. In other words, the solution x becomes very sensitive to noise in y and this sensitivity grows as l^4. Small amounts of noise can cause the calculated (recovered) x to be unreasonably large or contain negative, and therefore non-physical, components.

The 'difficulty' of an inverse problem is captured by the condition number which is the ratio of the largest and smallest singular value of the matrix. In this simple example it given by

$$\kappa(A) = \frac{2l^2+p^2}{p^2}.$$

For the matrices we deal with later in this thesis the condition number is typically 10^{18}.

This is characteristic of ill-posed inverse problems[75]. In a formal, functional analysis setting, the 'practical' concern about noise becomes the traditional mathematical pre-occupation with continuity and uniqueness. The issue of cross-talk that is the cause

of the difficulties above, becomes the mathematical concept of compactness. The continuum analogue of the problem above involves integral operators rather than matrix calculations, which then do not have bounded inverses. That is the solution does not depend continuously on the data and there is no unique solution. Rather than well-posed in the sense of Hadamard[76], the problem is ill-posed, and traditionally not regarded as worthy of serious study.

As often in the history of Mathematics, prejudice ultimately gives way to demands for answers. The geometric prejudice of the ancient greeks against anything but straightedge and compass gave way to the need for 'algebra' in running the Abassid empire[77]. The pre-renaissance notions on pre-determination were overcome in the development of probability theory by Pascal and Fermat[78]. In the case of inverse problems, the technological advances in Radar, Sonar, Seismography and Electrical Resistance Tomography (ERT) led Tikhonov[79] and others such as Phillips[80], Foster[81] and Hoerl[82] to introduce regularisation methods for ill-posed problems.

If a redesign of the experiment to minimize the cross-talk is not possible, then it is necessary to either rephrase the question or apply additional knowledge. It may be better to forgo the correct answer in favour of a reasonable and approximate one. This is the magic of regularisation.

Classical Tikhonov regularisation demands that the norm of the solution $||x||^2$ is small. The solution is now given by[83]

$$\tilde{x} = (A^T A + \alpha^2 I)^{-1} A^T y,$$

in which α is the regularisation parameter which sets the balance between fitting the data and keeping $||x||^2$ small.

In our simple example and for $p << l$ we have

$$(A^T A + \alpha^2 I)^{-1} A^T = \frac{1}{(\frac{2p^2}{l^6})^2 + \alpha^2(\frac{4}{l^4} + \alpha^2)} \frac{1}{l^2} \begin{pmatrix} \frac{2p^2}{l^6} + \alpha^2 & \alpha^2 - \frac{2p^2}{l^6} \\ \alpha^2 - \frac{2p^2}{l^6} & \alpha^2 + \frac{2p^2}{l^6} \end{pmatrix}.$$

If we choose α^2 to be larger than $2p^2/l^6$, but smaller than $4/l^4$, then

$$(A^T A + \alpha^2 I)^{-1} A^T = \frac{l^2}{4} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

The Tikhonov solution gives half the average of the two intensity readings y_1 and y_2 multiplied by the square of the distance l. In the difficult situation of $p << l$ and with only noisy measurements available, this is a sensible approach. Choosing smaller values of α^2, will recover weighted averages of the two intensities. But for very small $\alpha^2 << 2p^2/6$, this answer recovers the original, but possibly unstable $x = A^{-1}y$.

2.2 Inverse Coefficient Problems

Electrical Impedance Tomography (EIT) and Electrical Capacitance Tomography (ECT) aim to measure, or image, the distribution of conductivity or permittivity from measurements of current and or voltage from electrodes arranged around the object. In EIT this could be the body torso[38] and in ECT for instance a pipe carrying a mixture of gas and oil[60]. The EIT problem requires the solution of the following set of equations

$$\nabla \cdot \sigma \nabla u = 0 \tag{2.3}$$

$$\sigma \left. \frac{\partial u}{\partial n} \right|_{\partial\Omega} = j. \tag{2.4}$$

Where it is important to note that the conductivity varies in space. The forward problem is to find the source current density j from a known conductivity or permittivity profile $\sigma(x)$, or finding the potential distribution on the boundary for a given current density j. The inverse problem is to reconstruct the conductivity or permittivity profile from multiple measurements of j or u on the boundary[52] and is much researched in inverse problems[40].

The most straightforward approach to solving the inverse problem is local linearization by calculating the sensitivity of the current density j to small changes in the conductivity distribution σ on a pixel by pixel basis. This creates a Jacobian. $\sigma(x)$ can then by found from j by using the, possibly generalized and regularised, inverse of the Jacobian. The process can be visualised by considering the sensitivity between two specific electrodes. The current j is most sensitive to changes in permittivity at locations on a possibly curved line between the electrodes and the sensitivity falls off away from the line. The result is a banana shaped sensitivity profile. Such sensitivity profiles are used explicitly in EIT[38], ECT[84] and Optical Tomography[85]. The inversion

amounts to assigning changes in currents to changes in permittivity at locations in a weighed manner. In principle an iterative procedure can be used to improve the results but this can diverge and an important regularisation tool is a pre-emptive termination of the iteration procedure, sometimes after one step[86].

Linearization and sensitivity profiles are justified when looking for low contrast images, that is relatively small changes in the permittivity from the background value on a pixel by pixel basis. The assumption is that these small variations in permittivity cause only minimal change the overall pattern of current flow. When looking for larger variations in permittivity, an alternative to sensitivity profiles is to look for enclosed domains of constant, sometimes predetermined values of permittivity[87]. Such treatments then become similar to our next class of inverse problems.

2.3 Free Boundary Problems

An important class of inverse problems, including our own, occurs in situations where measurements, typically of potential and current, or of temperature and heat flux, are made in one part of a usually homogeneous domain, in an attempt to recover the location and shape of an unknown part of the boundary. The unknown boundary is usually characterized by a specific property such as an equipotential or zero flux.

2.3.1 Cauchy Problems

Many of these problem are formally characterized as Cauchy problems (that is recovering a potential from Cauchy data), and as such uniquely solvable. However the Cauchy problem is also ill-posed making making it a very hard inverse problem to solve.

The formulation typically involves Fredholm equations of the first kind which are classical inverse problems[88]. The Cauchy problem has been studied extensively, mainly on bounded domains in 2D[89]. It appears in Electrical Impedance Tomography (EIT)[90], crack detection[91, 92, 93, 94], corrosion detection[95] and continuous steel casting[96]. Akduman and Kress[97] studied 2D shape reconstruction from Cauchy data on an enclosing annulus. In 2D, conditional logarithmic stability estimates for the inverse boundary problem under a regularity assumption on the unknown boundary have been given[98, 99], and this theoretical result has been extended to 3D by Cheng *et al.*[100]. It is also important to note the similarities with the method of Fundamental

Solutions (MFS)[101] because it is essentially a method of solving the Cauchy problem in a chosen domain of approximation.

While the Cauchy problem for the Laplace equation is a near field problem, the reconstruction of scattering objects from far field data is an important class of free boundary problems that has given rise to important developments. The following section describes some of these are that are also are relevant for the near field problem.

2.3.2 Scattering

The use of acoustic, elastic or EM waves to 'see' things that are hidden from normal view is used in many different fields; these include ultrasonic imaging, geophysical survey or strength analysis of construction materials. The brief discussion here is based on Isakov[102], although the book by Colton and Kress[103] is regarded as the standard reference and there is an excellent recent review in[104]. The stationary incoming wave $u^i(x) = \exp(ik\xi \cdot x)$ of frequency k and incident direction $\xi \in \mathbb{R}^3$, $|\xi| = 1$ is scattered by an obstacle. The total field $u = u^i + u^s$ is a solution to the Helmholtz equation (scattering by an obstacle)

$$\nabla^2 u + k^2 u = 0 \quad \text{in} \quad D_e = \mathbb{R}^3 \setminus \overline{D}, \tag{2.5}$$

with the Dirichlet boundary data (soft obstacle D)

$$u = 0 \quad \text{on} \quad \partial D. \tag{2.6}$$

The scattered wave u^s must satisfy the Sommerfeld radiation condition

$$\sigma \cdot \nabla u^s - iku^s = O(r^{-2}) \qquad \text{as } r \to \infty, \tag{2.7}$$

where $r = |x|$, $\sigma = x/r$. It is well known[103] that u^s admits the representation

$$u^s(x; \xi, k) = r^{-1} \exp(ik \cdot \xi r) u^\infty(\sigma, \xi, k) + O(r^{-2}). \tag{2.8}$$

The function u^∞ is called the *scattering amplitude* (or the *(far field) scattering pattern*).

This representation follows from the fact that any solution v to the Helmholtz equation satisfying the radiation condition has the representation by a single layer

potential

$$u^s(x) \;=\; S_\Gamma\phi \;:=\; \int_\Gamma \Phi(x,y)\phi(y)\,\mathrm{d}s(y). \tag{2.9}$$

In which Γ is a surface enclosing the unknown object D and $\Phi(x,y)$ is the fundamental solution.

This representation gives rise to a two step inversion method[105, 106] in which it is assumed that a closed curve Γ wholly contained in D is known. This curve Γ serves as an initial guess. An injective far field operator is defined which takes a density ϕ to its corresponding far-field pattern. This operator is inverted and applied to the measured pattern to find ϕ (on Γ). The second step is then to find ∂D as zero-set of $u^i + S_\Gamma\phi$. The inversion is regularised with standard Tikhonov regularisation[103] and a defect minimization is used for finding a parameterised curve where $u^i + S_\Gamma\phi$ vanishes. The process can then be repeated with the new curve as initial guess.

This method cannot be applied directly to the electrostatic case because the far field resides at ∞ for $k \to 0$. However, similar layer potentials expressions can be derived for the near field. This then forms the basis of the inversion method we have used in chapter 6.

2.3.3 Point Source Method

A method that combines elements of the analytic continuation method of the previous section with aspects of the probe methods discussed below in section 2.4 is the source method, developed by Potthast[107] and used and analysed by several authors[108, 109, 110]. The method converts the measured far field data u^∞ into $w^\infty(\xi,z)$ which represents the artificial scattering amplitude from a point source at z. That is waves originating from the point source z, scattered at the object D, give rise to the far field pattern $w^\infty(\xi,z)$. The theory developed by Potthast[107] provides a mechanism by which to convert the measured u^∞ into $w^\infty(\xi,z)$ for different points z in the domain. For $z \in \overline{D}$, the scattered far field pattern must equal the far field pattern of the point source itself. By collecting points z for which this is true, the object D is recovered.

The actual treatment of the point source method is of course more precise. Kress and Päivärinta[110] use a factorisation for the far field patterns to provide a relatively straightforward derivation of the point source method.

Chandler-Wilde and Lines [108, 109] take the reciprocity relation that is at the

heart of the point source method a step further. They use a point-to-point reciprocity in which the Green's function $G_{1,h^*}(z, x^*)$ for a boundary value problem with known boundaries is used to establish a density function $\varphi^{\alpha}_{x^*}$ on the measurement line γ^*. The density $\varphi^{\alpha}_{x^*}$ is used to establish an (approximate) single layer potential for the total field of a point source at x^*. Moreover it is also used for a single layer potential with the 'real' Green's function $G(z, x^*)$ of a point source at x^* in the presence of the unknown boundary. By reciprocity $G(z, x^*) = G(x^*, z)$ and using a $\varphi^{\alpha}_{x^*}$ for every x^* allows the calculation of $G(x^*, z)$, the total field for the point source at z. The unknown boundary is then identified as the zero (or minimum) in $G(x^*, z)$. Note that this reconstructs the boundary from data on γ^* with just one source at z.

Recently, a similar point-to-point reciprocity relation was used by Pothastt[111] in the application of the point source method for acoustic scattering of bounded domains under single source illumination

2.3.4 Level Sets

Shape evolution is often described and driven within the framework of level sets. Litman *et al*[112] use level sets in the inverse scattering problem, building on earlier work by Santosa[113] and of course the work by Sethian[114, 115]. In level sets the concept of a velocity $V(\boldsymbol{x}, t) = \partial \boldsymbol{x}/\partial t$ plays an important role. Its choice is often a matter of preference. Litman*et al*[112] use the integrand of the Fréchet derivative of the cost function. The work in this thesis in chapter 6 also rests on the notion of level sets. There, the potential near the object itself is used as a velocity and drives the shape deformation.

2.4 Sampling and Probe Methods

The approaches to the free boundary problems described above rely on specific properties of the boundary as an equipotential surface or as a surface with a predetermined derivative. Also they generally employ calculations over the whole boundary and there is no numerical gain in looking for only a part of the boundary or for a low resolution reconstruction. Linear Sampling methods go some way towards addressing these limitations, although they do bring in new limitations of their own.

The basic idea of linear sampling is intuitive and old. A straightforward method

of establishing the size of an oil field is to drill holes over a large area and record the 'hits' on a map. The beauty of the method, and this will become clear as we apply it in chapter 7, is that the concept of a 'hit' is given abstract meaning in a linear function space.

Linear Sampling algorithms have been developed for electromagnetic or acoustic scattering[104, 116], and have also been used for some electrostatic problems[117, 118, 119]. This has resulted in different tests for 'hits', suitable for different problems and boundary conditions. For a recent review see [120]. Initially the methods, termed 'Linear Sampling', tested if the restriction of the fundamental solution (a point source) lies in the range of the forward map, to test for the edge ∂D of the object. The modification by Kirsch[116] used the range of the square root of the forward map instead to test for the whole of the object $\bar{D}$ and the name 'Factorisation Method' has been used for this. Instead of the fundamental solution as test function, Brühl[118] used the dipole term in the multipole expansion of the fundamental solution. In this work we will use the Factorisation Method to show that a test function based on the normal derivative of the Green's function on the plane ∂H can be used to find the object in this case.

Although a distinction is made between the Linear Sampling and the Factorisation Method as tests against the range of the forward map or the range of the square root of the forward map respectively, in the discrete approximation, the setting of any practical application, these ranges coincide. The parallels with the MUSIC algorithm in signal processing in finite dimensions have also been noted[121, 122]. It must be observed that formally the test against the range enjoys greater independence from the boundary condition on the unknown object, it is also true that formally the Linear Sampling Method has no knowledge about the inside of the object. Overall, although the theoretical tools used to derive the Linear Sampling or Factorisation Methods differ significantly and they formally detect different things, in practice the distinction between the methods may be debatable. Numerical experiments based on tests against the range of the square root or tests against the range perform differently[104, 123, 124, 125], but these differences are not easily explained by the formal differences between the methods. We return to this point in the chapter 7

3. THE FORWARD PROBLEM

This chapter contains a discussion of the electrostatics in a half space, that is the equations governing the cross capacitance sensing.

The forward problem is concerned with sensors located on a grounded plane and the change in capacitance between these as a result of the presence of grounded objects in the half space above the ground plane. We shall speak of the capacitance matrix when the emphasis is on the practical aspect of the applied voltage distribution on the plane, that is the voltages of the individual transmitter electrodes and on the currents measured on the ground plane at the receiver electrodes. In other words when we consider the measurement of the capacitive coupling between individual electrodes.

When considering the mapping between function spaces on the plane of the Dirichlet (voltage) and Neumann (current) data, we shall speak of the Dirichlet to Neumann map. We will be mainly concentrating on the *change* in the capacitance matrix or Dirichlet to Neumann map upon the introduction of the object in the half plane.

In this chapter, we will study the classical solution using the second Green identity and an appropriate Green's function. This provides an accessible and relatively self-contained treatment. Then, making a number of simplifications and approximations, closed form expressions are derived for the special circumstance of a sufficiently distant sphere. These provide insight and are useful in establishing an initial guess for reconstruction algorithms as described in chapter 5.

The forward solution is placed in context by a model problem, which does not feature a ground plane and instead of a voltage probe, features a charge probe. Closed form exact solutions are available for this case and these can be compared with single and double layer potential solutions that are applicable to the model problem and the situation of central importance in this thesis. This link is useful from theoretical point of view as well as from a practical point of view to test the computer code of the numerical implementations.

Finally a number of theorems on the solution of the weak problem are presented in a formal style.

3.1 Half Space Electrostatics

In the half space $H := \{x \in \mathbb{R}^3 : x_3 > 0\}$ above an infinite plane ∂H, consider a grounded and conductive object D of class C^2, whose closure is disjoint from the plane ∂H. A potential profile $f_H \in L^2(\partial H)$ is applied on the plane and the forward problem is to find the potential u, specifically to find the *change* in the charge distribution on the halfplane upon introduction of the object.

Formally this forward problem is split into two problems. The first problem defines a Dirichlet problem on an empty half space with L^2 boundary data on an infinite plane ∂H. The second problem defines an external Dirichlet problem with L^2 boundary data on a bounded object D. Throughout we denote the solutions to the first and second problem with u_0 and u_δ respectively. To ensure uniqueness it is important to set a boundary condition at infinity. This is straight forward for the second problem by demanding that the potential vanishes at infinity and decays as least as $1/|x|$ as $x \to \infty$. For the first problem however, this condition is not sufficient as the L^2 boundary data itself may not vanish at infinity. For instance a square integrable function can be constructed of increasingly sparse, narrow spikes that grow without bound as $x \to \infty$. Care must be taken in considering in what sense a solution of the Laplace equation can both satisfy such a function as boundary data and satisfy a condition at infinity such that the solution is unique. Following the description by Meyer and Coifman[126], we use a condition on the behaviour of the potential on cones standing tip-down on the plane ∂H. For some $M \in (0, 1)$ define the cone $\Gamma(x)$, $x \in \partial H$, by $\Gamma(x) = \{y : |y_3 - x_3| > M|x - y|\}$. The precise choice of M does not affect the result we use from [126]. The maximal function $u^\star$ is defined as:

$$u^\star(x) = \sup_{y \in \Gamma(x)} |u(y)| \tag{3.1}$$

With that the forward problem is defined as follows.

I. Given $f_H \in L^2(\partial H)$, find $u_0 \in C^2(H)$ such that

$$\nabla^2 u_0 = 0 \qquad \text{in } H, \tag{3.2}$$

where u_0 satisfies a Dirichlet boundary condition on ∂H in the L^2 sense given by

$$\lim_{x_3 \downarrow 0} \int_{\partial H} |f_H(x_1, x_2) - u_0(x_1, x_2, x_3)|^2 \, \mathrm{d}x_1 \, \mathrm{d}x_2 = 0, \tag{3.3}$$

and

$$u^\star \in L^2(\partial H). \tag{3.4}$$

II. Given $f_D \in L^2(\partial D)$, find $u_\delta \in C^2(\bar{H} \backslash \bar{D})$ for

$$\nabla^2 u_\delta = 0 \qquad \text{in } H \setminus \bar{D}. \tag{3.5}$$

Where u_δ satisfies a Dirichlet boundary condition on ∂D in the L^2 sense given by

$$\lim_{h \downarrow 0} \int_{\partial D} |f_D(x) - u(x + h\nu(x))|^2 \, \mathrm{d}s(x) = 0, \tag{3.6}$$

and

$$u_\delta(x) = \quad 0 \qquad\qquad \text{on } \partial H \tag{3.7}$$

$$u_\delta(x) = \quad O(1/|x|) \qquad \text{for } |x| \to \infty, \tag{3.8}$$

in particular $c(x)$ is defined as

$$c(x) = \frac{\partial u_\delta}{\partial n} \qquad \text{on } \partial H. \tag{3.9}$$

The solution to the overall forward problem is found by setting

$$f_D = u_0|_{\partial D},$$

in which case the solution to the overall problem is given by

$$u(x) = u_0(x) + u_\delta(x), \quad x \in H \setminus \bar{D}.$$

We will refer to $\frac{\partial u}{\partial x_3}$ as the surface charge density and note that c, given by (3.9), is the perturbation in the surface charge density due to introduction of the conductive object D.

The half space problem (problem I) and the exterior Dirichlet problem (Problem II) are well known and well researched problems. A standard body of literature exists for these. In this thesis we have relied mainly on Folland[127], particularly §2G, §2I, and Kress[105], particularly chapter 6, which discusses boundary data in L^2 in some detail. The constraint (3.4) plays the role of a boundary condition at infinity for problem I, as described in chapter 15 of the book by Meyer and Coifman[126]. The treatment in this book goes well beyond the scope of this thesis and discusses potential theory on (unbounded) Lipschitz domains with L^2 boundary data. The theory on boundary data in L^p spaces with $1 \leq p \leq \infty$ is discussed in greater detail by Taibleson[128] and extended to Lipschitz functions in L^p spaces. The case for bounded Lipschitz domains is discussed by Verchota[129].

Exterior Dirichlet problems for unbounded surfaces for the Helmholtz equation are also an area of active research for applications like ground penetrating radar or wave propagation over outdoor ground and sea surfaces. Uniqueness results have been established for a number of smoothness constraints on the surface and with the boundary data in L^p spaces with $p \geq 1$, see [130, 131, 132] and references therein.

These results would allow the formulation of the forward and inverse problem for more general objects than the smooth objects of class C^2 that we have required. However because our main objective here is the inverse problem of recovering the (approximate) object shape from measurements we have chosen to keep the forward problem relatively simple.

The solution to the forward problem is presented in rigorous style in section 3.3. In the following sections we use Green's Theorem in a narrative style to solve a slightly simpler problem, namely the classical problem for compactly supported continuous boundary data. This is done to develop the forward solution in a manner that is more familiar to engineers while, hopefully, satisfying the mathematical requirement for rigour in section 3.3.

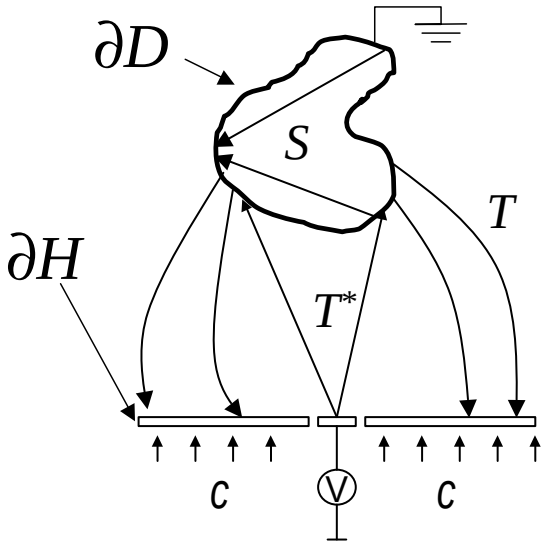

Fig. 3.1: Cross Capacitance Sensor Geometry

3.1.1 Electrodes in a ground plane

In this section the solution to the classical forward problem I is developed in a narrative fashion. This is done using the second and third Green identities and the Green's function for the half space $G(\cdot, y)$, the potential from a unit charge at y and its (negative) image at y', for which $y'_3 = -y_3$,

$$G(x, y) = \frac{1}{4\pi|x - y|} - \frac{1}{4\pi|x - y'|}. \tag{3.10}$$

For the purpose of the narrative, we limit ourselves in this section to boundary data on ∂H that is compactly supported and twice continuously differentiable. That is $f_H \in C_0^2(\partial H)$. The derived expressions are in fact valid for $f_H \in L^2(\partial H)$, as will be shown in the last section of this chapter.

The Green's function has two important properties. One is that

$$G(\cdot, y)\Big|_{\partial H} = 0, \tag{3.11}$$

and, for a continuous function $w(x)$,

$$\int_H w(x)\nabla^2 G(x,y)\,\mathrm{d}x = -w(y). \tag{3.12}$$

The first property is obvious. The second property can be found in, for instance, Kress[105]. We consider initially the bounded domain of the half ball B_R formed by a disk of radius R in the plane ∂H and a half sphere with the same radius, above it. The second Green identity [5, 6] states that

$$\int_{B_R} (w\nabla^2 v - v\nabla^2 w)\,\mathrm{d}x = \int_{\partial B_R} \left(w\frac{\partial v}{\partial n} - v\frac{\partial w}{\partial n} \right) \mathrm{d}s. \tag{3.13}$$

Because f_H is compactly supported when solving problem I we are looking for a harmonic solution u_0 for which $u_0 = O(1/|x|),\ x \to \infty,\ x_3 > 0$ and $\partial u/\partial n = O(1/|x|^2),\ x \to \infty, x_3 > 0$ (by proposition 2.75 of Folland[127]). It can also be shown that that $G(\cdot, y) = O(1/|x|^2),\ x \to \infty$ and $\partial G(\cdot, y)/\partial n = O(1/|x|^3),\ x \to \infty$ (cf. proposition C.1 on page 169). Hence for $v = G(\cdot, y)$ the surface integral over the top half sphere vanishes as $R \to \infty$, so that

$$\int_H (u_0\nabla^2 G(\cdot,y) - G(\cdot,y)\nabla^2 u_0)\,\mathrm{d}x = \int_{\partial H} \left(u_0\frac{\partial G(\cdot,y)}{\partial n} - G(\cdot,y))\frac{\partial u_0}{\partial n} \right) \mathrm{d}s. \tag{3.14}$$

Through the properties of u_0, the boundary conditions and the properties of G, equation (3.14) can be solved for u_0. The second term in the left hand integral can be dropped because u_0 is harmonic in H, that is $\nabla^2 u_0 = 0$. With the properties (3.11) and (3.12), and noting that $\partial G/\partial n = -\partial G/\partial x_3$, equation (3.14) becomes the so-called third Green identity

$$u_0(x) = \int_{\partial H} f_H(y)\frac{\partial G(x,y)}{\partial y_3}\,\mathrm{d}s(y). \tag{3.15}$$

The charge distribution on the ground plane induced by the voltage distribution

f_H is given by

$$\begin{aligned} c_0(x) := -g_H &:= \frac{\partial u_0(x)}{\partial x_3} \\ &= \frac{\partial}{\partial x_3}\int_{\partial H} f_H(y)\frac{\partial G(x,y)}{\partial y_3}\,\mathrm{d}s(y) \qquad (3.16) \\ &= \oint\!\!\!\!\!\!-_{\partial H} f_H(y)\frac{\partial^2 G(x,y)}{\partial x_3 \partial y_3}\,\mathrm{d}s(y) \\ &= \oint\!\!\!\!\!\!-_{\partial H} f_H(y)\left[\frac{1}{2\pi|x-y|^3}\right]\mathrm{d}s(y) \qquad (3.17) \\ &:= \Lambda_0 f_H. \end{aligned}$$

In this the integral $\oint\!\!\!\!\!\!-$ must be interpreted in the Hadamard finite part sense[133, 134], that is as the finite component of the corresponding divergent integral. For instance, if $f_H = 1$ on a unit disk around the origin and $f_H = 0$ outside the disk;

$$\begin{aligned} c_0(0) &= \oint\!\!\!\!\!\!-_{\partial H} f_H(y)\left[\frac{1}{2\pi|y|^3}\right]\mathrm{d}s(y) \\ &= \lim_{\epsilon\to 0}\left\{\int_\epsilon^1 \frac{1}{|\rho|^3}\rho\,\mathrm{d}\rho - \frac{1}{\epsilon}\right\} \\ &= -1. \qquad (3.18) \end{aligned}$$

In general f_H must satisfy a smoothness constraint for (3.17) to be calculable, $f_H \in (L^2(\partial H) \cap C^{1,\alpha}(\partial H))$ would be sufficient[135]. A well established literature exisits on the numerical evaluation of integrals of the type (3.17), see for instance[136].

3.1.2 The charge image of an object

A grounded object D with smooth boundary of class C^2 is now introduced as a hand phantom into the space above the plane, and Green's theorem is applied to the space $H\backslash D$. Writing $u = u_0 + u_\delta$, the result is

$$u(x) = \int_{\partial H} f_H(y)\frac{\partial G(x,y)}{\partial y_3}\,\mathrm{d}s(y) + \int_{\partial D} G(x,y)g_D(y)\,\mathrm{d}s(y)\ , \quad x \in H \setminus \bar{D}, \qquad (3.19)$$

in which $g_D = \partial u/\partial n$ is the charge distribution on the object. From (3.15) we see that the first term in (3.19) is $u_0(x)$ and the second term $u_\delta(x)$. The fact that $u(x) = 0$

on ∂D was used to remove the double layer potential contribution on ∂D in the first term on the RHS of (3.14). Equation (3.19) can also be used to evaluate $u(x)$ on the boundary of the object, although in that case a factor $1/2$ needs to be placed in front of the $u(x)$ [137]. The charge distribution g_D is found by solving (3.19) for $u(x) = 0$ on ∂D. This gives

$$-\int_{\partial D} G(x,y) g_D(y)\,\mathrm{d}s(y) = \int_{\partial H} f_H(y) \frac{\partial G(x,y)}{\partial y_3}\,\mathrm{d}s(y)\,,\ x \in \partial D. \tag{3.20}$$

The presence of the object decreases the charge density on the ground plane. We refer to this *decrease* as a charge image $c(x)$. It is given by the normal derivative of the second term in (3.19) and is written as

$$c(x) := \frac{\partial u_\delta}{\partial n} = -\int_{\partial D} \frac{\partial G(x,y)}{\partial x_3} g_D(y)\,\mathrm{d}s(y)\,,\ x \in \partial H, \tag{3.21}$$

where the negative sign appears because x_3 points into the half space. $c(x)$ is measured as a decrease in capacitance between the transmitter and electrodes in the ground plane.

We define operators $S : L^2(\partial D) \to L^2(\partial D)$, $T : L^2(\partial D) \to L^2(\partial H)$ and $T^* : L^2(\partial H) \to L^2(\partial D)$ by

$$\begin{aligned}
(Sg_D)(x) &= \int_{\partial D} G(x,y) g_D(y)\,\mathrm{d}s(y) \qquad x \in \partial D, && (3.22)\\
(Tg_D)(x) &= \int_{\partial D} \frac{\partial G(x,y)}{\partial x_3} g_D(y)\,\mathrm{d}s(y), \\
&= \int_{\partial D} \frac{y_3}{2\pi|x-y|^3} g_D(y)\,\mathrm{d}s(y), \quad x \in \partial H, && (3.23)\\
(T^* f_H)(x) &= \int_{\partial H} \frac{\partial G(x,y)}{\partial y_3} f_H(y)\,\mathrm{d}s(y), \\
&= \int_{\partial H} \frac{x_3}{2\pi|x-y|^3} f_H(y)\,\mathrm{d}s(y), \quad x \in \partial D. && (3.24)
\end{aligned}$$

The kernel $G(x,y)$ of S is real-valued, symmetric and weakly singular. S is therefore compact and the inverse on L^2 is unbounded. However, for sufficiently smooth boundary data f on ∂D, the single layer potential $f = S\phi$ uniquely solves the internal and external Dirichlet problem[127]. Therefore, as will be shown in more detail later (lemma 3.4), the extension of S to an operator $\tilde{S} : H^{-1/2}(\partial D) \to H^{1/2}(\partial D)$, creates an

operator that is bijective and has a bounded inverse[117, 138]. We can write (3.21) as

$$c = \Lambda_D f_H = T\tilde{S}^{-1}T^* f_H. \tag{3.25}$$

Both T and T^* have square integrable C^∞ kernels (∂D and ∂H are non intersecting surfaces), operator T^* is sufficiently smoothing and the operator $\Lambda_D : L^2(\partial H) \rightarrow L^2(\partial H)$ in (3.25) well defined. Further detail on this is provided in section 3.3 below.

3.1.3 Alternative Factorisation

The physical picture used in the solution above is that of a conductive object in which a charge redistribution takes place to neutralize the applied potential from the transmitter at the surface of the object. The solution in obtained by solving an integral equation for the charge distribution g_D and finding the charge image from that. An alternative approach is to imagine instead a distribution ϕ of dipoles on the surface that acts to neutralise the applied potential. While this physical picture may appear unnecessarily complex, the method has a sound mathematical foundation and certain advantages.

Defining integral operator $\tilde{K} : L^2(\partial D) \rightarrow L^2(\partial D)$ as

$$(\tilde{K}\phi)(x) = \int_{\partial D} \left[\frac{\partial G(x,y)}{\partial \nu(y)} + \frac{1}{|x - x_c|} - \frac{1}{|x - x_c'|} \right] \phi(y)\, \mathrm{d}s(y), \quad x \in \partial D, \tag{3.26}$$

in which $x_c \in D$ and x_c' is its mirror point in the plane ∂H. The moment ϕ is found from

$$(\tfrac{1}{2}I + \tilde{K})\phi = \int_{\partial H} f_H(y) \frac{\partial G(x,y)}{\partial y_3}\, \mathrm{d}s(y)\,, \; x \in \partial D. \tag{3.27}$$

The RHS is the same as (3.20). The factor $\frac{1}{2}I$ takes account of the jump relation of the integral operator $\tilde{K}$ at the surface. Using ϕ obtained from (3.27), the charge image is now calculated from

$$\begin{aligned} c(x) &= \int_{\partial D} \left[\frac{\partial^2 G(x,y)}{\partial x_3 \partial \nu(y)} + \frac{2x_{c3}}{|x - x_c|^3} \right] \phi(y)\, \mathrm{d}s(y) \quad x \in \partial H & (3.28) \\ &=: \tilde{M}\phi(x). & (3.29) \end{aligned}$$

The operators $\tilde{K}$ and $\tilde{M}$ are double layer potentials using the normal derivative of

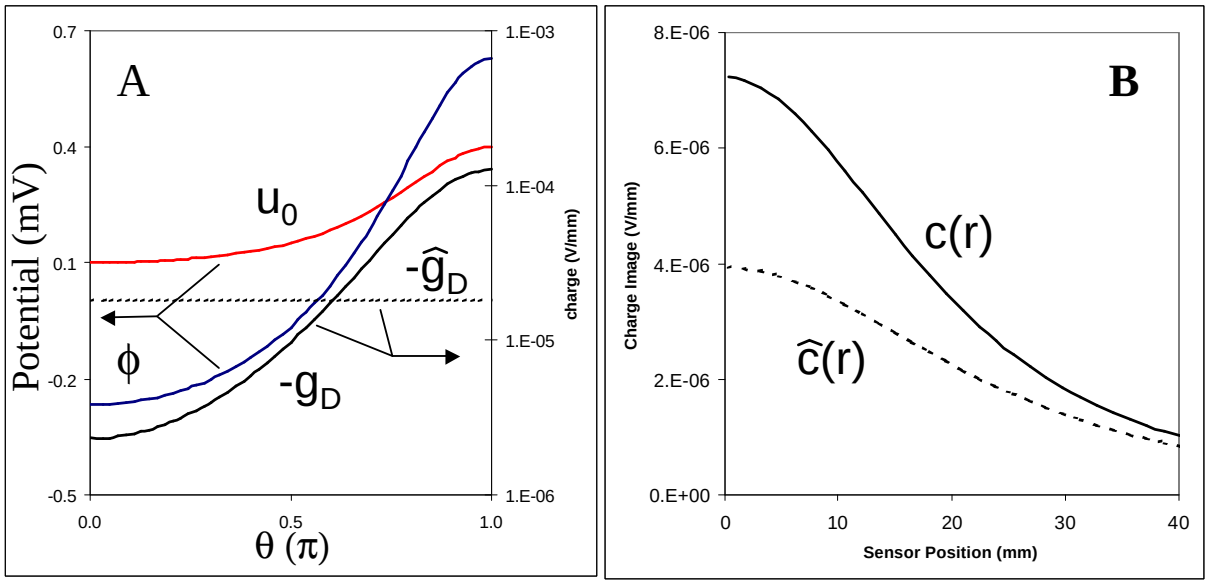

Fig. 3.2: A) Applied potential u_0, the charge distribution $-g_D$ and dipole moment ϕ . Dashed line indicates g_D from (3.32). B) Charge image $c(x)$.

$G(x, y)$. Following Kress [105], the modification $\frac{1}{|x-x_c|} - \frac{1}{|x-x'_c|}$ is required because without it (3.27) not uniquely solvable. The formal demonstration that (3.27) and (3.28) solve the forward problem is given later in this chapter.

With this we will obtain the alternative factorisation

$$\Lambda_D = \tilde{M}(\tfrac{1}{2}I + \tilde{K})^{-1}T^*. \tag{3.30}$$

The operator combination $\tilde{M}(\frac{1}{2}I + \tilde{K})^{-1}$ takes applied potential on the surface of the object to the charge image on the plane and is the solution operator of the exterior Dirichlet problem on the object. This solution operator is important in the discussion of the reconstruction method discussed in chapter 7 and is denoted as $\mathcal{G} : L^2(\partial D) \to L^2(\partial H)$. It is partly because of this that (3.30) is important. The modified double layer potential is further discussed in section 3.2.4 below, while the formal theorem for (3.30) is given in section 3.3 below as theorem 3.7.

To illustrate the forward solution a little, figure 3.2 takes the example of a sphere with radius a_o=10mm, located 30mm above a single transmitter of unit voltage. Shown in figure 3.2A are the applied potential u_0, the RHS of (3.20) and (3.27), the charge density $-g_D$ found by solving (3.20), and dipole moment ϕ obtained from (3.27). The

abscissa is the azimuth with the positive z-axis from the centre of the sphere, that is the point on the sphere closest to the plane (the front), is on the right . Shown in figure 3.2B is the charge image $c(x)$ calculated from (3.25) and (3.28). The dashed lines illustrates an approximate result described in the next section. Note that the dimension of the surface charge g_D and the charge image (capacitance change) is V/mm, that is the dimension of field. Details of the implementation are discussed in the next chapter.

3.1.4 Uniformly Charged Sphere

In this section we derive an approximate expression that aids our intuition for the forward problem. In particular we are interested how the capacitive coupling between electrodes in the ground plane (cf (3.17)) behaves with electrode size and spacing and how the charge image (cf (3.21)) behaves with object size and distance.

If we take, for illustration, a unit area transmitter at distance $|x| = d$ much larger than the diameter of the transmitter, then (3.17) becomes approximately

$$\hat{c}_0(x) = \frac{1}{\pi d^3}. \tag{3.31}$$

Hence, for electrodes in a ground plane, the capacitance falls off with the cube of the distance and scales with electrode area. We can compare this with the situation in free space and spherical electrodes where the capacitance is inversely proportional to the separation and scales linearly with electrode radius [34].

We now consider a small spherical object of radius a_o at position z above the ground plane, $z_3 >> a_o$ and a unity area transmitter at unity potential. If we assume that the object is sufficiently far removed from the transmitter electrode so that the charge density can be taken as constant, (3.20) becomes (cf. Lemma (C.2))

$$-\hat{g}_D a_o = \frac{z_3}{2\pi r_t^3}, \tag{3.32}$$

in which r_t is the distance between transmitter and object. Figure 3.2A compares the level of g_D calculated with this equation with the 'real' charge density from (3.20). Using the approximate $\hat{g}_D$ in (3.21) we obtain

$$\hat{c}(x) = \frac{a_o z_3^2}{\pi r_t^3 r_r^3}, \tag{3.33}$$

in which r_r is the distance between receiver (located at x) and object. We can see that the relative change in capacitance between the two electrodes, caused by an object above the ground plane, is inversely proportional to the altitude of the object above the ground plane and proportional to the size of the object. Considering the presence of the ground plane, and the cube power in (3.31), that is a surprisingly weak dependence. Figure 3.2B illustrates the charge image calculated with (3.33). As can be seen the $\hat{c}(x)$ has a similar shape to $c(x)$ although the approximate expression underestimates the capacitance change by about half. For objects further removed from the sensor plate the approximation becomes better and the agreement between $\hat{c}(x)$ and $c(x)$ improves considerably.

The expression (3.33) is useful in finding an initial guess for an unknown object as will be discussed in section 5.1.

3.2 A Model Problem

To put the solution of forward problem into context, we study a simpler, text book problem for which closed form solutions are available as well as the methods used above. This allows further study of the different solutions and enables a validation of the numerical implementation in the next chapter.

We consider a unit point charge at the origin and a grounded, conducting sphere that does not enclose the origin, centred at $z = (0, 0, z_3)$ and radius a. $(a < z_3)$.

The potential from the point charge is given by

$$u_0(x) = \frac{1}{4\pi|x|}. \tag{3.34}$$

Because the sphere is grounded, a net charge is created on the sphere to neutralise the potential from the point source. We are interested in the forward problem of calculating the change potential u_δ outside the sphere, defined by the relation

$$u_\delta(x) := u_0(x) - u(x),$$

in which u is the total field of point charge and grounded sphere. That is, we want to

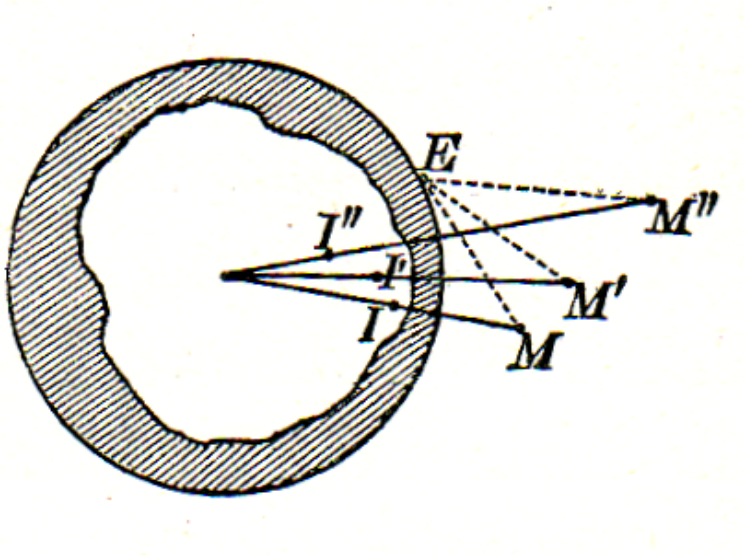

Fig. 3.3: Illustration from Lord Kelvin's 1849 Stockholm paper[139] on *Insulated sphere subject to the influence of a body of any form electrified in any given manner*, showing a solution by the method of images

solve a boundary value problem for the Laplace equation in which

$$\Delta u_\delta(x) = 0 \qquad x \in \mathbb{R}^3 \backslash \bar{D} \tag{3.35}$$

$$u_\delta(x) = f(x) \qquad x \in \partial D \tag{3.36}$$

$$u_\delta(x) \to 0 \qquad \text{for } x \to \infty. \tag{3.37}$$

The function $f(x) = u_0|_{\partial D}$ is the restriction to the boundary of the harmonic functions u_δ and u_0.

The problem is an old one, as the illustration in figure 3.3 from an 1849 paper by Lord Kelvin[139] suggests, although not all solution methods presented below are equally old. Below we present four different methods of solving this problem which reveal different aspects of the forward problem. The numerical agreement in the answers calculated with the different methods validates the implementation of the different integral operators in our c++ code. Details of this implementation are discussed in the next chapter.

3.2.1 Image Charge

The simplest, text book, method to solve the problem is to place an image charge of strength $-a/z_3$ inside the sphere at $(1 - a^2/z_3^2)z$. That is the change potential is given

by

$$u_\delta(x) = \frac{a}{z_3}\frac{1}{4\pi|x-(1-a^2/z_3^2)z|}. \tag{3.38}$$

It is straightforward that the difference between (3.34) and (3.38) is indeed zero on ∂D. Said differently, (3.38) is harmonic in $\mathbb{R}^3\backslash D$ and satisfies boundary conditions (3.36) and (3.37). With the uniqueness of the Dirichlet boundary value problem this means that (3.38) is the unique solution to the problem.

Figure 3.4 illustrates the potential u_δ in the $x_3 = 0$ plane (the plane that would contain the ground plane) and the $x_1 = 0$ plane (to see how u_δ behaves as we approach the surface of the sphere). They are labelled 'Image Charge'.

3.2.2 Poisson Kernel

The problem can be treated more generally using Green's theorem and a Green's function suitable for the specific geometry here. The function

$$G_\circ(x,y) = \frac{1}{4\pi|x-y|} - \frac{a}{|y-z|}\ \frac{1}{4\pi\left|x-\left(1-\frac{a^2}{|y-z|^2}\right)(y-z))\right|} \tag{3.39}$$

has the properties $G_\circ(\cdot,y)|_{\partial D} = 0$ and $\nabla_x^2 G_\circ(x,y) = -\delta(x-y)$ for $x \in \mathbb{R}^3\backslash D$. Applying Green's Theorem in a manner similar to that in section 3.1.1 to u_δ and $G_\circ$ over $\mathbb{R}^3\backslash D$, gives

$$u_\delta(x) = \int_{\partial D} \frac{\partial G_\circ}{\partial \nu} f(y)\,\mathrm{d}s(y), \tag{3.40}$$

in which

$$\frac{\partial G_\circ}{\partial \nu} = \frac{|a^2-|x-z|^2|}{4\pi a|x-y|^3}, \quad x \in \mathbb{R}^3\backslash D,\ \ y \in \partial D. \tag{3.41}$$

The term $\partial G_\circ/\partial\nu$ is called the 'Poisson Kernel' for which Poisson derived an expression by a different method in 1813 (see [140], p 360).

The kernel is harmonic and the expression (3.40) shows that u_δ is harmonic in $\mathbb{R}^3\backslash\bar{D}$. Is is also possible to demonstrate directly that $u_\delta(x \to \partial D) = f(x)$ if $f(x) \in C(\partial D)$ or that boundary condition (3.6) is satisfied if $L^2(\partial D)$ (see Theorem 2.48 in Folland[127] Hence (3.40) solves the boundary value problem.

Because the solution to the boundary value problem (3.35), (3.36) is unique, (3.38) and (3.40) are equal. Figure 3.4 illustrates this numerically by plotting the change

potential u_δ, calculated from (3.40), it is labelled 'Poisson Kernel'. As can be seen, the two methods discussed so far agree in the exterior of D ($z < 20$). Near the boundary of the object, the calculated potential displays a dip which is due to the limitations of the piecewise constant quadrature use to calculate (3.40). This is further discussed in the next chapter

The advantage of (3.40) is that it is valid for general boundary data $f(x) \in C(\partial D)$, provided $f(x)$ is restriction of a function harmonic in D. By the divergence theorem this means that

$$\int_{\partial D} f(x)\,\mathrm{d}s(x) = 0.$$

The particular benefit of using the normal derivative of $G_\circ$ in the kernel of the integral operator in (3.40) is that the double layer potential is continuous for as x approaches the boundary ∂D. This is not the case for the more general double layer potentials we encounter later, in which case jump relations at the boundary are used to describe the discontinuities.

These two methods can obviously only be used if the shape of the object admits a treatment by the method of image charges or when a closed form Green's function is available. We now turn to methods that can be used for more arbitrary shapes.

3.2.3 Single Layer Potential

One approach is, inspired by the physics of the problem, to work out the charge density q on the sphere that is required to neutralize the external potential. That is it is given by

$$\int_{\partial D} \Phi(x,y)q(y)\,\mathrm{d}s(y) = f(x), \qquad x \in \partial D. \tag{3.42}$$

This is a Fredholm equation of the first kind, solvable for $f(x) \in C^1(\partial D)$, see lemmas 3.3 and 3.4 on page 58 below. For $f(x) = u_0|_{\partial D}$ this condition is easily satisfied. The charge density can be calculated with a number of numerical techniques and the change in potential is found from

$$u_\delta(x) = \int_{\partial D} \Phi(x,y)q(y)\,\mathrm{d}s(y). \tag{3.43}$$

This function is obviously harmonic and satisfies the boundary condition as the single layer potential is continuous in $\mathbb{R}^n$ by theorem 6.14 in Kress[105].

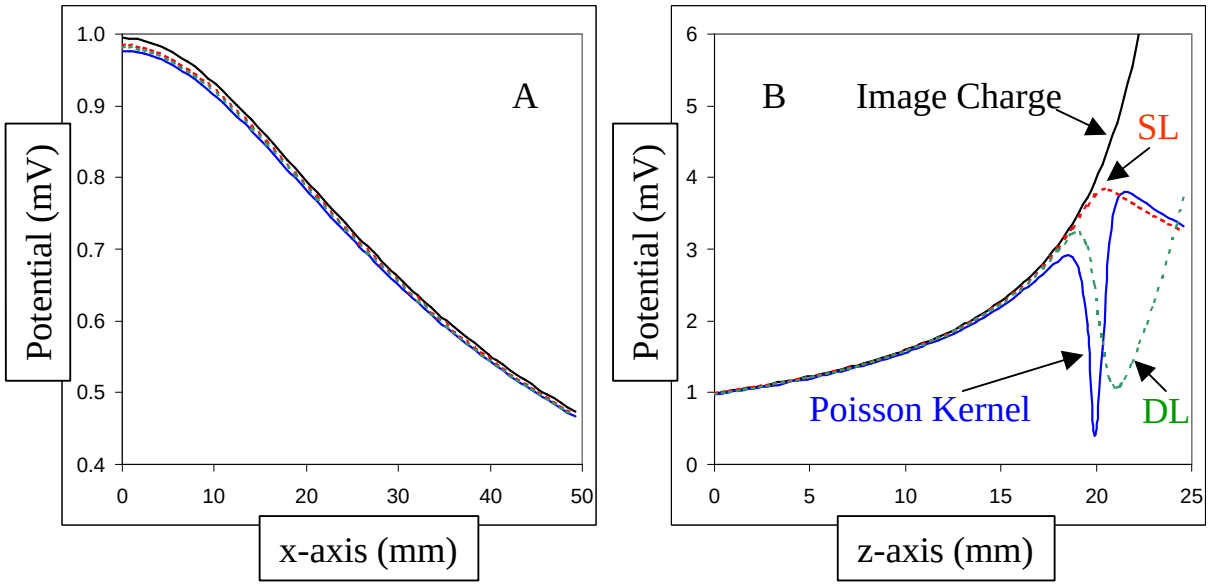

Fig. 3.4: Solutions to a simplified problem using 4 methods

The change in potential u_δ calculated with this method is also illustrated in figure 3.4, labelled 'SL' and again agrees for $z < 20$) with the other two methods. As can be seen, the derivative of the (3.43) shows a jump at the boundary.

3.2.4 Modified Double Layer Potential

Whilst the single layer potential method is valid for arbitrary objects with sufficiently smooth boundary (class C^2), the Fredholm equation of the first kind (3.42) cannot be solved if we want to consider $f \notin C^1(\partial D)$, or $f \notin H^{1/2}(\partial D)$. To find a solution of the external Dirichlet problem for a broader class of boundary data, we look for a Fredholm equation of the second kind.

One way to achieve this would be to write for the potential

$$u_\delta(x) = \int_{\partial D} \left[\frac{\partial \Phi(x,y)}{\partial n_y} \right] \phi(y)\, \mathrm{d}s(y). \tag{3.44}$$

The kernel of this double layer potential is discontinuous at the surface and the density

ϕ is found from a jump relation which satisfies

$$\frac{1}{2}\phi(x) + \int_{\partial D} \left[\frac{\partial \Phi(x,y)}{\partial n_y}\right] \phi \, \mathrm{d}s(y) = f(x), \quad x \in \partial D. \tag{3.45}$$

This equation for ϕ is the Fredholm equation of the second kind we were looking for

$$(\tfrac{1}{2}I + K)\phi = f. \tag{3.46}$$

However, for this to be uniquely solvable (invertable) for all $f \in C(\partial D)$ (classical result) or $f \in L^2(\partial D)$ (weak result), the null space must be trivial. This is because, by the Fredholm alternative, if the null space has a finite dimension, so has the adjoint $I + K'$ and f in equation (3.46) must be perpendicular to that null space of the adjoint. As pointed out by Kress [105], this condition that cannot be expected to be satisfied by arbitrary boundary data, and indeed may not be satisfied by the restriction to the boundary of potential (3.34).

Unfortunately (3.46) does have a non-trivial, finite null space. For $\phi = c$, we have, by proposition C.5, $K\phi = -\frac{1}{2}c$ and $f = 0$. Hence $\mathcal{N}(I + K) = \mathrm{span}\{1\}$, i.e. functions that are constant on ∂D.

Following theorem 6.23 in Kress [105], we introduce a modified double layer potential to create a Fredholm equation of the second find with a trivial null space. We write

$$u_\delta(x) = \int_{\partial D} \left[\frac{\partial \Phi(x,y)}{\partial n_y} + \frac{1}{|x-z|}\right] \phi(y) \, \mathrm{d}s(y). \tag{3.47}$$

Now ϕ is found from

$$\frac{1}{2}\phi(x) + \int_{\partial D} \left[\frac{\partial \Phi(x,y)}{\partial n_y} + \frac{1}{|x-z|}\right] \phi \, \mathrm{d}s(y) = f(x), \quad x \in \partial D, \tag{3.48}$$

or

$$(\tfrac{1}{2}I + \tilde{K})\phi = f.$$

The modification term decays to zero as $O(1/|x|$ as $x \to \infty$, whereas the $\partial\Phi/\partial n$ terms decays as $O(1/|x|^2)$ for $x \to \infty$, the homogenous equation $(\frac{1}{2}I + \tilde{K})\phi = 0$ has a solution only if

$$\int_{\partial D} \phi \, \mathrm{d}s = 0.$$

However if ϕ satisfies this relation then also $(\frac{1}{2}I + K)\phi = 0$. Then that also means that ϕ is constant and the only constant it can be is thus zero. Hence the null space is trivial. For more detail on this see the proof of lemma 3.6 on page 61 below.

The potential profile calculated with this method agrees perfectly with that of the other methods as illustrated in figure 3.4 for this case, it is labelled 'DL' and shows a jump at the boundary of the object. The point from which $|x|$ in the modification is measured can be varied inside the object.

3.3 Operator Properties

Having set out in the previous sections the classical solution of the forward problem in a narrative style, this section provides a number of rigorous theorems that support and extend the exposition above. Important in the extension of classical results is the completion of normed spaces and the extension of bounded linear operators on normed spaces to operators on their completions. In broad outline, the spaces of continuous functions are completed in Sobolev spaces. Theorems on operators between Sobolev spaces are then generally proven by demonstrating the relevant property for continuous functions and then relying on the fact that the continuous functions are dense in the Sobolev space, to apply the result to the operator between Sobolev spaces

Theorem 3.1. *The function*

$$u_0(x) = \int_{\partial H} f_H(y) \frac{\partial G(x,y)}{\partial y_3} \, \mathrm{d}s(y) \tag{3.15}$$

satisfies boundary problem I.

Proof. As shown in lemma C.4, the double layer satisfies the boundary condition in (3.3). By theorem 2 on page 260 of [126], (see C.9 and accompanying notes), the double layer potential (3.15) satisfies the maximal constraint (3.4) and the boundary data g 'for almost all $x \in \partial H$'. □

On page 45 of this thesis, the Dirichlet to Neumann map Λ_0, in the absence of an object was shown as a Hadamard finite part integral expression and this can be calculated only if f_H satisfies a smoothness constraint. By implication, the Dirichlet to Neumann map in the presence of an object must satisfy a similar constraint.

However the difference between these two maps, that is the operator Λ_D, which maps Dirichlet data to the change in Neumann data, can be defined as a bounded map $\Lambda_D : L^2(\partial H) \to L^2(\partial H)$. It is indeed compact. We work towards this result (theorem 3.8) in a number of steps. It is well known that integral operators with continuous kernels on bounded domains are compact (eg theorem 2.21 of [105]). However, on unbounded domains the operators may not even be bounded. The compactness, and by implication, boundedness, of the operator T is therefore of interest.

Lemma 3.2. *The operator $T : L^2(\partial D) \to L^2(\partial H)$, defined by*

$$(Tg)(x) := \int_{\partial D} \frac{y_3}{2\pi|x-y|^3} g(y)\,\mathrm{d}s(y), \quad x \in \partial H, \tag{3.49}$$

is compact, injective and has a well defined adjoint given by

$$(T^* f)(x) := \int_{\partial H} \frac{x_3}{2\pi|x-y|^3} f(y)\,\mathrm{d}s(y). \tag{3.50}$$

Proof. We demonstrate that the kernel is square integrable on $(\partial D \times \partial H)$. That is

$$\int_{\partial H}\int_{\partial D} \left|\frac{y_3}{2\pi|x-y|^3}\right|^2 \mathrm{d}s(y)\,\mathrm{d}s(x) < \infty. \tag{3.51}$$

Define on ∂H the bounded disk D_H of radius r, centered at the origin, containing the projection of D on ∂H, and write integral 3.51 as

$$\int_{D_H}\int_{\partial D} \left|\frac{y_3}{2\pi|x-y|^3}\right|^2 \mathrm{d}s(y)\,\mathrm{d}s(x) + \int_r^\infty \int_0^{2\pi} \int_{\partial D} \left|\frac{y_3}{2\pi|x(\rho,\phi)-y|^3}\right|^2 \mathrm{d}s(y)\rho\,\mathrm{d}\phi\,\mathrm{d}\rho. \tag{3.52}$$

Denote by $y = b$ the point on ∂D closest to the ∂H plane and by $y = t$ the point on ∂D furthest from the ∂H plane. Then $x(\rho,\phi) - y > \rho - r + b_3$ for all $y \in \partial D$ and $x \in \partial H \backslash D_H$. Therefore

$$\begin{aligned}
&\int_{\partial H}\int_{\partial D} \left|\frac{y_3}{2\pi|x-y|^3}\right|^2 \mathrm{d}s(y)\,\mathrm{d}s(x) \\
&< \int_{D_H}\int_{\partial D} \left|\frac{t_3}{2\pi b_3^3}\right|^2 \mathrm{d}s(y)\,\mathrm{d}s(x) + \int_r^\infty \int_0^{2\pi} \int_{\partial D} \left|\frac{t_3}{2\pi|\rho - r + b_3|^3}\right|^2 \mathrm{d}s(y)\rho\,\mathrm{d}\phi\,\mathrm{d}\rho \\
&< \infty.
\end{aligned} \tag{3.53}$$

Injectivity follows from the maximum principle. u_δ is harmonic and $u_\delta|_{\partial H} = 0$ by definition, therefore

$$Tg = \left.\frac{\partial u_\delta}{\partial n}\right|_{\partial H} = 0, \tag{3.54}$$

implies that $u_\delta = 0$ in $H\backslash D$ and therefore that $g = 0$.

The adjoint T^* is defined by

$$(Tg, f)_{\partial H} = (g, T^* f)_{\partial D}, \tag{3.55}$$

for every $g \in L^2(\partial D)$ and $f \in L^2(\partial H)$. We see from changing the order of integration

$$\int_{\partial H} f(x) \int_{\partial D} \frac{y_3}{2\pi|x-y|^3} g(y)\,\mathrm{d}s(y)\,\mathrm{d}s(x) = \int_{\partial D} g(x) \int_{\partial H} \frac{x_3}{2\pi|x-y|^3} f(y)\,\mathrm{d}s(y)\,\mathrm{d}s(x)$$

and so the adjoint is defined by

$$(T^* f)(x) := \int_{\partial H} \frac{x_3}{2\pi|x-y|^3} f(y)\,\mathrm{d}s(y).$$

As the adjoint of a bounded and compact operator, this operator is bounded and compact in $L^2(\partial D)$.

□

Lemma 3.3. *The operator $S : L^2(\partial D) \to L^2(\partial D)$ defined by*

$$(Sg_D)(x) : L^2(\partial D) \to L^2(\partial D) = \int_{\partial D} G(x,y) g_D(y)\,\mathrm{d}s(y), \qquad x \in \partial D. \tag{3.56}$$

is compact, self-adjoint and positive definite.

Proof. The kernel $G(x,y)$ of S is singular of order 1 and S is therefore compact by (cf Lemma C.7). In addition $G(x,y)$ is symmetric and real valued making S self adjoint.

Applying Green's theorem (Green's first identity) to the harmonic single layer potential

$$u_\delta(x) = \int_{\partial D} G(x,y) g_D(y)\,\mathrm{d}s(y), \tag{3.57}$$

both inside and outside the object D, gives

$$\int_D |\nabla u_\delta|^2 = \int_{\partial D} u_\delta^- \frac{\partial u_\delta^-}{\partial n}\,\mathrm{d}s, \tag{3.58}$$

and

$$\int_{H\setminus D} |\nabla u_\delta|^2 = -\int_{\partial D} u_\delta^+ \frac{\partial u_\delta^+}{\partial n}\,\mathrm{d}s. \tag{3.59}$$

In the second expression the boundary integral over the plane is ignored because $u = 0$ on ∂H. Adding these two expressions for the energy inside and outside the object

$$\begin{aligned} 0 \le \int_H |\nabla u_\delta|^2 &= \int_{\partial D} u_\delta^- \frac{\partial u_\delta^-}{\partial n} - u_\delta^+ \frac{\partial u_\delta^+}{\partial n}\,\mathrm{d}s \\ &= \int_{\partial D} u_\delta g_D\,\mathrm{d}s \\ &= \int_{\partial D} g_D(Sg_D)\,\mathrm{d}s. \end{aligned} \tag{3.60}$$

This establishes semi positive definiteness. The above steps show that $(g_D, Sg_D)_{L^2(\partial D)} = 0$ implies that $\nabla u_\delta = 0$ everywhere which implies that $g_D = 0$. So S is positive definite. □

The fact that S is positive definite means that $\mathcal{N}(S) = \{0\}$ implying that S has an inverse. The inverse is unbounded on $L^2(\partial D)$ because S is compact and the range is not closed. On the other hand, for sufficiently smooth boundary data f on ∂D, the single layer potential $f = S\phi$ uniquely solves the internal and external Dirichlet problem[127]. The L^2 norm is too weak to provide a closed image of S, which leads to the following assertion.

Lemma 3.4. *There is an extension of $S : L^2(\partial D) \to L^2(\partial D)$, denoted by $\tilde{S} : H^{-1/2}(\partial D) \to H^{1/2}(\partial D)$. This operator is bijective and has a bounded inverse.*

Proof. Theorem 2.31 in [141] implies that S maps $H^{-1/2}$ continuously unto $H^{1/2}$. Injectivity now follows from lemma 3.3. Defining for a given $f \in H^{1/2}(\partial D)$, the functions u^+ and u^- as the solutions to the internal and external Dirichlet problem with boundary data f we have $\partial u_-/\partial n \in H^{-1/2}(\partial D)$ and $\partial u_+/\partial n \in H^{-1/2}(\partial D)$ and so that $g_D := \partial u_-/\partial n - \partial u_+/\partial n \in H^{-1/2}(\partial D)$. At the same time we know that the single layer potential Sg_D satisfies the same Dirichlet and Neumann boundary conditions, that is $u^+|_{\partial D} = u^-|_{\partial D} = Sg_D$, and so $Sg_D = f$. □

Given the boundedness of $\tilde{S}$, (lemma 3.4) the compactness of $S : L^2(\partial D) \to L^2(\partial D)$ (lemma 3.3) would have followed immediately from the compactness of the embeddings $H^{1/2} \hookrightarrow L^2$ and $L^2 \hookrightarrow H^{-1/2}$. This approach was followed by Brühl[118].

For the theory below it is convenient to work with a further operator. We define a mapping $\mathcal{G} : L^2(\partial D) \to L^2(\partial H)$ as the solution operator of the exterior Dirichlet problem which maps the potential on the object to the field on the ground plane.

$$\mathcal{G} : u_\delta\Big|_{\partial D} \to \frac{\partial u_\delta}{\partial x_3}\Big|_{\partial H}. \tag{3.61}$$

The u_δ is equivalent to the scattered field in scattering theory. The charge image $c(x) = -\partial u_\delta/\partial x_3$ is the change in electric field on the ground plane. We point out that $\mathcal{G}$ includes the contribution from the mirror object in the lower halfspace, in line with the definition of the Green's function (3.10), this implies that $u_\delta|_{\partial H} = 0$.

Theorem 3.5. *The solution operator of the external Dirichlet problem $\mathcal{G} : L^2(\partial D) \to L^2(\partial H)$ is compact and injective.*

Proof. There is a z such that the ball $B(z; a)$ of radius a, centered at z encloses the bounded object. $D \subset B \subset H$. By elliptic regularity and the maximum principle, u_δ and $\partial u_\delta/\partial n$ are continuous and bounded on ∂B for L^2 boundary data on ∂D. Hence the solution operators $\mathcal{G}_\circ : L^2(\partial D) \to C(\partial B)$ and $\mathcal{G}'_\circ : L^2(\partial D) \to C^2(\partial B)$, defined by

$$\mathcal{G}_\circ : u_\delta\Big|_{\partial D} \to u_\delta\Big|_{\partial B}, \tag{3.62}$$

and

$$\mathcal{G}'_\circ : u_\delta\Big|_{\partial D} \to \frac{\partial u_\delta}{\partial x_3}\Big|_{\partial B} \tag{3.63}$$

are bounded operators.

By Green's theorem

$$u_\delta(x) = \int_{\partial B} G(x,y)\frac{\partial u_\delta(y)}{\partial n}\,\mathrm{d}s(y) - \int_{\partial B} u_\delta(y)\frac{\partial G(x,y)}{\partial n(y)}\,\mathrm{d}s(y),$$

and therefore

$$\frac{\partial u_\delta}{\partial x_3}\Big|_{\partial H} = \int_{\partial B} T_\circ(x,y)u_\delta(y)\,\mathrm{d}s(y) + \int_{\partial B} K_\circ(x,y)\frac{\partial u_\delta(y)}{\partial n}\,\mathrm{d}s(y). \tag{3.64}$$

In which the kernels are given by:

$$T_\circ(x,y) = \frac{y_3}{2\pi|x-y|^3} \tag{3.65}$$

$$K_\circ(x,y) = \frac{\partial^2 G}{\partial x_3 \partial n} \quad x \in \partial H,\ y \in \partial B(z,a) \tag{3.66}$$

Using the Poisson Kernel for a sphere centred at $z = (0,0,z_3)$ and radius a. $(a < z_3)$ [127].

$$K_\circ(x,y) := \frac{\partial^2 G}{\partial x_3 \partial \nu} = \frac{\partial}{\partial x_3}\left[\frac{a^2-|x-z|^2}{4\pi|x-y|^3} - \frac{a^2-|x-z'|^2}{4\pi|x-y'|^3}\right], \quad x \in \partial H,\ \ y \in \partial D,$$

and using

$$\frac{\partial}{\partial x_3}|x-y|^p = p\ |x-y|^{p-2}(x_3-y_3),$$

we get

$$K_\circ(x,y) = \frac{-2(x_3-z_3)}{4\pi|x-y|^3} - \frac{-2(x_3-z_3')}{4\pi|x-y'|^3} - 3(x_3-y_3)\frac{a^2-|x-z|^2}{4\pi|x-y|^5} + 3(x_3-y_3')\frac{a^2-|x-z'|^2}{4\pi|x-y'|^5}.$$

Noting that $z_3 = -z_3'$ by definition and that on ∂H, we have $x_3 = 0$ and therefore $|x-y| = x-y'$, thus

$$K_\circ(x,y) = \frac{z_3}{\pi|x-y|^3} + 3y_3\frac{a^2-|x-z|^2}{4\pi|x-y|^5}. \tag{3.67}$$

It is now clear that both $T_\circ$ and $K_\circ$ are C^∞ and $o(1/|x|^3)$ for $|x| \to \infty$. It follows that both integral operators in (3.64) are Hilbert-Schmidt. The operator $\mathcal{G}$ is therefore the sum of products of bounded ($\mathcal{G}_\circ$ and $\mathcal{G}_\circ'$) and compact operators and as such compact.

The injectivity follows from the uniqueness of the external Dirichlet problem. For given boundary data $u_\delta \in L^2\partial D$ the potential u_δ is uniquely determined in $H\backslash\overline{D}$. We have $u_\delta|_{\partial H} = 0$ and $\partial u_\delta/\partial x_3|_{\partial H}$ is uniquely determined. Thus $\mathcal{G}$ is injective □

Having established that the external Dirichlet solution operator is compact, we wish to go further and find an explicit operator expression, or factorisation for this. This is done by applying the Fredholm theory to the solution of the Dirichlet problem in this case.

Lemma 3.6. *The operator* $(\frac{1}{2}I + \tilde{K}) : L^2(\partial D) \to L^2(\partial D)$ *introduced in (3.26) and defined by*

$$((\tfrac{1}{2}I + \tilde{K})\phi)(x) = \tfrac{1}{2}\phi(x) + \int_{\partial D} \left[\frac{\partial G(x,y)}{\partial \nu(y)} + \frac{1}{|x - x_c|} - \frac{1}{|x - x_c'|}\right] \phi(y)\, \mathrm{d}s(y) \quad x \in \partial D,$$

in which $x_c \in D$ *and* x_c' *its mirror point, is bijective and has a bounded inverse*

Proof. The operator $\tilde{K}$ is compact and by lemma C.8, $(\frac{1}{2}I + \tilde{K})$ is Fredholm with index zero. We therefore only need to demonstrate that the homogeneous equation $(\frac{1}{2}I + \tilde{K})\phi = 0$ is satisfied only for $\phi = 0$.

Assume that there is a non zero ϕ for which $(\frac{1}{2}I + \tilde{K})\phi = 0$, then the double layer potential given by

$$u_\delta(x) = \int_{\partial D} \left[\frac{\partial G(x,y)}{\partial \nu(y)} + \frac{1}{|x - x_c|} - \frac{1}{|x - x_c'|}\right] \phi(y)\, \mathrm{d}s(y), \tag{3.68}$$

will be zero on ∂D. For $|x| \to \infty$ we have $u_\delta(x) \to 0$, and so, by the maximum principle we have $u_\delta(x) = 0$ in $H \backslash \overline{D}$.

The asymptotic behaviour of the modification term is $O(\cos\theta/|x|^2)$ and using the asymptotic behaviour of the Green's function (cf lemma C.1), we have

$$|x|^2 u_\delta(x) = \cos\theta \int_{\partial D} \phi\, \mathrm{d}s(y) + O\left(\frac{1}{|x|}\right) \tag{3.69}$$

From this, since $u_\delta(x) = 0$ in $H \backslash \overline{D}$, follows $\int_{\partial D} \phi\, \mathrm{d}s = 0$. We now have

$$0 = \tfrac{1}{2}\phi(x) + \int_{\partial D} \left[\frac{\partial G(x,y)}{\partial \nu(y)}\right] \phi(y)\, \mathrm{d}s(y) \quad x \in \partial D. \tag{3.70}$$

Corollary C.6 in the appendix on page 174 implies that this can only be satisfied for $\phi =$ constant and the assumption that ϕ is nonzero leads to a contraction with $\int_{\partial D} \phi\, \mathrm{d}s = 0$. It must therefore be that $\phi = 0$. □

The structure of the proof demonstrates the requirement of the modification term. Without it, the requirement that $\int_{\partial D} \phi\, \mathrm{d}s = 0$ can not be obtained from 3.69 and the equation $(\frac{1}{2}I + \tilde{K})\phi = 0$ can have a non-trivial solution. The use of a modification term like is not universal in the literature. The use here was suggested by Kress[105],

other authors, such as Folland[127] or Verchota[129] avoid its use by a priori limiting boundary data to data with $\int_{\partial D} \phi \, \mathrm{d}s = 0$.

Theorem 3.7. *The factorisation $\mathcal{G} = \tilde{M}(\frac{1}{2}I + \tilde{K})^{-1}$, in which $\tilde{M} : L^2(\partial D) \to L^2(\partial H)$ is defined as*

$$(\tilde{M}\phi)(x) = \int_{\partial D} \left[\frac{\partial^2 G(x,y)}{\partial x_3 \partial \nu(y)} + \frac{2x_{c3}}{|x - x_c|^3} \right] \phi(y) \, \mathrm{d}s(y) \quad x \in \partial H, \tag{3.71}$$

and provides the unique Dirichlet to Dirichlet map from ∂D to ∂H.

Proof. By lemma 3.6, the function $\phi = (\frac{1}{2}I + \tilde{K})^{-1} f$ exists and the double layer potential

$$u_\delta(x) = \int_{\partial D} \left[\frac{\partial G(x,y)}{\partial \nu(y)} + \frac{1}{|x - x_c|} - \frac{1}{|x - x_c'|} \right] \phi(y) \, \mathrm{d}s(y), \tag{3.72}$$

is the unique solution to the boundary value problem for $u_\delta|_{\partial D} = u_0$ on ∂D. Direct differentiation with respect to the normal on the plane ∂H yields the expression for $\tilde{M}$ above □

Theorem 3.8. *The operator $\Lambda_D : L^2(\partial H) \to L^2(\partial H)$ which maps Dirichlet data on ∂H to the change in Neumann Data on ∂H as a result of the presence of object D, is given by operator factorisations $\Lambda_D = T\tilde{S}^{-1}T^* = \mathcal{G}T^* = \mathcal{G}S\mathcal{G}^*$ and is bounded, injective, compact and positive definite.*

Proof. The first factorisation was demonstrated in the narrative leading up to (3.25). Lemma 3.2 demonstrates that T^* is well defined on L^2. With the smoothing properties of T^*, lemma 3.4, this extends $\Lambda_D = T\tilde{S}^{-1}T^*$ to functions in $L^2(\partial H)$. By definition of $\mathcal{G}$ in theorem 3.5 we have $T = \mathcal{G}S$ and consequently $T^* = S\mathcal{G}^*$. Inserting these relations in $T\tilde{S}^{-1}T^*$, we obtain the second and third factorisations.

The product of a compact operator and bounded operator is compact. The product of injective operators is injective. Using the operator properties demonstrated above, the result follows from any of the three factorisations $\Lambda_D = T\tilde{S}^{-1}T^* = \mathcal{G}T^* = \mathcal{G}S\mathcal{G}^*$. Finally we have

$$\int_{\partial H} g \, (\Lambda_D \, g) \, \mathrm{d}s = \int_{\partial H} g \, \mathcal{G}S(\mathcal{G}^* g) \, \mathrm{d}s = \int_{\partial D} (\mathcal{G}^* g) \, S \, (\mathcal{G}^* g) \, \mathrm{d}s > 0. \tag{3.73}$$

□

The injectivity of Λ_D is the property that, given one object D, each charge image $c(x)$ is uniquely associated with one transmitter voltage pattern f_H. The next theorem establishes a different kind of uniqueness.

Theorem 3.9. *For two objects D_1 and D_2 the operators $\Lambda_1 = \Lambda_2$ if and only if $D_1 = D_2$.*

Proof. Suppose $D_1 \neq D_2$ but for any f_H we have $c(x) = \Lambda_1 f_H = \Lambda_2 f_H$. If necessary we can exchange D_1 and D_2 to obtain the nonempty domain $B = (D_1 \cup D_2) \backslash D_2 \neq \emptyset$. Because f_H and $g_H = c_0 + c(x)$ is the same for D_1 and D_2, uniqueness of the Cauchy problem with boundary data on ∂H demands that $u_1 = u_2$ in $H \backslash (D_1 \cup D_2)$. We therefore have $u_2|_{\partial B} = 0$ and so $\Delta u_2 = \nabla u_2 = u_2 = 0$ in B. Call $\partial\Gamma$ the portion of ∂B that does not coincide with ∂D_2. Because $\nabla u_2|_{\partial\Gamma} = 0$, if u_2 grows to a positive value away from Γ, then it must grow to a negative value in a different direction. Hence if u_2 is non-zero in some portion of $H \backslash D_2$, then there must be some portion of $H \backslash D_2$ in which u_2 has opposite sign. This is in contradiction to the maximum principle for harmonic functions if f_H is unipolar. If $u_2 = 0$ in $H \backslash D_2$ then this is not compatible with non-zero f_H. The domain B must therefore be empty and $D_1 = D_2$. □

We note that the last paragraph of this proof could have been simplified my referring to the analyticity of harmonic functions (cf Kress[105], page 70).

Lemma 3.10. *$(f, \Lambda_D f)$ is the energy of introducing the grounded object in the half space.*

Proof. Writing u for the potential in the presence of the object D and u_0 the potential without the object, and using $f = u|_{\partial H} = u_0|_{\partial H}$, as well as the fact that $u|_{\partial D} = 0$, we write

$$\begin{aligned}
(f, \Lambda_D f) &= \int_{\partial H} u \left(\frac{\partial u}{\partial \nu} - \frac{\partial u_0}{\partial \nu} \right) \mathrm{d}s \\
&= \int_{\partial H} u \frac{\partial u}{\partial \nu} \mathrm{d}s - \int_{\partial H} u_0 \frac{\partial u_0}{\partial \nu} \mathrm{d}s \\
&= \int_{H \backslash D} |\nabla u|^2 - \int_H |\nabla u_0|^2. \qquad (3.74)
\end{aligned}$$

□

Note that the semi positive definiteness of Λ_D implies that this energy is positive. It seems intuitive that introducing a smaller object requires less energy than introducing a big object. If the larger object does not overlap the smaller object completely, then this observation may depend on the drive signals used. For example in the case of a single transmitter, bringing a smaller object close to the transmitters may require more energy than a very large object far away. However, if the smaller object is contained inside the bigger one, then the larger object will require more energy for any drive signal used. This observation leads to the following monotonicity property for Λ_D, which will be useful later.

Theorem 3.11. *For two objects D_i and D_e, such that $D_i \subset D_e$, the operator $\Lambda_{D_e} - \Lambda_{D_i}$ is positive definite.*

Proof. Consider the solutions u_i and u_e to the boundary value problem in the presence of D_i and D_e respectively. That is $u_0|_{\partial H} = u_i|_{\partial H} = u_e|_{\partial H}$ and $u_i|_{\partial D_i} = 0$, $u_e|_{\partial D_e} = 0$. Before proceeding to the main demonstration, we need some inequalities.

$$\begin{aligned} \int_{\partial D_e} u_i \frac{\partial u_i}{\partial \nu} \, \mathrm{d}s - \int_{\partial D_e} u_i \frac{\partial u_e}{\partial \nu} \, \mathrm{d}s &= \int_{\partial D_e} (u_i - u_e) \frac{\partial (u_i - u_e)}{\partial \nu} \, \mathrm{d}s \\ &= \int_{H \setminus D_e} |\nabla (u_i - u_e)|^2 \\ &\geq 0 \end{aligned} \tag{3.75}$$

Where the normal ν points into the domain D_e. The second step above uses the fact that $(u_i - u_e)|_{\partial H} = 0$. We also need

$$\int_{\partial D_e} u_i \frac{\partial u_i}{\partial \nu} \, \mathrm{d}s = - \int_{D_e \setminus D_i} |\nabla u_i|^2 < 0. \tag{3.76}$$

Where the minus sign appears because ν again points into the domain $D_e \setminus D_i$. The inequality is strict because the domain $D_e \setminus D_i$ is nonempty. Combining these two inequalities we obtain

$$\int_{\partial D_e} u_i \frac{\partial u_e}{\partial \nu} \, \mathrm{d}s < 0. \tag{3.77}$$

We now proceed towards the main result by applying the second Green's identity to

u_i and u_e in the space $H\backslash D_e$, and write

$$\int_{\partial[H\backslash D_e]} u_i \frac{\partial u_e}{\partial \nu} - u_e \frac{\partial u_i}{\partial \nu} = \int_{[H\backslash D_e]} u_i \Delta u_e - u_e \Delta u_i \tag{3.78}$$

The RHS is zero because both u_i and u_e are harmonic in $H\backslash D_e$, using $u_0|_{\partial H} = u_e|_{\partial H} = u_i|_{\partial H}$;

$$\int_{\partial H} u_0 \left(\frac{\partial u_e}{\partial \nu} - \frac{\partial u_i}{\partial \nu} \right) \mathrm{d}s + \int_{\partial D_e} u_i \frac{\partial u_e}{\partial \nu} \mathrm{d}s = 0. \tag{3.79}$$

Now

$$\begin{aligned} (f, (\Lambda_{D_e} - \Lambda_{D_i}) f) &= \int_{\partial H} u_0 \left(\frac{\partial u_e}{\partial \nu} - \frac{\partial u_i}{\partial \nu} \right) \mathrm{d}s \\ &= -\int_{\partial D_e} u_i \frac{\partial u_e}{\partial \nu} \mathrm{d}s \\ &> 0. \end{aligned} \tag{3.80}$$

□

4. NUMERICAL IMPLEMENTATION

The forward solution as set out in the previous chapter, as well as in the inverse algorithms developed later, require only boundary-to-boundary integral operators. The numerical implementation therefore uses the Boundary Element Method (BEM), a powerful approach to solving partial differential equations numerically[137], particularly for homogeneous media. In contrast to the Finite Element Method (FEM), which uses a discrete mesh over the whole space of interest, the BEM discretises and solves only on the boundary. This requires less points and the BEM solves a much smaller, albeit dense, system of linear equations rather than the larger, sparse, FEM matrix, and can be more efficient. For problems in potential theory, the Fast Multipole Method[142, 143] can provide a frame work for very effecient numerics and in future it may be necessary to use this in real time applications of the algorithms developed in this thesis.

The implementation here will be using collocation quadrature, that is all functions are assumed to be piecewise constant. This chapter provides expressions for the different operator matrices, comments on the implementation and discusses some aspects of accuracy.

Note that whereas in the rest of this thesis x, y, z are used to denote separate 3D points. In this section, (x, y, z) is used to refer to individual cartesian coordinates and 3D vectors are given in bold $\boldsymbol{x}, \boldsymbol{r}$ etc. Bold letters will also be used to denote the discrete approximation of operators and subscript indices to denote the individual matrix elements. For instance the matrix $\boldsymbol{S}$, the element S_{ij} and the operator S. A small deviation from this rule will be made for the matrix $\tilde{\boldsymbol{K}}$ and elements $\tilde{K}_{ij}$ which represent the operator sum $\frac{1}{2}I + \tilde{K}$.

4.1 Triangulated 3D Objects

The 3D object ∂D is represented by a triangle mesh. Each triangle ∂D_j is characterized by 3 vertices $\boldsymbol{r}_1, \boldsymbol{r}_2$ and $\boldsymbol{r}_3$. The centre point $\boldsymbol{x}_j$ of the triangle is given by

$$\boldsymbol{x}_i = \frac{1}{3}(\boldsymbol{r}_1 + \boldsymbol{r}_2 + \boldsymbol{r}_3). \tag{4.1}$$

The area of a triangle with sides a, b and c is given by

$$A_i = \int_{\partial D_i} \mathrm{d}s = \frac{1}{4}\sqrt{2a^2b^2 + 2a^2c^2 + 2b^2c^2 - a^4 - b^4 - c^4}. \tag{4.2}$$

The potential and its first derivative on the object are represented by vectors $\boldsymbol{v}$ ($v_i = A_i\phi(\boldsymbol{x}_i)$) and $\boldsymbol{q}$ ($q_i = A_i g_D(\boldsymbol{x}_i)$). The advantage of scaling the functions by the area of each triangle is that the matrix $\boldsymbol{S}$ is symmetric under certain quadrature rules and the matrices $\boldsymbol{T}$ and $\boldsymbol{T}^*$ are adjoint to each other.

On the plane ∂H are located a number of electrodes with area R_i, the potential on the plane and the *change* in the first derivative (charge image) are given by vectors $\boldsymbol{w}$ ($w_i = f_H(\boldsymbol{x}_i)R_i$) and $\boldsymbol{c}$ ($c_i = c(\boldsymbol{x}_i)$). We note that the charge image is not scaled by the area of the electrodes. It is also worthwhile to point out that the electrodes R_i do not necessarily cover the plane ∂H. The Dirichlet data on the plane is, in effect, compactly supported.

4.1.1 Matrix S

The single layer potential operator S defined in (3.22) is discretised as follows.

$$\begin{aligned}
(Sg_D)(\boldsymbol{x}_i) &= \int_{\partial D} G(\boldsymbol{x}_i, \boldsymbol{y}) g_D(\boldsymbol{y})\, \mathrm{d}s(\boldsymbol{y}) \\
&= \sum_j \int_{\partial D_j} G(\boldsymbol{x}_i, \boldsymbol{y}) g_D(\boldsymbol{y})\, \mathrm{d}s(\boldsymbol{y}) \\
&\approx \sum_j g_D(\boldsymbol{y}_j) \int_{\partial D_j} G(\boldsymbol{x}_i, \boldsymbol{y})\, \mathrm{d}s(\boldsymbol{y}).
\end{aligned} \tag{4.3}$$

The integral of $G(\boldsymbol{x}_i, \cdot)$ over each element ∂D_j is performed differently depending on whether the calculation concerns a diagonal element ($i = j$), whether triangles i and

j are near to eachother, or whether they are relatively remote. For diagonal elements the Cauchy principal value is calculated over multiple circle segments of small angle ϕ_k centred at $\boldsymbol{x}_i$, with radius r_k such that the arc of the segments intersects the side of the triangle and the area of the arc segment equals the area of the triangle segment. That is

$$\begin{aligned}\int_{\partial D_i} G(\boldsymbol{x}_i, \boldsymbol{y})\,\mathrm{d}s(\boldsymbol{y}) &\approx \frac{1}{4\pi}\left[\sum_k \int_0^{r_k}\int_0^{\phi_k} \frac{\rho\,\mathrm{d}\phi\,\mathrm{d}\rho}{\rho} - \frac{1}{\sqrt{4z_j^2}}\int_{\partial D_i}\mathrm{d}s(\boldsymbol{y})\right] \\ &= \frac{1}{4\pi}\sum_k r_k\phi_k - \frac{A_j}{8\pi z_j}. \end{aligned} \tag{4.4}$$

For triangles i and j that are near to each other, the integral over ∂D_j is calculated using a piecewise constant approximation of $G(\boldsymbol{x}_i, \cdot)$ over a dense (typically 1,000) subdivision of ∂D_j. The density of the subdivision decreases as the triangles i and j are further apart. For remote pairs, a single piecewise constant approximation is taken over the entire area of ∂D_j. That is

$$\int_{\partial D_j} G(\boldsymbol{x}_i, \boldsymbol{y})\,\mathrm{d}s(\boldsymbol{y}) \approx G(\boldsymbol{x}_i, \boldsymbol{y}_j)\int_{\partial D_j}\mathrm{d}s(\boldsymbol{y}) = G(\boldsymbol{x}_i, \boldsymbol{y}_j)A_j.$$

In axi-symmetric configurations, used for some of the results in chapter 6, the integral of $G(\boldsymbol{x}_i, \cdot)$ over each element ∂D_j is performed numerically with NAG[144] routine `D01AJF`, a general purpose adaptive routine.

With the definition of the vector $\boldsymbol{q}$ ($q_i = -A_i g_D(\boldsymbol{x}_i)$), the matrix $\boldsymbol{S}$ is defined by the quadrature above. Note that because of the definition of q_i, the columns of $\boldsymbol{S}$ are normalized by A_j.

In the interest of execution speed, it is on occasion convenient to make the assumption of a constant $G(\boldsymbol{x}_i, \cdot)$ over the entire element ∂D_j for every $j \neq i$. Moreover, an approximate Cauchy principal value is calculated for $j = i$, for just one single segment k in (4.4). This means taking a disk at $\boldsymbol{x}_i$, with an area equal to the area of the triangle. That is, a disk with radius $a_i^2 = A_i/\pi$. In that case we obtain for the diagonal

element;

$$\begin{aligned}\int_{\partial D_i} G(\boldsymbol{x}_i, \boldsymbol{y})\,\mathrm{d}s(\boldsymbol{y}) &\approx \frac{1}{4\pi}\Big[\int_0^{a_j}\int_0^{2\pi} \frac{\rho\,\mathrm{d}\phi\,\mathrm{d}\rho}{\rho} - \frac{A_j}{\sqrt{4z_j^2}}\Big] \\ &= \frac{1}{2}\sqrt{\frac{A_j}{\pi}} - \frac{A_j}{8\pi z_j}. \end{aligned} \tag{4.5}$$

In that case we have for the matrix elements

$$\begin{aligned}\hat{S}_{i,j} &= \frac{1}{2}\sqrt{\frac{1}{A_i\pi}} - \frac{1}{8\pi z_i}, && \text{for } i = j, \\ &= \frac{1}{4\pi\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}} \\ &\quad -\frac{1}{4\pi\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i + z_j)^2}}, && \text{for } i \neq j. \end{aligned} \tag{4.6}$$

This efficiently calculated but relatively poor quadrature matrix is clearly symmetric. It is estimated that this expression gives a typical error of 0.3%, compared with a typical error of 0.06% for the more accurate quadrature, please see the discussion around figure 4.1B below.

4.1.2 Matrices $\mathbf{T}$ *and* $\mathbf{T}^*$

The operators T defined in (3.23) and (3.24) are discretised as follows,

$$\begin{aligned}(Tg_D)(\boldsymbol{x}_i) : &= \int_{\partial D} \frac{\partial G(\boldsymbol{x}_i, \boldsymbol{y})}{\partial z_i} g_D(\boldsymbol{y})\,\mathrm{d}s(\boldsymbol{y}) \\ &\approx \sum_j g_D(\boldsymbol{y}_j) \int_{\partial D_j} \frac{z}{2\pi((x - x_i)^2 + (y - y_i)^2 + z^2)^{3/2}}\,\mathrm{d}s(\boldsymbol{y}) \\ &\approx \sum_j g_D(\boldsymbol{y}_j) A_j \frac{z_j}{2\pi((x_j - x_i)^2 + (y_j - y_i)^2 + z_j^2)^{3/2}}. \end{aligned}$$

$$
\begin{aligned}
(T^* f_H)(\boldsymbol{x}_i) : &= \int_{\partial H} \frac{\partial G(\boldsymbol{x}_i, \boldsymbol{y})}{\partial z} f_H(\boldsymbol{y}) \, \mathrm{d}s(\boldsymbol{y}) \\
&\approx \sum_j f_H(\boldsymbol{x}_j) \int_{\partial H_j} \frac{z_i}{2\pi((x_i - x)^2 + (y_i - y)^2 + z_i^2)^{3/2}} \, \mathrm{d}s(\boldsymbol{y}) \\
&\approx \sum_j f_H(\boldsymbol{x}_j) R_j \frac{z_i}{2\pi((x_i - x_j)^2 + (y_i - y_j)^2 + z_i^2)^{3/2}},
\end{aligned}
$$

and so the adjoint matrices are

$$
T_{i,j} = T^*_{j,i} = \frac{z_j}{2\pi((x_j - x_i)^2 + (y_j - y_i)^2 + z_j^2)^{3/2}}.
$$

4.1.3 The Matrices $\tilde{\mathbf{K}}$ and $\tilde{\mathbf{M}}$

The modified double layer potential operator $\tilde{K}$ defined in (3.26) is discretised as follows,

$$
\begin{aligned}
(\tilde{K}\phi)(\boldsymbol{x}_i) &= \int_{\partial D} \left[\frac{\partial G(\boldsymbol{x}_i, \boldsymbol{y})}{\partial \nu(\boldsymbol{y})} + \frac{1}{|\boldsymbol{x}_i - \boldsymbol{x}_c|} - \frac{1}{|\boldsymbol{x}_i - \boldsymbol{x}'_c|} \right] \phi(\boldsymbol{y}) \, \mathrm{d}s(\boldsymbol{y}) \quad \boldsymbol{x} \in \partial D \\
&\approx \sum_{j \neq i} \left[\frac{\partial G(\boldsymbol{x}_i, \boldsymbol{y}_j)}{\partial \nu(\boldsymbol{y}_j)} + \frac{1}{|\boldsymbol{x}_i - \boldsymbol{x}_c|} - \frac{1}{|\boldsymbol{x}_i - \boldsymbol{x}'_c|} \right] A_j \phi(\boldsymbol{y}_j) \\
&\quad + \left[\frac{1}{A_i} \int_{\partial D_i} \frac{\partial G(\boldsymbol{x}_i, \boldsymbol{y})}{\partial \nu(\boldsymbol{y})} \, \mathrm{d}s(\boldsymbol{y}) + \frac{1}{|\boldsymbol{x}_i - \boldsymbol{x}_c|} - \frac{1}{|\boldsymbol{x}_i - \boldsymbol{x}'_c|} \right] A_i \phi(\boldsymbol{y}_i). \quad (4.7)
\end{aligned}
$$

Here $\boldsymbol{x}_c \in D$ and $\boldsymbol{x}'_c$ is its mirror point. In the calculations $\boldsymbol{x}_c$ is chosen at the centre of the object

$$
\boldsymbol{x}_c = \frac{\sum_i A_i \boldsymbol{x}_i}{\sum_i A_i}. \quad (4.8)
$$

For $i \neq j$, the terms $\partial G(\boldsymbol{x}_i, \boldsymbol{y}_j)/\partial \nu(\boldsymbol{y}_j)$ are calculated numerically.

$$
\frac{\partial G(\boldsymbol{x}_i, \boldsymbol{y}_j)}{\partial \nu(\boldsymbol{y}_j)} = \frac{1}{\delta} \left[G(\boldsymbol{x}_i, \boldsymbol{y}_j + \delta \nu(\boldsymbol{y}_j)) - G(\boldsymbol{x}_i, \boldsymbol{y}_j) \right]. \quad (4.9)
$$

In which δ is a suitably chosen small number. To calculate the singular term in the diagonal elements we use corollary C.6 on page 174, which gives

$$
\int_{\partial D_i} \frac{\partial G(\boldsymbol{x}_i, \boldsymbol{y})}{\partial \nu(\boldsymbol{y})} \, \mathrm{d}s(\boldsymbol{y}) = -\frac{1}{2} - \sum_{j \neq i} A_j \frac{\partial G(\boldsymbol{x}_i, \boldsymbol{y}_j)}{\partial \nu(\boldsymbol{y}_j)}. \quad (4.10)
$$

In which the terms of the sum are worked out as in (4.9). With the vector $A_i\phi(\boldsymbol{x}_i)$, the discrete approximation of $\frac{1}{2}I + \tilde{K}$, the matrix $\tilde{\boldsymbol{K}}$ is given by

$$\begin{aligned}
\tilde{K}_{i,j} &= -\frac{1}{A_i}\sum_{j\neq i} A_j \frac{\partial G(\boldsymbol{x}_i, \boldsymbol{y}_j)}{\partial \nu(\boldsymbol{y}_j)} \\
&\quad + \quad \frac{1}{|\boldsymbol{x}_i - \boldsymbol{x}_c|} - \frac{1}{|\boldsymbol{x}_i - \boldsymbol{x}_c'|} \qquad \text{for } i = j \\
&= \frac{\partial G(\boldsymbol{x}_i, \boldsymbol{y}_j)}{\partial \nu(\boldsymbol{y}_j)} + \frac{1}{|\boldsymbol{x}_i - \boldsymbol{x}_c|} - \frac{1}{|\boldsymbol{x}_i - \boldsymbol{x}_c'|} \qquad \text{for } i \neq j.
\end{aligned}$$

Note that there is no $\frac{1}{2}$ term in the diagonal terms because the $-\frac{1}{2}$ in (4.10) is compensated by the $\frac{1}{2}I$ term of the operator combination $\frac{1}{2}I + \tilde{K}$.

The operator $\tilde{M}$ defined in (3.28) is calculated from

$$\begin{aligned}
(\tilde{M}\phi)(\boldsymbol{x}_i) &= \int_{\partial D} \left[\frac{\partial^2 G(\boldsymbol{x}_i, \boldsymbol{y})}{\partial z_i \partial \nu(\boldsymbol{y})} + \frac{2z_c}{|\boldsymbol{x}_i - \boldsymbol{x}_c|^3}\right] \phi(\boldsymbol{y})\, \mathrm{d}s(\boldsymbol{y}) \qquad \boldsymbol{x}_i \in \partial H \\
&= \sum_j \int_{D_j} \left[\frac{\partial^2 G(\boldsymbol{x}_i, \boldsymbol{y})}{\partial z_i \partial \nu(\boldsymbol{y})} + \frac{2z_c}{|\boldsymbol{x}_i - \boldsymbol{x}_c|^3}\right] \phi(\boldsymbol{y})\, \mathrm{d}s(\boldsymbol{y}) \\
&\approx \sum_j \left[\frac{\partial^2 G(\boldsymbol{x}_i, \boldsymbol{y}_j)}{\partial z_i \partial \nu(\boldsymbol{y}_j)} + \frac{2z_c}{|\boldsymbol{x}_i - \boldsymbol{x}_c|^3}\right] A_j \phi(\boldsymbol{y}_j).
\end{aligned} \tag{4.11}$$

Where again the normal derivatives are calculated numerically

$$\frac{\partial^2 G(\boldsymbol{x}_i, \boldsymbol{y}_j)}{\partial z_i \partial \nu(\boldsymbol{y}_j)} \approx \frac{1}{\delta}\left[\frac{\partial G(\boldsymbol{x}_i, \boldsymbol{y}_j + \delta\nu(\boldsymbol{y}_j))}{\partial z_i} - \frac{\partial G(\boldsymbol{x}_i, \boldsymbol{y}_j)}{\partial z_i}\right].$$

The expression for $\partial G(\boldsymbol{x}_i, \cdot)/\partial z_i$ has already been given in the discretization of T above.

4.2 Quadrature Tests

To verify the calculation of matrix elements and the validity of the piecewise constant quadrature, a number of tests have been been carried out.

For a uniformly charged sphere centred at $\boldsymbol{x}_c$ of radius a_0, the single layer potential and the normal derivative on the plane can be calculated exactly (cf lemma C.2). This

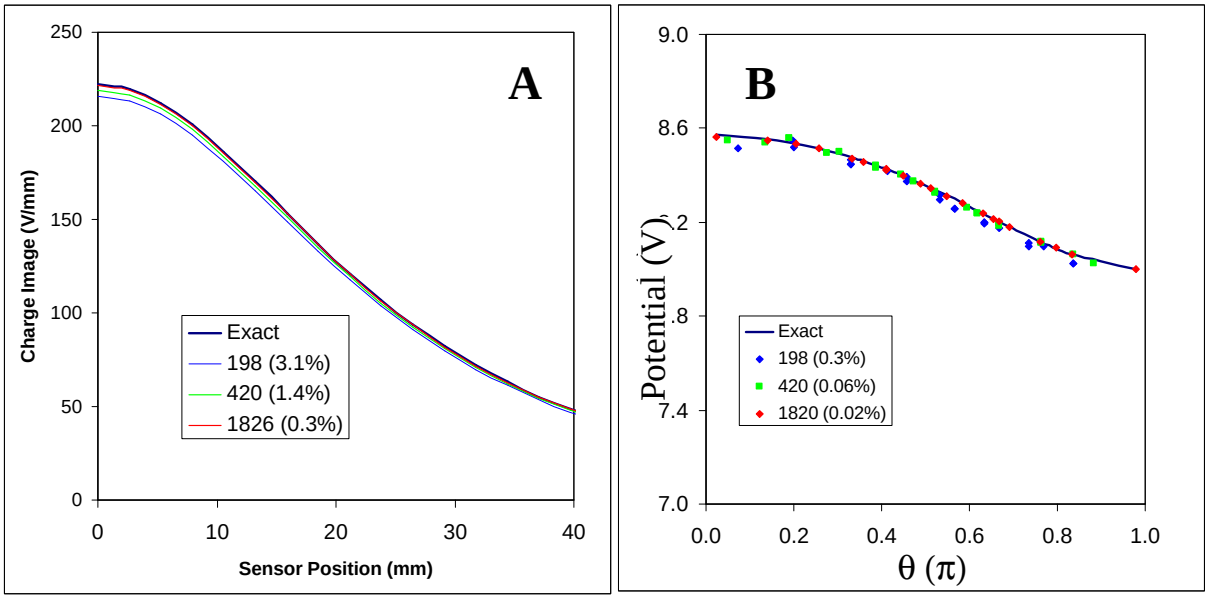

Fig. 4.1: Comparison of exact and numerical quadrature. A) Charge images calculated from (4.12) (labelled 'Exact') and (4.13) for different triangulation densities. B) Comparison of (4.14) and (4.15). Percentages in brackets indicate error with respect to the exact result.

gives

$$
\begin{aligned}
(T\ 1)(\boldsymbol{x}_i) &= \int_{\partial H} \frac{\partial G(\boldsymbol{x}_i, \boldsymbol{y})}{\partial z}\, \mathrm{d}s(\boldsymbol{y}) \quad \boldsymbol{x}_i \in \partial H \\
&= \frac{2a_0^2 z_c}{\sqrt{|\boldsymbol{x}_i|^2 + z_c^2}} \qquad (4.12) \\
&\approx \sum_j T_{i,j} A_j. \qquad (4.13)
\end{aligned}
$$

Figure 4.1A compares (4.12) and (4.13) for a sphere located at $\boldsymbol{x}_c = (0, 0, 30)$ and $a_0 = 10$. An object of this size and location is quite typical in the numerical experiments discussed later in this thesis. The receiver sensors, that is the points $\boldsymbol{x}_i$ are located on a radial line from the centre of the half plane ∂H and the abscissa in figure 4.1A represents this radial position. The thicker top blue line illustrates the exact result according to (4.12), while thinner coloured lines beneath that illustrate the approximate results for a sphere consisting of 198 triangles, 420 triangles and 1826 triangles respectively. The exact and approximate results appear quite close together in the graph of figure 4.1A,

but this is primarily due to the dynamic range on the ordinate. The error in percentage terms in row sums of $\boldsymbol{T}$ is given in brackets and decreases from 3.1% to 0.3%. This error is quite large, but is almost entirely explained by the difference in area between the triangulated object and the perfect sphere which it approximates. For instance the area of the 420 triangle polyheadron is 1.5% smaller than the $4\pi \cdot 10^2$ mm^2 of the sphere.

The quadrature for $\boldsymbol{S}$ can be tested in a similar way. Taking again the single layer potential of a uniformly charged sphere at $\boldsymbol{x}_c = (0, 0, z_c)$ we can also calculate

$$\begin{aligned}(S\ 1)(\boldsymbol{x}_i) &= \int_{\partial D} G(\boldsymbol{x}_i, \boldsymbol{y})\, \mathrm{d}s(\boldsymbol{y}) \quad \boldsymbol{x} \in \partial D \\ &= a_0 - \frac{a_0^2}{|\boldsymbol{x}_i - \boldsymbol{x}_c'|} \qquad (4.14) \\ &\approx \sum_j S_{i,j} A_j. \qquad (4.15)\end{aligned}$$

Figure 4.1B compares (4.14) and (4.15) for the same sphere located at $\boldsymbol{x}_c = (0, 0, 30)$ and $a_0 = 10$mm. The full line illustrates the exact result according to (4.14), while the diamonds, squares and triangles illustrate the approximate results from (4.15). As can be seen the approximate results are very close to the exact results with much smaller percentage errors than those observed in figure 4.1A. The reason for this is that the single layer potential close to the object depends primarily on the radius a_0 as can be seen in (4.14), whereas the single layer potential further away from the object depends on area as can be seen from the a_0^2 term in the numerator of (4.12). For the 420 triangle polyheadron, the average radius is 0.1% smaller than the 10mm of the ideal object. This tallies fairly well with the error reported in figure 4.1B.

The quadrature for $\tilde{K}_{i,j}$ and $\tilde{M}_{i,j}$ is tested with the divergence theorem for the electric field, or Gauss's law. For a point source at a point $\boldsymbol{z}$ we have

$$\Theta(\boldsymbol{z}) = -\int_{\partial D} \frac{\partial}{\partial \nu(\boldsymbol{y})} \frac{1}{|\boldsymbol{y} - \boldsymbol{z}|}\, \mathrm{d}s(\boldsymbol{y}) = \int_D \delta(\boldsymbol{x} - \boldsymbol{z})\, \mathrm{d}\boldsymbol{x}. \qquad (4.16)$$

It is clear that $\Theta(z) = 1$ for $z \in D$ and $\Theta(z) = 0$ for $z \notin D$. In the program the function $\Theta(z)$ is part of the standard output and the surface integral in (4.16) is calculated using the same piecewise constant quadrature and numerical differentiation (4.9) for $\tilde{K}$ and

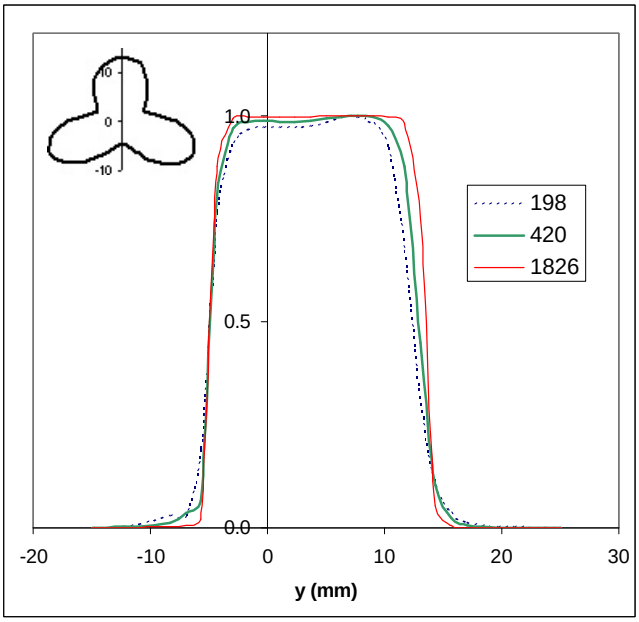

Fig. 4.2: Calculation of total electric field flux (4.16) as function of position of the point charge, for different discretization levels

$\tilde{M}$. Figure 4.2 plots $\Theta(z)$ on a line through a test object that is illustrated in the inset. As can be seen, $\Theta(z) \approx 1$ inside the object and falls off rapidly towards zero at the edge. The sharpness of the transistion depends on the triangulation density of the object.

4.3 Matrix Solvers

With the matrices and vectors defined above, the steps in the numerical calculation of the forward charge image can be formulated. First, the applied potential at each triangle of the object is calculated; $\boldsymbol{t} = \boldsymbol{T}^*\boldsymbol{w}$. Then the charge at each triangle is found by solving $\boldsymbol{S}\boldsymbol{q} = \boldsymbol{t}$ for $\boldsymbol{q}$. With that, the charge image is then calculated $\boldsymbol{c} = \boldsymbol{T}\boldsymbol{q}$. Alternatively it is possible to solve $\tilde{\boldsymbol{K}}\boldsymbol{v} = \boldsymbol{t}$ for the dipole moment at each triangle and then calculate the charge image from $\boldsymbol{c} = \tilde{\boldsymbol{M}}\boldsymbol{v}$.

In each of these cases, it is the second step that matters most. The numerical solution of Fredholm Equations of the first kind has been the subject of much research, see, for instance, [145, 146, 147] for historical overviews. In our case, the object D typically consists of 420 triangles and $\boldsymbol{S}$ and $\tilde{\boldsymbol{K}}$ are 420×420 matrices. We use the well tested and well documented NAG[144, 148] routines to solve equation $\boldsymbol{t} = \boldsymbol{S}\boldsymbol{q}$, either directly or through the calculation of an inverse. Both methods can also be used

to solve $\tilde{\boldsymbol{K}}\boldsymbol{v} = \boldsymbol{t}$, but in the numerical experiments presented in this work, the equation is always solved by calculating an inverse.

For direct solution of the matrix equation $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ with a square matrix $\boldsymbol{A}$, the program uses NAG routine `F04ATF`. The routine uses an LU factorization with partial pivoting, and iterative refinement until full machine accuracy is obtained. This method is efficient when we are interested in solving the equation just once, that is when obtaining the charge image for just one single transmitter, or a single transmitter drive pattern f_H. When calculating a discrete approximate of a forward map Λ_D for multiple transmitters, it is more convenient to use the inverse of the matrix $\boldsymbol{S}$, that is

$$\boldsymbol{\Lambda_D} = \boldsymbol{T}(\boldsymbol{S}^{-1}\boldsymbol{T}^{*}). \tag{4.17}$$

Two NAG routines are used for the calculation of the inverse, both of which return the inverse matrix rather than the LU factorisation. The latter can be more efficient, but using the inverse was more convenient from a coding perspective. `F01BLF` calculates the rank and pseudo-inverse of a general $m \times n$ real matrix, $m \geq n$, using QR factorization with column interchanges. The routine `F01BLF` is also used to calculate the inverse of $\tilde{\boldsymbol{K}}$. Because $\hat{\boldsymbol{S}}$ defined in (4.6) is symmetric, the inverse in that case can be calculated with `F01ABF`. This provides the inverse of a real $n \times n$ symmetric positive definite matrix with guaranteed accuracy. That is, with iterative refinement. The routine is about 7 times slower than `F01BLF` although there is not much discernable difference in the forward map, and no discernable difference in the reconstruction.

Although the methods used to solve the inverse problem are discussed in subsequent chapters of this thesis, it is appropriate to briefly mention some of the numerical methods used to solve the matrix equations that arise. Central to the iterative method in chapter 6 is the minimization of the Tikhonov functional

$$f = \min_{f} ||Af - b||^2 + \alpha^2 ||f - f_0||_P^2. \tag{4.18}$$

In which f_0 is the prior and P the penalty norm. For standard Tikhonov regularisation $P = I$. In much of this work a discrete approximation of a first order Sobolev norm is

used, that is

$$\begin{aligned} ||f||_P^2 &= \int_{\partial D} f^2 \,\mathrm{d}s + \int_{\partial D} \nabla f \cdot \nabla f \,\mathrm{d}s \\ &= \int_{\partial D} fPf \,\mathrm{d}s(\boldsymbol{x}) \\ &\approx \sum_i A_i f(\boldsymbol{x}_i) \sum_j P_{i,j} f(\boldsymbol{x}_j). \end{aligned}$$

The matrix $\boldsymbol{P}$ is given by $\boldsymbol{P} = \boldsymbol{I} + \boldsymbol{L}^{\boldsymbol{T}}\boldsymbol{L}$, in which

$$\begin{aligned} L_{i,i} &= \frac{1}{d_1} + \frac{1}{d_2} + \frac{1}{d_3} \\ L_{i,i_1} &= -\frac{1}{d_1} \\ L_{i,i_2} &= -\frac{1}{d_2} \\ L_{i,i_3} &= -\frac{1}{d_3}. \end{aligned} \tag{4.19}$$

Here i_1, i_2 and i_3 are the triangle indices of the three nearest neighbours of triangle i and d_1, d_2 and d_3 the distances between triangle i and its three nearest neighbours.

The Tikhonov functional is minimized by the solution to the equation

$$(A^T A + P)x = A^T b + Px_0. \tag{4.20}$$

The Tikhonov problem can be solved efficiently through a bidiagonalisation scheme or an iterated scheme[149], and the explicit use of (4.20) is not generally employed in large systems. In our case however it was convenient to trade off the high reliability of NAG against the inefficiency of using (4.20). Because $(A^T A + P)$ is symmetric positive-definite, the equation is solved with NAG routine `F04ASF` which calculates the accurate solution of a set of real symmetric positive-definite linear equations, using a Cholesky factorization and iterative refinement. When, either because P is defined differently or because of numerical difficulties, the system is non-symmetric, an automatic switch is made to the routine `F04ATF` mentioned above, and the change logged.

The range tests used in chapter 7, as well as the monotonicity method in chapter 8 require the calculation of a Singular Value Decomposition (SVD). This is calculated

with `F02WEF`, which obtains the SVD through a sequence of transformations and Givens plane rotations.

4.4 Forward Maps

The solution of the forward problem, and the setup of the boundary element method discussed above provides us with three different methods of numerically calculating the discrete approximation to the Dirichlet to Neumann map. The first method, called for convenience the Physicist's method uses the $\Lambda_D = T\hat{S}^{-1}T^*$ factorisation in (3.25) to calculate the matrix equation (4.17). The second method takes the perspective more of the experimentalist, who switches on just one transmitter electrode R_i and measures the resulting charge image c which then forms the i-th column of the forward matrix. This is then repeated for the next transmitter and so on. In matrix terms we can write,

$$\boldsymbol{\Lambda_D} = \boldsymbol{T}\ (\boldsymbol{S} \setminus \boldsymbol{T^*}).$$

In which we use the MatLab backslash notation to denote that we solve the equation $\boldsymbol{Sq} = \boldsymbol{t}$ for each column $\boldsymbol{t}$ of $\boldsymbol{T^*}$. The experimentalist method is generally preferred when there are not the same number of transmitters and receivers and they are not colocated. The definitions of $\boldsymbol{T}$ and $\boldsymbol{T^*}$ must then be adjusted adjusted appropriately and are no longer adjoint. Finally the last method takes the mathematically more attractive approach of solving a Fredholm equation of the second kind to find the induced dipole distribution on the object and calculates

$$\boldsymbol{\Lambda_D} = \tilde{\boldsymbol{M}}\ (\tilde{\boldsymbol{K}}^{-1}\ \boldsymbol{T^*}). \tag{4.21}$$

It is important to verify that these three methods give the same numerical answer for the forward Dirichlet to Neumann map. The answer is that they do, with a typical variation between them of 0.1%. The discussion below explores this in more detail and does so specifically by looking at the variation in elements of the singular value decomposition. The singular value decomposition is useful here because we want to investigate an entire forward map. The variation in a single forward charge image has already been explored superficially in the discussion around figure 4.1A, where a dependence on the discretization of the test object was found. Here we use the

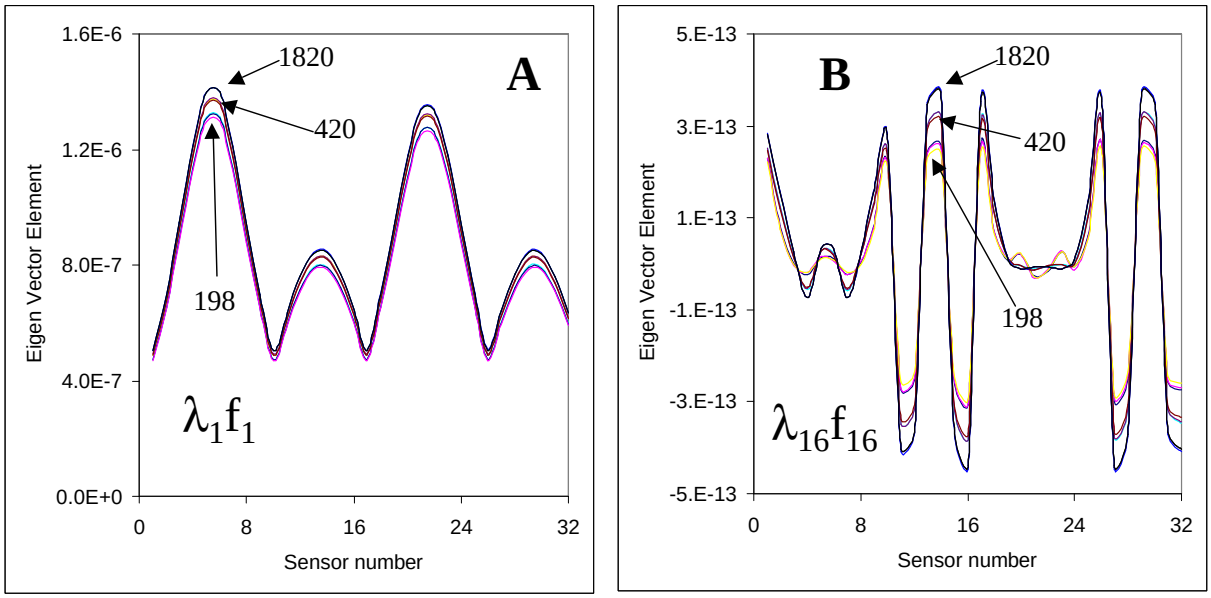

Fig. 4.3: Eigenvectors f_1 and f_{16} of the forward matrix scaled by eigenvalue, for different methods and for different discretization levels (1820, 420 and 198 triangles) of the object. Sensor electrodes, forward map and test object are illustrated in figure 4.6.

example of 32 transmitters and collocated receivers arranged in a linear array at the edge of display, and a test object located 30mm above the plane of the sensors. For an illustration of the electrode layout, the test object used and a visual representation of the forward map, see figure 4.6 below.

A straightforward comparison between the forward matrices resulting from the different methods, can be made by looking at the matrix norm, or first eigenvalue. For 420 triangles $\bar{\lambda}_1 = 5.013 \cdot 10^{-6}\text{mm}^{-1}$, with a standard deviation of 0.1% between the methods. More variation between the forward matrices can be detected when looking at the higher order eigenvalues, for instance $\bar{\lambda}_{16} = 1.29 \cdot 10^{-12}\text{mm}^{-1}$, with a standard deviation of 2% between the methods. These trends are illustrated in figure 4.3 which illustrates two of the eigenvectors f_1 and f_{16} of the forward matrix scaled by eigenvalue.

The figure also illustrates the effect of different discretization levels of the object. The variation between discretization level is an order of magnitude larger than that between the methods. That is 3% for λ_1 and 19% for λ_{16}. It must be pointed out that the variation with triangulation density must not be interpreted as an estimate of error in the numerical implementation. Objects at different triangle densities simply are different objects. Another trend worth noting is that at higher triangulation densities

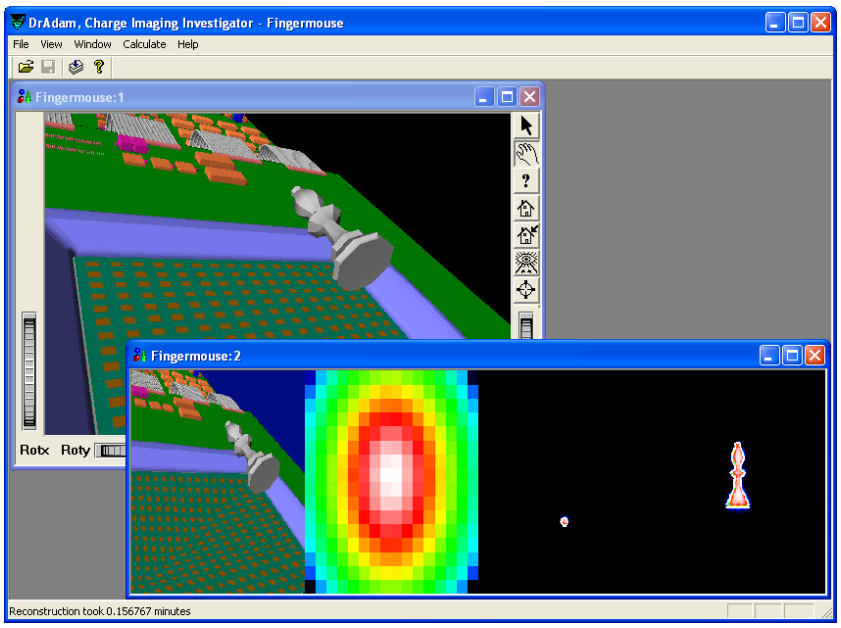

Fig. 4.4: Screen shot of the Program DrAdam, left shows the Open Inventor Examiner window, bottom right shows the standard output strip for forward problem solution of the bishop chess piece object and a 2D array of receivers arranged over the area of the display. A single transmiter at the centre of the display area is used. The second image in the output strip shows the charge image. The third and fourth image are object cross sections with a plot of the flux function $\Theta(z)$.

the variation in λ_{16} between the methods decreases.

4.5 Computer Program

A Windows program, called DrAdam, was written to perform the BEM calculations. It is written in c++ and uses the Microsoft Foundation Library (MFC) for initialisation, windows and messaging. To load and manipulate the 3D objects the program uses the Open Inventor[150] 3D graphics library. Figure 4.4 provides a screen shot of the program. The program is available for installation, provided its use is non-commercial. For details on this, please see appendix A.

Figure 4.4 gives a screen shot. The program bases the physical configuration of display, sensing electrodes and objects on the sensor prototype fingermouse system described in section 1.6. A model of the system is shown in an Open Inventor Examiner window and a test object is positioned above the display / sensor combination. The number of electrodes can be chosen at will and different electrode layouts can be used.

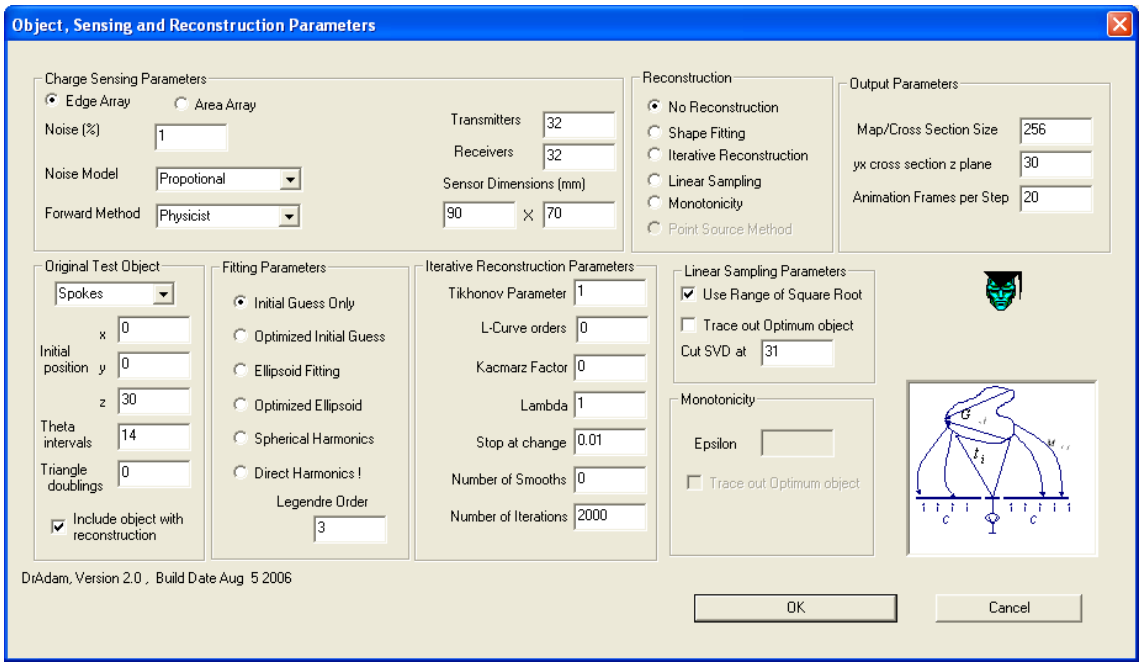

Fig. 4.5: Adam's File—OnNew Dialog box. This dialog box sets all options for forward and inverse problem calculations. The calculation itself is then triggered from the main frame menu item `Calculate`. The options are described in detail in appendix A.

In particular it possible to arrange the electrodes only around the edge of the display, or to distribute them in a 2D array over the surface of the display itself. Different test models are available and it also is possible to load test objects in VRML format from file. Figure 4.4 illustrates a situation where a bishop chess piece has been loaded as test object and a 2D array configuration of the sensor electrodes has been chosen.

A dialog box to set these and other options, as shown in 4.5, is brought up at start-up, details of the options are described in appendix A.

The calculation of the forward charge image, or Dirichlet to Neumann map is triggered from the menu (`Calculate`). The bottom righthand child window in figure 4.4 illustrates the output that is then created. This is the result strip of four images shown in the window. A bitmap image of the strip is automatically saved to disk. The first image is a perspective view of the test object above the sensor configuration. If a reconstruction algorithm has been selected in the dialog box (figure 4.5), then this image shows the reconstructed object in red, either alone or in combination with the original object. The second image in the result strip shows the charge image. The example in figure 4.4 shows the case for a single transmitter located at the centre of

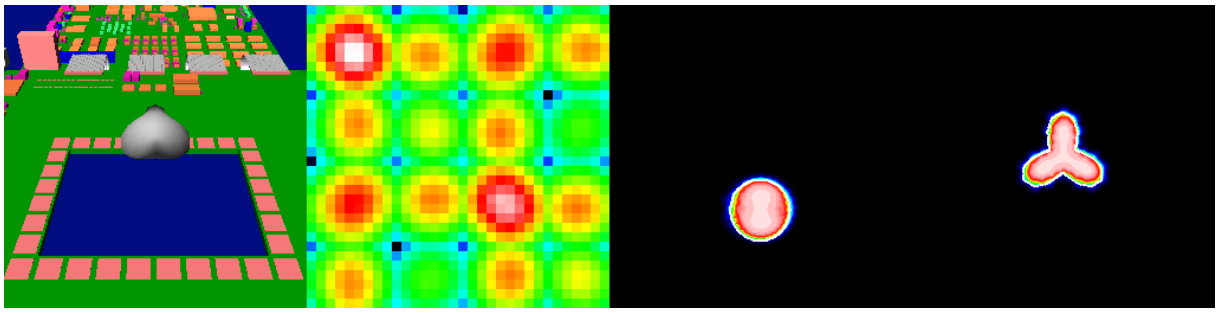

Fig. 4.6: Illustration of the standard output strip of DrAdam for the case of the forward problem solution of the spoke object located 30mm above the plane of the sensors and a linear array 32 electrodes around the edge of the display area. Each electrode in turn acts as transmitter and 32 charge images are collected. The combined forward map is illustrated in the second image in which each column is the charge image with one active transmitter.

the display area combined with a 2D array of 16 receiver electrodes distributed over the 90mm $\times$ 70mm area. The image is a scaled logarithmic representation of the charge image. White represents maximum and black the minimum (typically zero).

Random, normally distributed noise is added to the charge image in an algorithm due to Kindermann *et al*[151]. Two approaches have been employed to calculate the standard deviation of the noise for each transmitter-receiver pair. In the 'Uniform' noise model, the same standard deviation is used for all measurements. The standard deviation can be expressed as a percentage of the maximum signal in a given set of measurements or as proportion of the matrix norm of the forward matrix $||\mathbf{\Lambda_D}||$. This approach to noise essentially assumes that the electronics in the receiver electrodes operates identically regardless of the proximity to the transmitting electrode or the change in signal resulting from the object. Alternatively, in the 'Proportional Model', noise is added to the signal of each transmitter and receiver as a percentage proportion of that signal. The assumption underlying this mode is that the measurement system is designed such that the gain is adjusted for each measurement in response to the signal strength. Noise of this type preserves the dynamic range of the overall measurement set and in general much larger percentage noise levels of this type can be tolerated to achieve comparable reconstruction results.

The third and fourth images in the result strip show cross sections in the $z_2 = 0$ and $z_3 = 30$ planes respectively of the object. The cross sections also show the total electric flux $\Theta(z)$ from a point source through the surface of the object as function

of the position of the point source as defined in (4.16). In the image, white indicates $\Theta(z) = 1$ and black $\Theta(z) = 0$. As can be seen the map follows the figure outline quite closely but there are some intermediate values near the surface of the object. These effects decrease with increasing triangulation density of the object. In this way, the image permits a qualitative assessment of the quadrature used in solving the forward problem. The third and fourth images show the reconstructed object if a reconstruction has taken place and different output may be plotted.

If the `Calculate` menu item is selected a second time, the program performs a series of charge image calculations (and reconstructions if necessary). The object is initially stationary, allowing the user to examine the difference between different noise draws on the charge image and the reconstruction. The object is then moved over the sensor area, allowing the examination of the effect of position with respect to the electrode positions. Typically 200 animation frames are created and these are written to disk. These can then be combined into avi or mp4 movies or any other format. Animations of some results presented in this thesis, as well as additional results, are available on disk.

A second example of the results strip from the program is shown in figure 4.6, This figure shows the spoke object. Whereas the bishop object shown in figure 4.4 was loaded from file in VRML format, this object is build inside the program itself using spherical harmonics. This has the advantage that the triangulation is well defined and the discretization level is easily controlled. This means that the spoke object has been used for many of the numerical experiments in this thesis. It has already been used in figure 4.2 and in section 4.4 above. The electrode layout used in figure 4.6 is that of a linear array of 32 electrode around the edge of the display area. This is illustrated in the first image of the output strip. Each electrode can be both transmitter and receiver and the forward Dirichilet to Neuman map Λ_D is a 32×32 matrix. The matrix is illustrated schematically in the second image of the strip, each column j represents the signals from all electrodes as receivers while electrode j acts as transmitter. The map is symmetric in principle although the added noise destroys this symmetry.

5. SHAPE FITTING

This chapter explores a straightforward approach to an inverse problem; given a (black box) forward model, adjust or fit a set of parameters that describe a shape, until the calculated charge image or charge images (forward map) of that shape resembles, as closely as possible, the original measured data.

The results in this chapter indicate that it is possible to do this and that we can trade-off the measurement accuracy against the amount of detail in the shape fitting. For instance, it is possible to find the approximate size and the position of the object, from measurements contaminated with as much as 40% noise. If the degrees of freedom are increased to include object orientation, the acceptable noise level goes down to 20% or 15% depending on the amount of prior knowledge that is used. More general object parameterisation with spherical harmonics provides satisfactory results only at 5% noise.

For reference, a superficial description of the modified Gauss-Newton optimization is given here. The Newton Method for finding the zero of a single variate scalar function $f(x)$ is illustrated in figure 5.1. The update at each iteration step is given by

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}. \tag{5.1}$$

We are interested in optimizing multiple parameters that describe the position, orientation and shape of an object and this means minimizing a cost function that is a multivariate sum of squares given by

$$F(\xi) = \sum_i \int_{\partial H} [(\Lambda_D - \Lambda_\xi) f_i]^2 \, \mathrm{d}s(x). \tag{5.2}$$

In which $\Lambda_\xi : L^2(\partial H) \to L^2(\partial H)$ is the forward map of the parameterized object. The n-tuple ξ represents the collection of different fitting parameters, f_i is a given

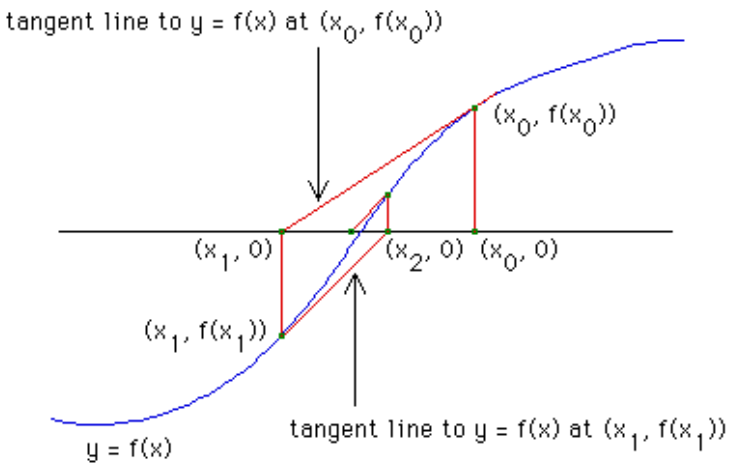

Fig. 5.1: Illustration of the univariate Newton method. This provides a useful mental picture for the Modified Gauss-Newton optimizations of object shape used in this chapter.

basis of $L^2(\partial H)$ determined by the pattern of transmitters used. It follows from the uniqueness of the Cauchy problem that the cost function is uniquely minimized for one f_i by one unique object. However as both the set of fitting parameters ξ and the number of sampling points of the integral, that is the number of receivers, is limited, this minimum may not be part of the parameter space or uniquely determined by the measurements.

For a multivariate sum of squares, the problem of finding the zero becomes that of finding a minimum and the derivative f' becomes the Jacobian matrix. The Gauss-Newton iteration is

$$\xi_{k+1} = \xi_k - J_k^{\dagger} F(\xi_k), \tag{5.3}$$

in which $J_k^{\dagger}$ is the Moore-Penrose generalised inverse of the Jacobian. It is more common to write this not as an inverse but to provide an equation for the decent direction $J_k^T J_k p_k = -J_k^T f_k$. It must also be noted that his equation arises because in the Modified Gauss-Newton method, $J_k^T J_k$ is taken as a good estimate for the Hessian (curvature) term. Hence our brief sketch in which $J^{\dagger}$ is presented a straightforward generalisation of $1/f'(x)$ is a slight of hand. We refer to Gill[148] for the appropriate detail.

Shape optimization is a branch of optimal control theory which is based on the calculus of variations. Under this topic the work goes back to the early 1900s when Hadamard coumputed the derivative of the Greens function of the Laplace operator with respect to the normal vector of the surface of a domain[152]. A substantial body of literature exist on the solution of inverse problems using iterative methods

of the kind used here. The shape Fréchet differentiability of the forward operator has been investigated in potential theory[153], elastics[154], EM scattering[112] and acoustics[155]. An efficient calculation of the Jacobian can be obtained from the domain derivative[153, 156], which is defined on any vector field a. Using this the Jacobian can be obtained using a range of different boundary conditions on a fixed object. This is a much more efficient way of calculating the Jacobian than calculating it for fixed boundary conditions on a range of objects. There is also some discussion on the merit of using the $J_k^T J_k$ term to estimate the second order derivatives. For instance, Haber et al[157] use a full Newton method with the second order derivatives to solve the inverse coefficient problem for geophysical EM. The relatively high cost of calculating the second order derivatives is alleviated by formulating the problem directly in the differential equation domain.

If the Jacobian in (5.3) is ill-conditioned, it may be necessary to use a regularised inverse for $J_k^\dagger$ [87]. Examples of this are the Levenburg-Marquardt, see for instance [158, 159] method or the iteratively regularized Gauss-Newton method[160], of which there is also a continuous analogue[161].

Regularisation in effect limits the number of active degrees of freedom by suppressing degrees that cause instability. This thesis takes a more structured approach by step wise introducing only those degrees of freedom that are really required. The next section explores fitting a sphere, the section after that the orientation of an ellipsoid and so on.

In the shape fitting tests below, comprehensive modified Gauss-Newton algorithms as implemented in NAG[144] routines `E04GDF` and `E04FCF` is used. For `E04GDF` the values of the Jacobian must be supplied at each iteration point, while `E04FCF` uses repeated calls to the multi-value, multi-variate $F(\xi)$ itself, to estimate the Jacobian internally.

5.1 The Initial Guess

The simplest optimization problem is that of finding the size and position of a spherical object that reproduces most closely the capacitance data of the original shape. With only four degrees of freedom, this is relatively straightforward and provides an initial guess for further optimization and for the iterative reconstruction algorithm in the next

chapter. Assuming that the transmitters and receivers are small in size and are well separated from each other, we also assume that the object is sufficiently far removed from the electrodes for the uniform charge assumption in section 3.1.4 to be valid. From (3.33) the charge image measured at a receiver located at x, when a transmitter located at y is active is given by

$$\hat{c}(x,y) = \frac{a_o z_3^2}{\pi r_t^3 r_r^3},$$

with

$$\begin{aligned} r_r^2 &= (x_1 - z_1)^2 + (x_2 - z_2)^2 + z_3^2, \\ r_t^2 &= (y_1 - z_1)^2 + (y_2 - z_2)^2 + z_3^2. \end{aligned}$$

As cost function is then used defined by

$$F(z) = \sum_{t,r} F_{t,r}^2 = \sum_{t,r} \left\{ \left[\frac{c}{c_0} \right]_{t,r} - \left[\frac{\hat{c}}{\hat{c}_0} \right]_{t,r} \right\}^2, \tag{5.4}$$

in which the sum is over all transmitter and receiver pairs (t,r). The term c_0 is the measured capacitive coupling between the electrodes in the absence of the object and $\hat{c}_0$ the calculated capacitive coupling as given by (3.31). This scaling is convenient from an experimental point of view because relative capacitance change is more readily available than absolute capacitance measurements. The scaling also helps the Modified Gauss-Newton routine work over a larger variety of different electrode numbers and configurations, without change in the control parameters. Note that the scaling means that diagonal elements of the cross capacitance matrix are ignored.

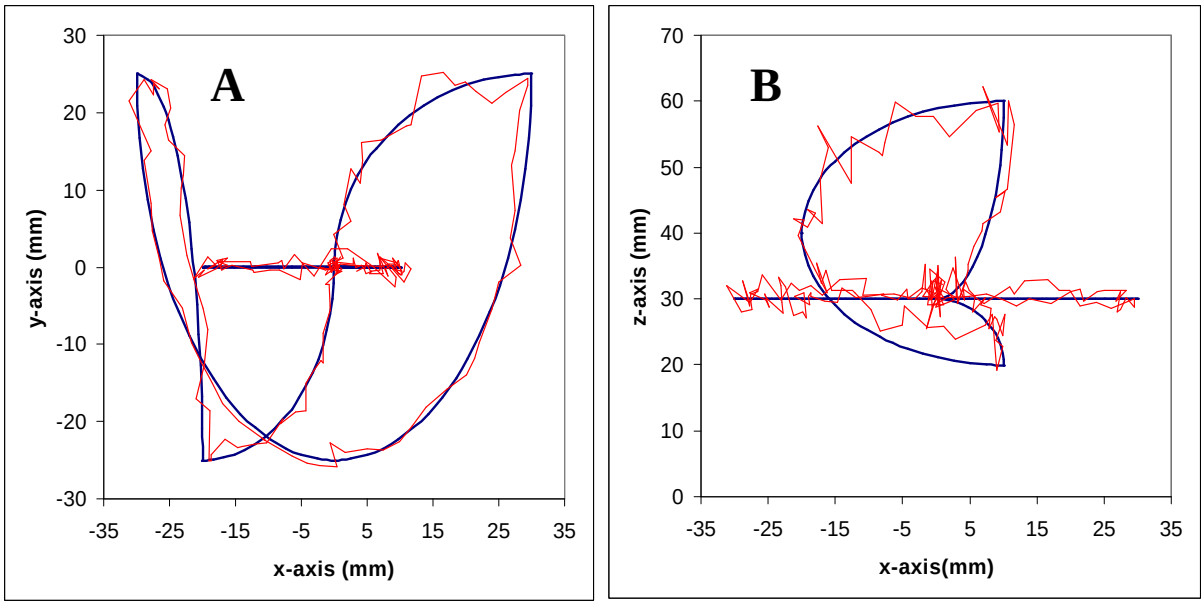

Fig. 5.2: Trace of an original object (smooth blue line) and tracked intial guess (jagged red) in the xy and xz plane. Random noise at 40% of the signal value was used.

The elements of the Jacobian are given by

$$\begin{aligned}
\frac{\partial F_{t,r}}{\partial z_1} &= 3\left[\frac{c}{c_0}\right]_{t,r}\left[\frac{y_1 - z_1}{r_t^2} + \frac{x_1 - z_1}{r_r^2}\right], \\
\frac{\partial F_{t,r}}{\partial z_2} &= 3\left[\frac{c}{c_0}\right]_{t,r}\left[\frac{y_2 - z_2}{r_t^2} + \frac{x_2 - z_2}{r_r^2}\right], \\
\frac{\partial F_{t,r}}{\partial z_3} &= 3\left[\frac{c}{c_0}\right]_{t,r}\left[\frac{2}{3z_3} - \frac{z_3}{r_t^2} - \frac{z_3}{r_r^2}\right], \\
\frac{\partial F_{t,r}}{\partial a_o} &= \left[\frac{c}{c_0}\right]_{t,r}\frac{1}{a_o}.
\end{aligned}$$

Because there is an explicit expression for the Jacobian, routine `E04GDF` [144] is used for the optimization.

Figure 5.2 illustrates how this strategy can track an object moving above the linear array of 32 sensors arranged around the edge of a 4" display. As discussed before, all electrodes act as transmitters in turn for the complete set of 32 receivers, collecting 32×32 data points. Normally distributed random noise using the proportional model, with a standard deviation of 40% of the signal strength for each signal is added to each signal. This represents a very high noise level, but can be tolerated because the

proportional noise model preserves dynamic range. If the alternative noise model is used in which uniform noise with the same standard deviation is added to all signals, a noise level corresponding approximately 15% of the maximum signal produces results comparable to those in figure 5.2. The average condition number of the Jacobian in figure 5.2 is $\bar{\kappa}(J_0) = 11$.

A refinement could be made on the initial guess by dropping the uniform charge assumption of 3.1.4 and calculate the full forward matrix for the cost function (5.2) to optimize the position and size of the sphere. This then uses `E04FCF` in which the Jacobian is estimated internally from multiple least squares evaluation calls. This requires substantial computing effort which is not justified by a commensurate improvement in the positional and size accuracy.

5.2 Ellipsoids

Following recovery of the object position with the initial guess, the next step is to obtain orientation information. It is important in the practical application to know the screen position that a finger points to rather than the position of the whole hand. To the user, whose expectation is shaped by conventional touchscreens, it seems natural that the 3D, touchless interactivity of the cross capacitance sensing can sense the tip of the finger. However the electrostatics of the cross capacitance sensing involve the whole hand, and the position of the initial guess typically corresponds to the location of the palm or wrist. The location of the tip of the finger can be obtained if a degree of object reconstruction is introduced. To explore the feasibility of this we demonstrate optimizing the position, orientation and axis of an ellipsoidal object against simulated data.

An ellipsoid with axis (a, b, c) centred at z_1, z_2, z_3 is given by

$$\frac{(x_1 - z_1)^2}{a^2} + \frac{(x_2 - z_2)^2}{b^2} + \frac{(x_3 - z_3)^2}{c^2} = 1. \tag{5.5}$$

Orientation is determined using rotation through Euler Angles,

$$R(\alpha, \beta, \gamma) = R_3(\gamma)\ R_2(\beta)\ R_3(\alpha),$$

in which $R_3(\alpha)$ is a rotation around the x_3 axis through angle α and so on (for details

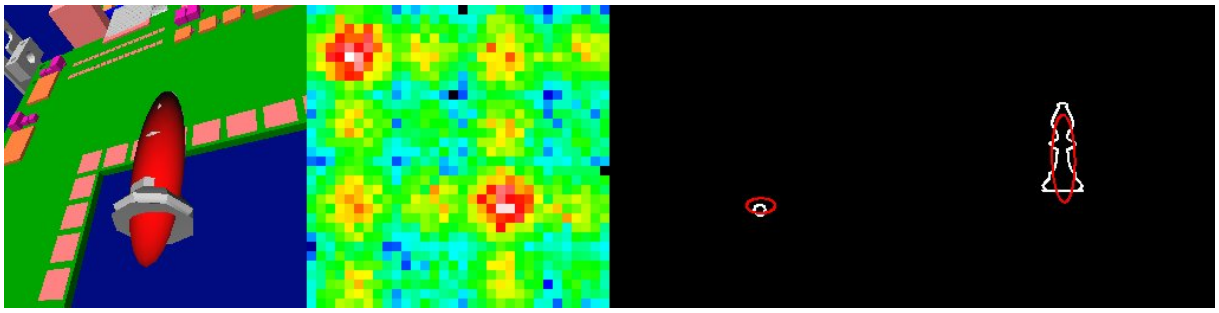

Fig. 5.3: Approximation of the Bishop Object by an optimized ellipsoid based on data from 32 edge electrode data with 20% added noise.

see Arfken [162] page 179). With this, the optimization is done in steps.

- Find the initial guess from the previous section. That is find z_1, z_2, z_3 and the overall size parameter a_o.
- Change the sphere of the initial guess into an ellipsoid with axis $2a_o, a_o/2, a_o/2$
- Optimize Euler angles α and β.
- Optimize position (z_1, z_2, z_3), axis lengths(a, b, c) and Euler angles (α, β, γ) simultaneously.

The third step has only two degrees of freedom and requires advance knowledge of the original shape, a realistic requirement in the example of finger pointing. The third step has 9 degrees of freedom and this may appear to be more than is required, but the use off all axis lengths and all Euler angles avoids local minima in the optimization. Figure 5.3 shows an example of the optimization in this case. A bishop chess object acts here as phantom for a finger. Sensing is done by a linear array of electrodes as in figure 5.2 but slightly less noise is added of 20% of individual signal strength. The difference in 'difficulty' between the third (2 parameters to optimize) and fourth (9 parameters) step is illustrated by the condition number of the Jacobian which is $\kappa(J_2) = 14$ for the third step and $\kappa(J_9) = 1,400$ for the fourth step.

5.3 Spherical Harmonics

In situations with more general objects or where prior information about an ellipsoid is lacking, more degrees of freedom are required. One possibility is to describe shape

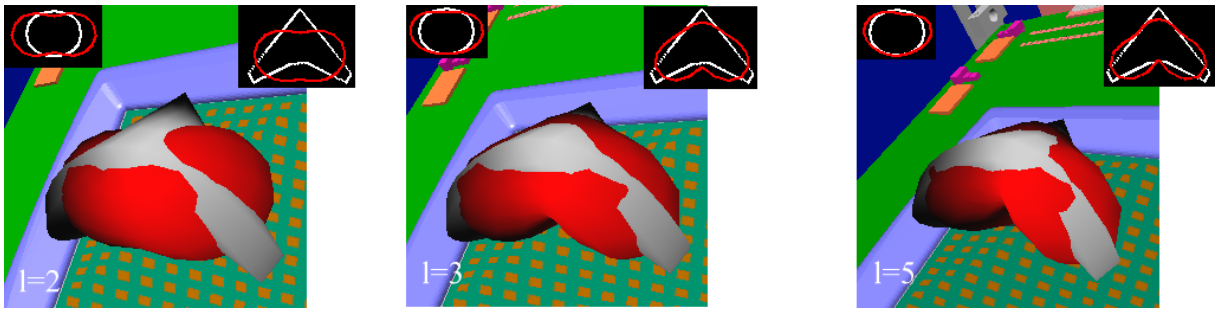

Fig. 5.4: Optimization of Spherical Harmonics from data of a 2D array of 256 electrodes with 5% noise, of a kite shaped object. Illustrated are reconstructions with different Legendre orders.

through spherical harmonics. Sperical harmonics have been used for instance recently in Acoustic scattering[163], Optical Tomography[164], Electromagnetic Scattering[165, 166] and EIT[167].

From a single point, the distance to any point on a surface is described by

$$r(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=0}^{l} \left[C_l^m \cos(m\phi) + S_l^m \sin(m\phi)\right] \; P_l^m(\cos\theta). \tag{5.6}$$

Where $P_l^m(x)$ are Associated Legendre Functions, given by

$$P_l^m(x) = \frac{(-1)^m}{2^l l!}(1-x^2)^{m/2}\frac{d^{l+m}}{dx^{l+m}}(x^2-1)^l. \tag{5.7}$$

Objects that admit this description are called star shapes. The function $r(\theta,\phi)$ is used to calculate the vertices of triangles, using a regular triangulation of (θ,ϕ) on a unit sphere. The spherical harmonics are also used to create the test spoke object introduced above in section 4.5, for which $C_0^0 = 9$ and $S_3^3 = 0.3$.

When used for fitting, the initial guess is used to set the centre of the object and the radius of the initial guess to set the initial value for C_0^0. The fitting parameters C_l^m and S_l^m, excluding the S_l^0 terms, are then optimized with the Gauss-Newton routine `E04FCF`. The program first fits 4 parameters for $l \leq 1$, then 9 parameters for $l \leq 2$, 16 parameters for $l \leq 3$ and so on. This strategy aids convergence in the presence of the rapidly increasing number of fitting parameters. The centre point of the object remains fixed to the value set by the initial guess.

Because the number of parameters increases rapidly with Legendre order, the

method is suitable for situations in which a large number of charge images are collected. Figure 5.4 illustrates a situation in which data is collected from a 2D array of 16×16 electrodes creating a 256×256 forward matrix. Noise at 5% of signal strength is added to the data. A kite shaped object is used as phantom and the figure illustrates the reconstructions for different Legendre orders. As can be seen, higher order optimizations reveal more detail but the rapid increase in the number of parameters makes the fitting of higher orders computationally prohibitively expensive, the optimization can take many hours to perform. The condition numbers of the Jacobian also increases rapidly. In the example of figure 5.4 $\kappa(J_{l=2}) = 52$, $\kappa(J_{l=3}) = 850$ and $\kappa(J_{l=5}) = 30,229$ while the residuals decrease only marginally.

Increased levels of noise reduce the amount of detail in the reconstruction. For instance, increasing the noise level to 10% results in reconstructions similar to figure 5.4A for all orders while setting to noise to zero improves the level of detail recovered in 5.4C.

Although the expansion in spherical harmonics (5.6) describes any star shaped object accurately, from a practical point of view the question arises how many orders are needed for a satisfactory shape description of an object, regardless of the input data, or indeed the sensing technology. Moreover, how well does a finite expansion approximate a shape that is not a star shape? A few numerical experiments that explore these questions are presented in appendix B.

5.4 Discussion

The results in this chapter show that optimization using closed form expressions for the forward matrix and Jacobian provide a very good method of finding the position of an object and its approximate size. The method is robust against noise and is flexible enough to work with arbitrary sensor configurations. The method is obviously limited to the topological configuration of one single object and would fail to indicate the presence of multiple objects, for instance two hands in front of the display screen. It is nevertheless useful in providing an initial guess for reconstruction algorithms discussed later in this thesis.

The optimization method also works well to detect orientation if prior knowledge on the ellipsoidal nature of an object is available. Indications are that moderate noise

levels can be tolerated for this. For the extraction of more detailed shape information, the optimization methods do not appear to work well. The approach of optimizing the coefficients of a low order spherical harmonics expansion does recover some shape information, but the method requires large forward matrices, low noise levels and is very slow. As will be shown in subsequent chapters, better methods are available.

It is worth noting that the interactive display system mentioned in figure 1.3 has an extremely low number of electrodes providing just a 2×2 forward matrix. In that case none of the optimization methods work in a satisfactory way and the hand position is found through a much simpler heuristic method[34].

The heuristic method used for the 2×2 method is simply a 3×4 matrix A that transforms the four element signal vector s into the three element vector of position $x = As$. The elements of the matrix are established through a least squares calibration procedure in which the user holds the finger or hand in pre-determined spatial positions. The advantage of this method is that it requires very little computation.

The initial guess optimization method does not work for the 2×2 sensor configuration but it does work very well for the 14×8 method of 14 transmitters and 8 receivers shown in figure 1.5. The method is efficient and can run at the measurement frame rate (10 f/s) on a high end PC.

An alternative approach to finding the initial guess is suggested by the spherical search method of Kim *et al*[168]. The method was developed for bounded domains, but can be expanded to the unbounded half space used here. In that case, and the notation used here, the method finds for every point z' in the *negative* half space, the distance $r_{z'}$ given by

$$r_{z'} = \frac{\int_{\partial H} c(y)\, \mathrm{d}s(y)}{\int_{\partial H} \frac{c(y)}{|y-z'|}\, \mathrm{d}s(y)}. \tag{5.8}$$

The ball $\partial B(z'; r_{z'})$ divides or cuts through, the object. The intersection of two balls from different points z' finds a line through the object. Three balls find one, or more, points inside the object and so on. The method will not necessarily find the centre nor provide information on the size of the object, but tests indicate that the method could work in satisfactory way with suitable data sets. The drawback of the method is that suitable data sets are defined by the fact that knowledge of the charge image $c(x)$ over the entire plane is required to calculate the integrals (5.8) and that the method does not take advantage of data taken with different active transmitters.

6. ITERATIVE RECONSTRUCTION

In this chapter the inverse problem is solved by identifying ∂D as the zero potential contour of the unique solution of the Cauchy problem for the Laplace equation. The zero potential contour is found with an iterative method. At each iteration there is an approximation ∂D_k to ∂D on which approximate Cauchy data can be calculated by solving a Tikhonov regularized linear system. This data is used to modify ∂D_k by extrapolation towards the zero-surface giving the next approximation ∂D_{k+1}.

Figure 6.1 illustrates the simulated charge image $c(x)$ for a number of axi-symmetric 3D objects. In this case there is a single transmitter at the origin. Using rotational symmetry has the advantage of computational efficiency as well as representational expediency while at the same time remaining close, in terms of the Green's function for example, to the full 3D situation that we are ultimately interested in. For clarity, images are plotted noise free, although in the reconstructions below, 1% or 10% white noise is added. The shape of the three objects is illustrated in the figure.

The charge images for the three different shapes may be different, but it is not immediately clear that these differences are enough to distinguish between them or even reconstruct the different shapes from the charge images alone. The results in this chapter show that it is in fact possible to reconstruct with reasonable accuracy the approximate shapes of simply connected but not necessarily convex objects. As might be expected the shape is more accurately determined near the plane of measurement than away from it.

The iterative reconstruction by analytic continuation method we use here is similar to a method used in inverse acoustic scattering[169, 106, 105, 103]. Here we have a near field charge distribution on the ground plane rather than a far field pattern and we adjust the scattering method[169, 106, 105, 103] by introducing a mixed single and double layer potential description that takes account of both the charge and potential distribution on the iterated object. This leads to a system of coupled equations and a

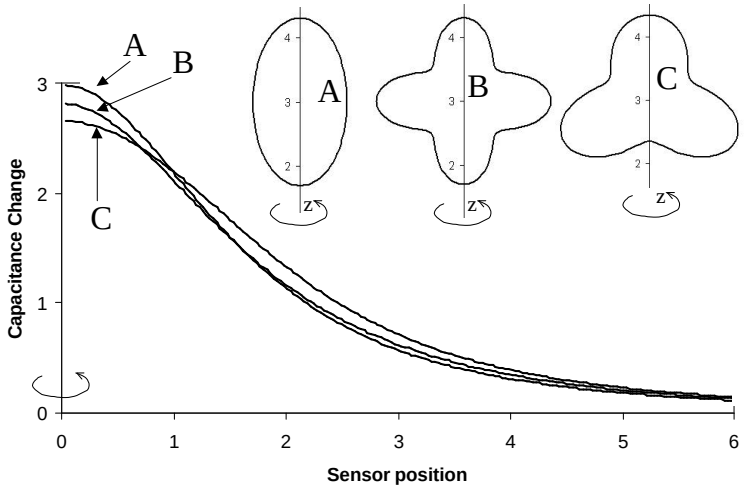

Fig. 6.1: Simulated charge image $c(x)$ for three axi-symmetric objects. Indicated is the axis of rotation and the shape of the three objects used.

natural regularisation requirement in terms of the potential on the object. The method is then further developed here by considering a simplified system of equations and a simplified functional that is more convenient to solve numerically. For this simplified system it is important to adjust the regularisation through a normalisation scheme that we introduce. Regularisation is applied at three levels; in terms of the norms in the penalty terms, in terms of the regularisation parameters and, finally, in terms of the extrapolation to the zero level set that identifies the reconstructed object. We also consider two ways to extend the method towards using data from multiple transmitter experiments to improve the object reconstruction.

6.1 Inverse Charge Imaging

The iterative reconstruction is illustrated in figure 6.2B. At each step k, we first use the data $c(x)$ to find a potential and charge distribution on an object ∂D_k, the current guess. That is the Cauchy date of the plane is taken to Cauchy data on the current guess, and this is illustrated by the set of arrows going from ∂H to ∂D. This ill-posed step is the subject of this section. Once the Cauchy data on the current guess has been obtained, this is then used to deform ∂D_k towards the real object, yielding ∂D_{k+1}. This step is the subject of section 6.2 below.

We make the assumption that we have a current guess ∂D_k which encloses all singularities of the analytic continuation of u. The initial guess can be any one of the optimized shapes described in chapter 5, but in the examples here the initial guess

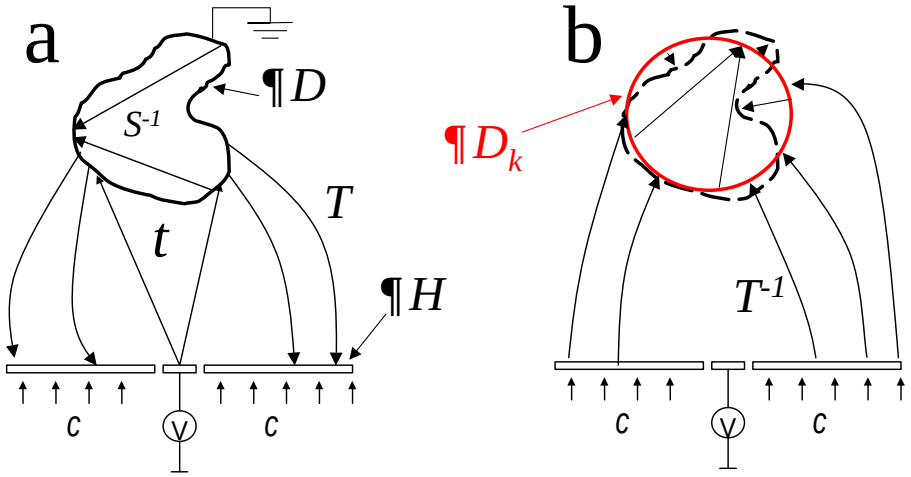

Fig. 6.2: A) Schematic illustration of cross capacitance charge imaging. B) Schematic illustration of the cross capacitance object reconstruction.

described in section 5.1 is used. Writing f_{D_k} for the potential on the guess and g_{D_k} for the charge distribution, we obtain from Green's theorem

$$\begin{aligned} -u(x) \quad &= -\int_{\partial H} f_H(y)\frac{\partial G(x,y)}{\partial y_3}\,\mathrm{d}s(y) + \int_{\partial D_k} f_{D_k}(y)\frac{\partial G(x,y)}{\partial n(y)}\,\mathrm{d}s(y) \\ &\quad -\int_{\partial D_k} G(x,y)g_{D_k}(y)\,\mathrm{d}s(y)\;, \quad x \in H \setminus \bar{D}_k \end{aligned} \tag{6.1}$$

In contrast to (3.19), the double layer potential contribution on ∂D_k has now been included, because the assumption that $u(x) = 0$ does not hold on ∂D_k.

We can evaluate (6.1) at the boundaries to obtain a system of equations for the boundary conditions. Specifically we evaluate (6.1) on the boundary of the object to find an equation for f_{D_k}. Here we need to introduce the factor 1/2 to account for the evaluation on the boundary. The normal derivative of (6.1) on the ground plane can be used to obtain an expression for g_H. This then gives the following system of two simultaneous integral equations for the two unknown functions f_{D_k} and g_{D_k}, from

which the functions can be determined,

$$
\begin{aligned}
-\frac{1}{2} f_{D_k}(x) \quad &= -\int_{\partial H} f_H(y) \frac{\partial G(x,y)}{\partial y_3} \,\mathrm{d}s(y) + \int_{\partial D_k} f_{D_k}(y) \frac{\partial G(x,y)}{\partial n(y)} \,\mathrm{d}s(y) \\
&\quad - \int_{\partial D_k} G(x,y) g_{D_k}(y) \,\mathrm{d}s(y) \;, \quad x \in \partial D_k \qquad (6.2) \\
g_H(x) \quad &= -\int_{\partial H} f_H(y) \frac{\partial^2 G(x,y)}{\partial y_3 \partial x_3} \,\mathrm{d}s(y) + \int_{\partial D_k} f_{D_k}(y) \frac{\partial^2 G(x,y)}{\partial x_3 \partial n(y)} \,\mathrm{d}s(y) \\
&\quad - \int_{\partial D_k} \frac{\partial G(x,y)}{\partial x_3} g_{D_k}(y) \,\mathrm{d}s(y) \;, \quad x \in \partial H. \qquad (6.3)
\end{aligned}
$$

In previous chapters of this thesis we have already introduced several operator and function definitions for terms in this system of equations. We now also define

$$
\begin{aligned}
v(x) &= f_{D_k}(x) && (6.4) \\
q(x) &= -g_{D_k}(x) && (6.5) \\
t(x) &= \int_{\partial H} f_H(y) \frac{\partial G(x,y)}{\partial y_3} \,\mathrm{d}s(y). && (6.6)
\end{aligned}
$$

In addition to

$$
\begin{aligned}
(Kv)(x) &= \int_{\partial D_k} f_{D_k}(y) \frac{\partial G(x,y)}{\partial n(y)} \,\mathrm{d}s(y) && (6.7) \\
(Sq)(x) &= -\int_{\partial D_k} G(x,y) g_{D_k}(y) \,\mathrm{d}s(y) && (6.8) \\
c(x) &= \int_{\partial H} f_H(y) \frac{\partial^2 G(x,y)}{\partial y_3 \partial x_3} \,\mathrm{d}s(y) - g_H(x) && (6.9) \\
(Mv)(x) &= \int_{\partial D_k} f_{D_k}(y) \frac{\partial^2 G(x,y)}{\partial x_3 \partial n(y)} \,\mathrm{d}s(y) && (6.10) \\
(Tq)(x) &= -\int_{\partial D_k} \frac{\partial G(x,y)}{\partial x_3} g_{D_k}(y) \,\mathrm{d}s(y). && (6.11)
\end{aligned}
$$

With these the equations (6.2) and (6.3) become

$$
\begin{pmatrix} \frac{1}{2} I + K & S \\ M & T \end{pmatrix} \begin{pmatrix} v \\ q \end{pmatrix} = \begin{pmatrix} t \\ c \end{pmatrix} \qquad (6.12)
$$

The operators and functions defined above have close parallels in matrices and vectors

defined in chapter 4. Indeed 6.12 can either be read as an operator equation or as a block matrix equation.

The system describes the mapping of the Cauchy data on the object (v, q) to the functions (t, c) which are derived from the Cauchy data on the sensor plane. The inverse problem of finding (v, q) from (t, c) is the ill-posed problem of finding Cauchy data on a part of the boundary.

It is therefore necessary to use regularisation to solve (6.12) to find (v, q) from (t, c). Although a number of different regularisation techniques exist, see for instance [170, 171, 172, 173] for recent examples, in this case it is natural to use Tikhonov regularisation. Provided that ∂D_k is close to the real object, the potential on ∂D_k is small. The norm $||v||$ must therefore be small and this requirement is added to the solution to (6.12). Hence v, q are found through minimization of the Tikhonov functional:

$$||\frac{1}{2}v + Kv + Sq - t||^2 + ||Mv + Tq - c||^2 + \alpha_v^2||v||_P^2 \tag{6.13}$$

In which α_v is the Tikhonov regularisation parameter and the first order Sobolev norm is given by

$$||f||_P^2 = (f, Pf) = \int_{\partial D} f^2 \, \mathrm{d}s + \int_{\partial D} \nabla f \cdot \nabla f \, \mathrm{d}s.$$

Here∇ is the grad in the plane of the surface element of ∂D only. As mentioned in chapter 4, in operator terms, $P = I + L^T L$ and the discrete matrix approximation for L of ∇ was given in equation (4.19). We note that there is only overt regularisation for v in (6.13), control on the norm of q is provided the second term in functional (6.13). In numerical experiments we have found that minimizing this functional works well to recover approximate Dirichlet and Neumann data on a guess ∂D_k. In the next section we will discuss how this data can be used to deform the object ∂D_k towards an improved approximation ∂D_{k+1}.

One feature of the regularisation is that, for noise free data and for the final object $\partial D_k = \partial D$, the Tikhnov functional solves the genuine Cauchy problem and not some problem close to it. This is a feature also seen in iterated Tikhonov regularisation[174, 175] in which the solution at a previous iteration step is used as the prior for the current step. As the iterated Tikhonov proceeeds, the regularisation terms fades to zero. Here we will see the same effect, though not as in [174, 175] for the iterated Tikhonov solution of a single problem, but over multiple problems for different objects

∂D_k.

We also consider a simplified functional

$$||v + Sq - t||^2 + 4||Tq - c||^2 + \alpha_q^2||q - q_p||^2_{P_Q} \tag{6.14}$$

The Kv and Mv terms in (6.13) are approximated with $\frac{1}{2}v$ and $qT - c$ respectively. This substitution is motivated by numerical efficiency. No K and M matrices have to be calculated and the first term in the functional can be minimized independently of the last two. The simplification can be justified by the fact that as ∂D_k approaches the correct ∂D, the variation in f_{D_k} on ∂D_k becomes small. Using lemma C.5 we can approximate

$$(Kv)(x) \approx f_{D_k}(x) \int_{\partial D_k} \frac{\partial G(x,y)}{\partial n(y)} \,\mathrm{d}s(y) = \frac{1}{2}v(x)$$

At the same time, it can be shown that in the second term of (6.13), $qT - c \approx Mv$ as $\partial D_k \to \partial D$.

Lemma 6.1.

$$\lim_{\partial D_k \to \partial D} (Tq - c) \to Mv$$

Proof.

$$\begin{aligned} Tq - c &= \int_{\partial D_k} \frac{\partial G(x,y)}{\partial x_3} g_{D_k}(y) \,\mathrm{d}s(y) - \int_{\partial D} \frac{\partial G(x,y)}{\partial x_3} g_D(y) \,\mathrm{d}s(y) \\ &= \int_{\partial D_k} \frac{\partial G(x,y)}{\partial x_3} g_{D_k}(y) \,\mathrm{d}s(y) - \int_{\partial D_k} \frac{\partial G(x,y+\delta)}{\partial x_3} g_D(y+\delta) \,\mathrm{d}s(y+\delta). \end{aligned}$$

Because u is harmonic between D and D_k, we have $g_{D_k}(y) \,\mathrm{d}s(y) = g_D(y+\delta) \,\mathrm{d}s(y+\delta)$ and so we have

$$\begin{aligned} \lim_{\partial D_k \to \partial D} (Tq - c) &= -\int_{\partial D_k} \frac{\partial^2 G(x,y)}{\partial x_3 \partial n(y)} \delta g_{D_k}(y) \,\mathrm{d}s(y) \\ &= \int_{\partial D_k} f_{D_k}(y) \frac{\partial^2 G(x,y)}{\partial x_3 \partial n(y)} \,\mathrm{d}s(y) \\ &= Mv. \end{aligned}$$

□

We note that the second term in (6.14) is equivalent to (2.9) in section 2.3.2 on

the two step method for inverse scattering. Indeed the entire functional (6.14) is now equivalent to a regularised version of that method, though we have arrived at this as an approximation of (6.13). With regard to the regularisation, the second term in (6.14) is independent of $v(x)$ and an alternative regularisation is required to that used in (6.13). We now require that the charge distribution q on ∂D_k is close to a charge distribution corresponding to $v(x) = 0$ on ∂D_k. This implies a prior for the charge distribution that satisfies (3.20) on ∂D_k, that is $q_p(x) = (S^{-1}t)(x)$.

Moreover, the penalty term is given by a *normalized* first order Sobolev norm on the prior.

$$||f||^2_{P_Q} = (f, P_Q f) = \int_{\partial D} \left(\frac{f}{q_p}\right)^2 \mathrm{d}s + \int_{\partial D} \nabla\left(\frac{f}{q_p}\right) \cdot \nabla\left(\frac{f}{q_p}\right) \mathrm{d}s \tag{6.15}$$

In operator terms we now have $P_Q = Q_p^{-1} P Q_p^{-1}$ and Q_p^{-1} is given by

$$(Q_p^{-1} f)(x) = \frac{f(x)}{q_p(x)} \qquad x \in \partial D_k$$

we note that $||Q_p^{-1}|| = \sup 1/q_p(x) > 0$. The reason for the use of the normalisation operator Q_p^{-1} is that $q_p(x)$ will vary by several orders of magnitude over an object and will be smallest on the parts of ∂D_k furthest away from the transmitter and sensor plate. It is there where the instability in $q(x)$ is greatest and a pre-conditioned penalty term is required to constrain the solution there. Another way of interpreting the use of the $Q_p^{-1} P Q_p^{-1}$ norm is that in a statistical sense the prior has a covariance Q_p^2 and an expected value of $q_p(x)$ [176].

6.2 Object Reconstruction

Having obtained the potential and charge distribution we now seek to deform the current guess towards the real object guided by the knowledge that this real object is at ground potential. One method that readily suggests itself is to use (6.1) to find a zero contour, or at least a minimum contour, near the object and identify that as our next best guess. Figure 6.3 provides an illustration of the current guess ∂D_k and the (unknown) object ∂D to be constructed. A variation on the method used in acoustic scatteric would be use the potential v and the gradient q to make a extrapolation from

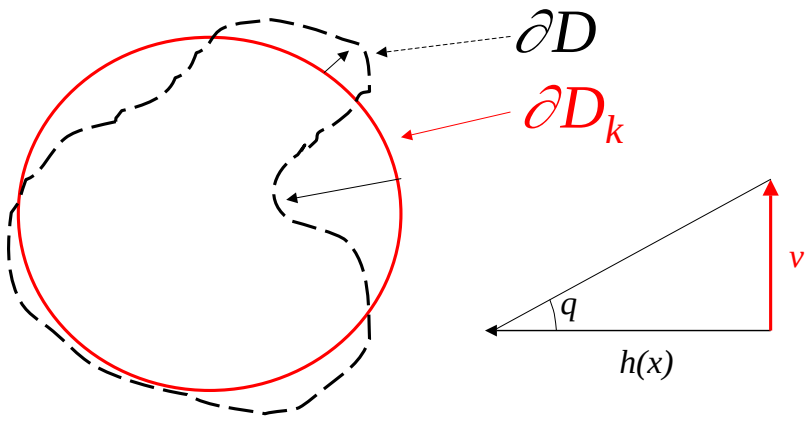

Fig. 6.3: Illustration of the object reconstruction, using the potential on the current guess, to extrapolate to the zero potential contour.

∂D_k towards the zero potential contour. That is define a dilation function $h(x)$ which gives the deformation at x on the object along the outward pointing normal, given by

$$h(x) \propto -\frac{v(x)}{q(x)} \quad x \in \partial D_k.$$

In numerical experiments we have found that this does not work well. The reason is that, because $q(x)$ appears in the denominator, extrapolating towards the zero potential contour will induce the strongest changes in ∂D_k where $q(x)$ is smallest, i.e. those parts for which the conditioning of (6.12) is worst. It is therefore necessary to pre-condition the extrapolation with the operator Q_p. For the linear extrapolation this then leads to a simple dilation function $h(x)$ which gives the deformation at x on the object along the outward pointing normal as

$$h(x) = -\lambda v(x) \quad x \in \partial D_k. \tag{6.16}$$

The attenuation, or relaxation, factor λ avoids overshoot and it is a second, independent regularisation parameter.

6.3 Stopping Criterion

We thus have a cross capacitance reconstruction algorithm in which the shape is found from a level set evolution driven by the potential u itself. Level sets have been used in a variety of shape reconstruction problems [113, 177, 178]. Briefly stated the algorithm here starts with an initial guess and then, through repeated application of either eqns (6.13) and (6.16), or eqns (6.14) and (6.16), evolves the object until a stopping criterion is reached.

A common method for deciding when a best fit has been obtained is the so called 'Morozov Discrepancy Principle', which states that a good fit is obtained when the difference between measured and fitted image, the so called cost function, is similar to the noise in the measured image. A drawback of this criterion is however that it requires prior knowledge of the noise in the measurements and is in any case only valid for true white noise. Instead we use the change in cost function for our stopping criterion. Thus the iteration is stopped when the iteration to iteration change in the cost function falls below a certain threshold. The criterion is expressed as

$$\left| \frac{r_k - r_{k-1}}{r_k} \right| < E. \tag{6.17}$$

In which E is the threshold. The cost function, or residue, r_k at iteration k is defined in terms of the piecewise constant elements of the charge image used in the numerical experiments,

$$r_k = \sqrt{\frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \left(\frac{\Lambda_{D_k i,j} - \Lambda_{Di,j}}{\Lambda_{D_k i,j} + \Lambda_{Di,j}} \right)^2}. \tag{6.18}$$

Which is well defined as $\Lambda_{Di,j} > 0$. Strictly speaking the stopping criterion is only sensible and can only be guaranteed to stop the iteration if r_k is strictly decreasing with iteration number. Currently we lack a formal proof for this, but a strictly decreasing behaviour in the numerical experiments is always observed. The threshold E is an additional regularisation parameter. Choosing E too large has obvious drawbacks, but also at excessively small values poor reconstruction is obtained. We have found that a value of $E = 0.01$ worked well in all numerical experiments. We stress that with this construction, the actual noise level is not an input parameter to the stopping criterion.

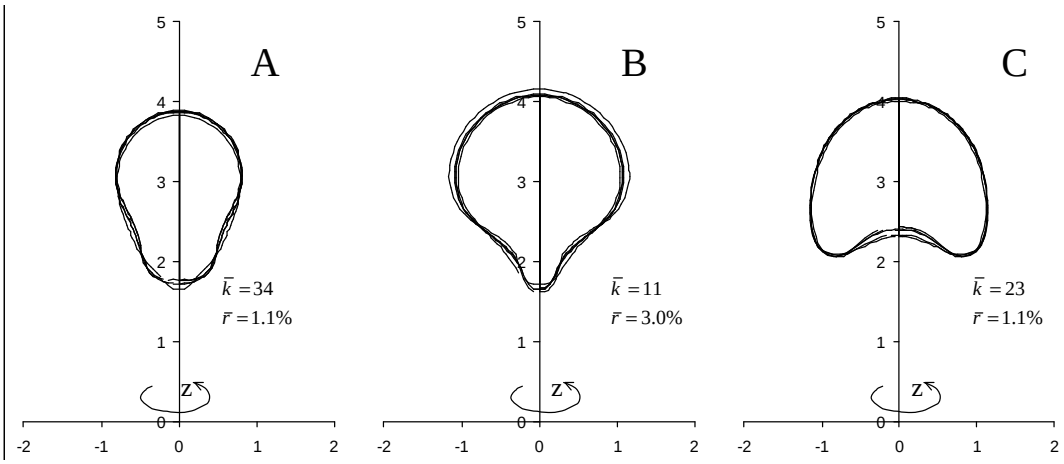

Fig. 6.4: Reconstructions from the charge images in figure 6.1 using functional (6.13). The average iteration number ($\bar{k}$) and residue at termination ($\bar{r}$) are shown in insets.

6.4 Parameter Scaling

Before we turn to the numerical results in the next section a refinement to equations (6.13)-(6.16) is made by introducing scaled regularisation parameters $\hat{\alpha}$ and $\hat{\lambda}$ that are insensitive to the scale and discretization level and can be used in either functional (6.13) or (6.14)

$$\alpha_v = \hat{\alpha}\frac{||M||}{\sqrt{||P||}} \tag{6.19}$$

$$\alpha_q = \hat{\alpha}\frac{||T||}{||Q_p^{-1}||\sqrt{||P||}} \tag{6.20}$$

$$\lambda = \frac{\hat{\lambda}}{||Q_p||} \tag{6.21}$$

These factors allow comparison of the effectiveness of the penalty terms in functionals (6.13) and (6.14). The factors $\hat{\alpha}$ and $\hat{\lambda}$ are used as input to the cross capacitance object reconstruction and the values of λ, α_v or α_q and are then calculated once only, using the initial guess to calculate the operator norms. For the results in section 6.5, the values for α_v and α_q then remain fixed through out the iterative procedure. For the results in section 6.6, however, an L-curve criterion[179] is used to establish an optimal value for the regularisation parameter. The L-curve is one method among a collection

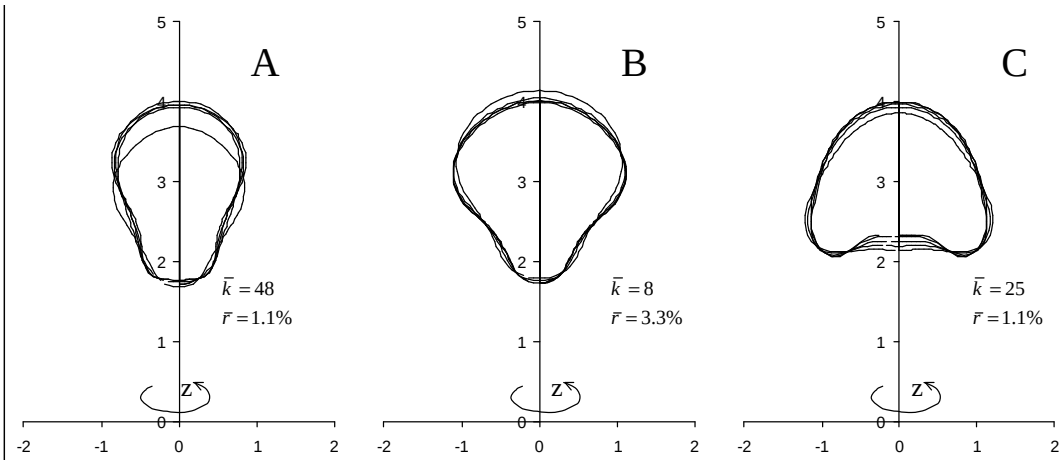

Fig. 6.5: Reconstructions from the charge images in figure 6.1 using functional (6.14). The average iteration number ($\bar{k}$) and residue at termination ($\bar{r}$) are shown in insets.

of methods for choosing the regularisation parameters such as Morozov's discrepancy principle or generalized cross validation. See [180, 181] for reviews. Though discredited as a universal method by some counter examples[182], the L-curve remains a much used method.

6.5 Numerical Results

Numerical simulations were performed on a personal computer using c++ code with NAG library support. The simulations are calculated using a Boundary Element Method (BEM) formulation [137, 183] as described in chapter 4. One variant of the code was implemented for an axi-symmetric problem in which the object and the sample points of $c(x)$ have rotational symmetry around the x_3-axis. This reduces the problem to a two dimensional one in which the Green's function $G(x, y)$ and its derivative on the ground plane are expressed in complete elliptic integrals of the first and second kind [184]. The second variant of the code implemented the full 3D problem for non-symmetric triangulated wire frame models of the objects.

Figure 6.4 illustrates reconstructions using functional (6.13) for the axi-symmetric objects of the (axi-symmetric) charge images illustrated in figure 6.1. Each object was reconstructed five times for different draws of 1% normally distributed noise added to the input image. The objects were defined in 51 linear line segments and the charge

image $c(x)$ was sampled at 200 equidistant radial points from $\rho = 0$ to $\rho = 6$. In each case a unit sphere centred at $x_3 = 3$ was used as initial guess. Fixed regularisation parameters $\hat{\alpha} = 1$ and $\hat{\lambda} = 1$ were used.

The objects and reconstructions show the variation in reconstruction attributable to the noise in the input images and the reconstructions here have been chosen to illustrate both what can be and what cannot be reconstructed. As is perhaps obvious, no significant reconstruction is achieved on the side of the object facing away from the sensor plane. However, fair reconstruction is obtained at the facing side.

Figure 6.5 illustrates reconstructions using simplified functional (6.14) for the same axi-symmetric objects. Again each object was reconstructed five times for different draws of 1% Gaussian noise added to the input image and regularisation parameters $\hat{\alpha} = 1$ and $\hat{\lambda} = 1$ were again used. As can be seen, the reconstructions are very similar to those shown in figure 6.4. These reconstructions typically took a third of the CPU time of the reconstructions shown in figure 6.4. We stress that the same $\hat{\alpha}$ and $\hat{\lambda}$ were used for the results in Figures 6.4 and 6.5. The actual regularisation parameters α_v and α_q used in functionals (6.13) and (6.14) differed by two orders of magnitude. Results at $3\times$ discretization or $100\times$ scale, again with the same $\hat{\alpha}$ and $\hat{\lambda}$, but with very different values of α_v, α_q and λ, yield virtually identical results. These observations support the approximate equivalence of the functionals and the scaling of the regularisation parameters.

The results in figure 6.6 explore the effect of the input noise and the Tikhonov regularisation parameter $\hat{\alpha}$ on the reconstruction. Shown are three reconstructions of object C in figure 6.1, each again for 5 draws of the input noise, which is this time set at 10% rather that 1%. The reconstruction in figure 6.6A was done with the same regularisation parameters as before ($\hat{\alpha} = 1$, $\hat{\lambda} = 1$) and shows that increased noise on the input data deteriorates the reconstruction. This deterioration is particularly noticeable as an increased variation in the reconstructed object between noise draws. In figure 6.6B this variation is reduced by increasing the Tikhonov regularisation parameter ($\hat{\alpha} = 10$, $\hat{\lambda} = 1$). Figure 6.6C shows that increasing the regularisation parameter further ($\hat{\alpha} = 100$, $\hat{\lambda} = 1$) reduces the variation to zero but also obliterates meaningful reconstruction. Note that the average error at termination is close to the 10% input noise level for all results in figure 6.6.

To probe the numerical results further, figure 6.7 illustrates the cost function r_k of

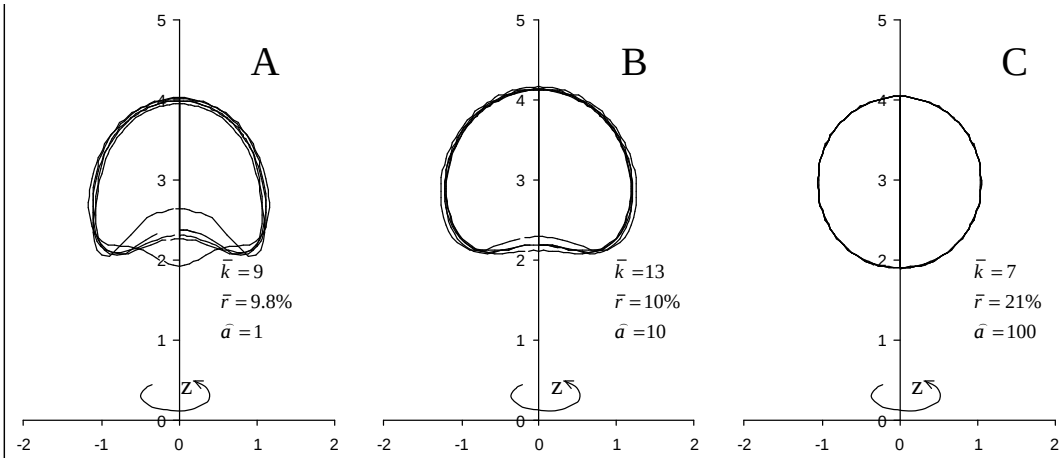

Fig. 6.6: Reconstructions at 10% input noise level for different regularisation parameters using functional (6.13). The average iteration number ($\overline{k}$) and residue at termination ($\overline{r}$) and the regularisation parameter ($\hat{\alpha}$) are shown in insets.

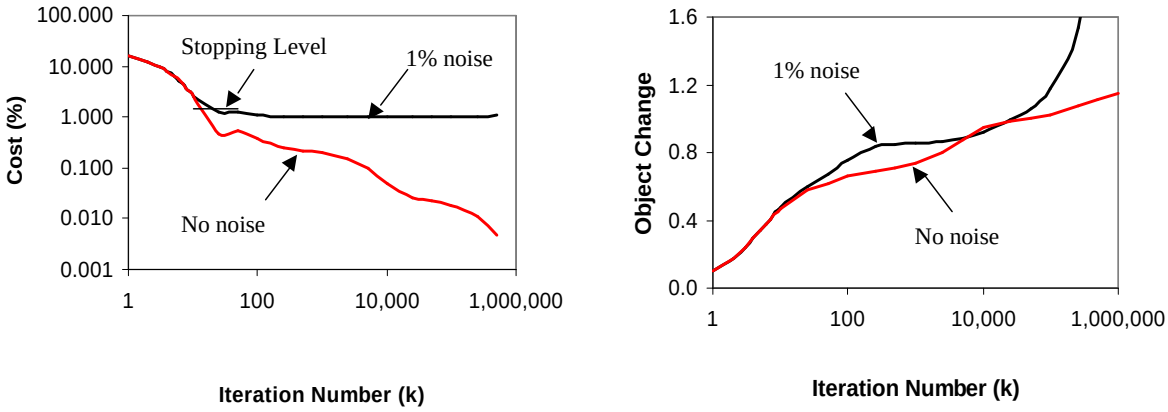

Fig. 6.7: Cost function (A) and cumulative dilation (B) for the reconstruction of object C in figure 6.1 from noisy and noiseless data.

(6.18) and the cumulative dilation of object C in figure 6.1. The cumulative dilation at iteration k is defined as

$$\sum_{j=1}^{k} ||h_j(x)|| \tag{6.22}$$

It is intended as a qualitative measure of the object change. Figure 6.7 illustrates these functions for iterations well beyond the stopping criterion used for the examples in figures 6.4 and 6.5. As can be seen, the cost function reaches a plateau when it reaches the noise level. The cumulative object change also reaches a plateau, but then, for very large iteration numbers, the object change increases rapidly on the log iteration scale of figure 6.7 and the object disintegrates. Also illustrated is the case for a reconstruction from a noiseless charge image. Clearly the cost function reaches much lower levels and a closer fit to the original object is achieved. In this case too however, beyond the scale shown here, the reconstruction eventually goes to destruction.

6.6 Multiple Transmitters

In the situation of practical interest the experimental data are obtained from a linear array of electrodes arranged around the edge of a display surface, instead of being distributed over the whole surface of the display itself. Each electrode i can function as transmitter and a charge image c_i is collected from all electrodes in each case. In this way n separate charge images are obtained time sequentially. This influences the object reconstruction in a number of ways. On the one hand the reduced aperture of the measurements will make the problem harder. On the other hand, the multiple charge images c_i provide a richer data set in analogy with multiple incident wave directions in acoustic scattering[103] or multiple illumination sources in machine vision.

Now (6.12) must be solved for each transmitter i, using simplified functional (6.14), we obtain the sum

$$\sum_i ||v_i + Sq_i - t_i||^2 + 4||Tq_i - c_i||^2 + \alpha_i^2 ||q_i - q_{p,i}||^2_{P_{Q,i}} \tag{6.23}$$

In which the sum is over n transmitter experiments and a subscript has been added to each quantity that depends on the specific transmitter being used. In particular the regularisation parameter α_i is determined using an L-curve criterion for each transmit-

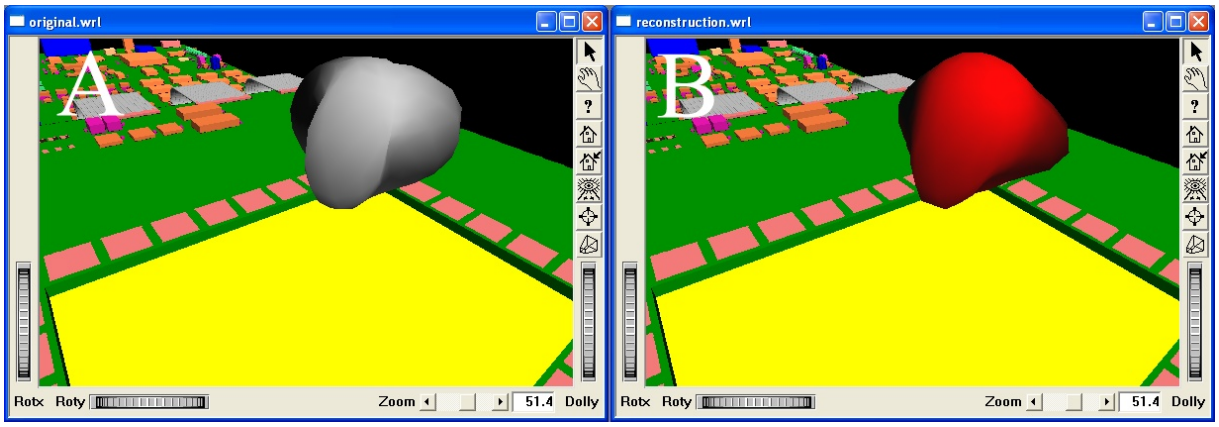

Fig. 6.8: Perspective views of original (A) and reconstructed object (B) above a model experimental system.

ter problem independently.

Functional (6.23) is minimized by n independent under-determined regularised solutions q_i which provide n estimates of v_i. The dilation function is now given by

$$h(x) = -\lambda \frac{1}{n} \sum_i v_i(x) \quad x \in D_k \tag{6.24}$$

Figure 6.8 provides an example of a reconstruction for this situation. Figure 6.8A illustrates a simulated object located 30mm above the model of a 4" display surrounded by 32 electrodes. The charge images c_i for each transmitter where calculated as described in chapter 4 consisting typically of 420 triangles. Doubling or halving the triangle mesh does not affect the results. The computer code used is described in detail in chapter 4 and is quite separate from the axi-symmetric code used to calculate the results shown in the previous sections of this chapter. 1% noise was added to the measurements from the 32 edge electrodes. The white object in figure 6.8A illustrates the original object while the object in figure 6.8B illustrates the reconstructed object. An initial guess obtained from Gauss-Newton optimization of the position and size of a sphere as described in section 5.1, was used and 12 iterations applied subsequently. Clearly the object is reconstructed in the correct place with some shape features. If the object is moved and rotated above the display, the location and orientation continue be to reconstructed correctly.

Figure 6.9 provides more detail on the shape reconstruction. Figure 6.9A shows cross sections in the $x_3 = 30mm$ plane and 6.9B shows cross sections in the $x_2 = 0$ plane. The full line illustrates the original object and the dashed lines illustrate five reconstructions for different draws of the noise. As before, $\hat{\lambda} = 1, E = 0.01$. The geometric average of the regularisation parameter $\overline{\hat{\alpha}} = 0.4$.

Here only $\hat{\alpha}$ is determined with a L-curve criterion. It may be possible to also determine the reguralisation parameters $\hat{\lambda}$ and E in a multidimensional generalized L-curve framework as proposed by Belge *el al*[185]. However this approach will be computationally expensive. It must also be pointed out that use of the L-curve does not actually improve the quality of reconstruction. Similar results to those in figures 6.9A and 6.9B are obtained with $\hat{\alpha}$ fixed at 0.1. The L-curve criterion mainly facilitates the process of finding the best regularisation parameters. Behaviour as illustrated in figure 6.6 was observed for increased regularisation or noise.

The approach above solves the inverse problems for each charge image c_i independently using $q_{p,i} = S^{-1}t_i$ for each case. An alternative would be a Kaczmarz like approach. The Kaczmarz approach is an iterative method used, for instance, in Computerized Tomography[186] in which an orthogonal projection of the current solution is applied into the affine subspace of the set of equations provided by the next incident direction. Here we can use the solution (v_{i-1}, q_{i-1}) for transmitter $i-1$ to provide a prior for the solution (v_i, q_i) at transmitter i. That is, in (6.23) we replace $q_{p,i}$ with

$$q_{K,i} = \tilde{S}^{-1}(t_i + v_{i-1}) \tag{6.25}$$

This uses the fact that the dilation function (6.16), of the solution (v_{i-1}, q_{i-1}) should be identical to the dilation function of the next solution (v_i, q_i). Hence rather than using $v_i = 0$ as prior information, we use $v_i = v_{i-1}$ instead for $i \geq 2$. For $i = 1$ the prior $v_1 = 0$ is used. The dilation is now not calculated from a sum as in (6.24) but only from the last solution. That is, $h = -\lambda v_n$. Normalisation of the regularisation as defined in (6.20) and (6.21), continues to be used.

Figures 6.9C and 6.9D illustrate cross sections of the reconstructed object for this case. Here $\overline{\hat{\alpha}} = 1.8$, and $\hat{\lambda} = 1$.

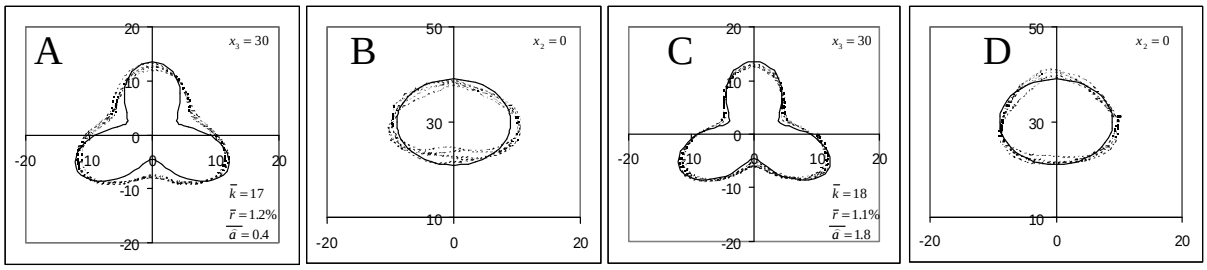

Fig. 6.9: Cross sections of the orginal (bold) and reconstructed object (dashed) 3D object from edge capacitance data with multiple transmitters. A and B: using (6.24), C and D: using (6.25)

6.7 Conclusions

In this chapter we have presented an iterative algorithm with three levels of regularisation that can recover shape information from noisy data. The algorithm is regularised by 1) the assumption that the initial guess is close to the real object, 2) by the relaxation parameter in the deformation of the guess and 3) by the stopping criterion. We have shown that a simplification can be made to the minimization functional, which yields similar reconstruction results but makes the numerical execution three times faster. We have introduced a scaling of the regularisation parameters that makes the algorithm robust across a range of object and sensor sizes, as well as different discretization levels.

The reconstruction works for realistic sensor configurations.

7. INDICATOR FUNCTION

In this chapter we are interested in fast algorithms that use an indicator function to test if a single point lies inside or outside the object. The setup is schematically illustrated in figure 7.1a. As before, the grounded object is located above the plane in which voltage drive electrodes and current measurement electrodes are embedded. The voltage drive electrodes are termed transmitters and the measurement electrodes, which are held at virtual ground, are called receivers. The algorithms are useful in a situation when multiple sequential measurements using different transmitter electrodes provide knowledge of the $f_H \to c$ mapping for arbitrary $f_H \in L^2(\partial H)$. The so-called Dirichlet to Neumann map.

The following section sets out the Factorisation Method which establishes a link between a test function and the range of the forward Dirichlet to Neumann map. Based on this, section 7.2 sets out the practical algorithm for object reconstruction, which is tested in reconstructions from synthetic data. Both idealised synthetic data are used and noisy data from limited measurements that simulate what might be obtained with a practical system. Section 7.4 examines the numerical experiments more closely and provides insight into the range test that we have employed.

7.1 The Factorisation Method

We recall from chaper 3 the single layer operator T and the trace of the single layer operator $\tilde{S}$, as well as the defintion of $\mathcal{G}$ as the solution operator of the external Dirichlet problem $\mathcal{G} : L^2(\partial D) \to L^2(\partial H)$ which, by lemma 3.5 is compact and injective. With these operators we obtained the factorisations

$$\Lambda_D = T\tilde{S}^{-1}T^* = \mathcal{G}S\mathcal{G}^* \tag{7.1}$$

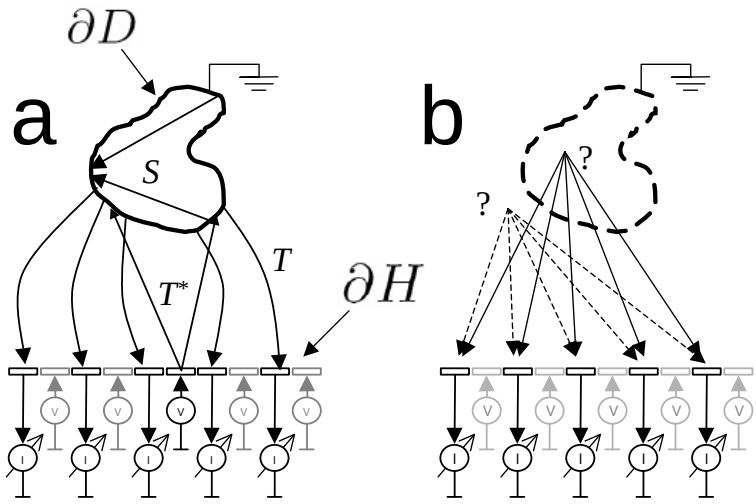

Fig. 7.1: Schematic illustration of the reconstruction in which a test function is used to establish if points lie inside or outside the unkown domain.

These factorisations solved the forward problem. Assuming that Λ_D is obtained from measurements, the inverse problem is to reconstruct the object from knowledge of the forward map Λ_D. We do this by applying to the situation here, with our own test function, the Factorisation Method that Kirsch[116], Hähner[117] and Brühl[118], developed for the Helmholz equation in electromagnetic or acoustic scattering and for the Laplace equation in EIT. We note in particlar that here the Factorisation method used for object reconstruction in an unbounded half space without the restriction to boundary date in a weighted Sobolev space[53, 119, 187].

We recall from lemma 3.8 that the linear operator Λ_D is self-adjoint, compact and positive definite. The spectral theorem for compact, self-adjoint and positive definite operators now implies that Λ_D generates a singular system (λ_n, f_n) in which (f_n) is an orthonormal basis in L^2 and (λ_n) a sequence of non-increasing, positive real eigenvalues. The fact that Λ_D has a discrete spectrum cleary relies on the fact that the object D is bounded. Moreover Λ_D admits a square root;

$$\Lambda_D^{1/2} f = \sum_n \sqrt{\lambda_n}(f, f_n) f_n$$

In this electrostatic setting it is possible to link the range of $\Lambda_D^{1/2}$ with that of a mapping $\partial D \to \partial H$, in a relatively straightforward manner[117].

Theorem 7.1.

$$\mathcal{G}S^{1/2}(L^2(\partial D)) = \Lambda_D^{1/2}(L^2(\partial H)) \subset L^2(\partial H)$$

Proof. From the orthonormal basis (f_n) on ∂H it is possible to define (φ_n) on ∂D

$$\varphi_i = \frac{1}{\sqrt{\lambda_i}} S^{1/2} \mathcal{G}^* f_i$$

It is easily verified that $(\varphi_i, \varphi_j) = \delta_{i,j}$. Moreover, let $\varphi \in L^2(\partial D)$ satisfy $(\varphi, \varphi_i) = 0$ for all i. Then, through

$$0 = (\varphi, \varphi_i)_{\partial D} = (\mathcal{G} S^{1/2} \varphi, f_i)_{\partial H}$$

this implies that $\mathcal{G} S^{1/2} \varphi = 0$ and therefore $\varphi = 0$ by the injectivity of $\mathcal{G}$ and S. Therefore, (φ_n) is an orthonormal basis[116, 117].

Take $g \in \mathcal{G} S^{1/2}(L^2(\partial D))$, there is a $\varphi \in L^2(\partial D)$ for which $g = \mathcal{G} S^{1/2} \varphi$. Because $\mathcal{G} S^{1/2}$ is bounded, we have

$$g = \sum_n (g, f_n) f_n = \sum_n (\mathcal{G} S^{1/2} \varphi, f_n) f_n = \sum_n (\varphi, S^{1/2} \mathcal{G}^* f_n) = \sum_n \sqrt{\lambda_n} (\varphi, \varphi_n) f_n.$$

Which shows that $\mathcal{G} S^{1/2} : L^2(\partial D) \to L^2(\partial H)$ can be expressed by a singular system $(\sqrt{\lambda_n}, \varphi_n, f_n)$. There is now a function $f \in L^2(\partial H)$ with fourier components $(f, fn) = (\varphi, \varphi_n)$ for which $g = \Lambda_D^{1/2} f$ and therefore $g \in \Lambda_D^{1/2}(L^2(\partial H))$. Thus $\mathcal{R}(\mathcal{G} S^{1/2}(L^2(\partial D))) = \mathcal{R}(\Lambda_D^{1/2}(L^2(\partial H)))$. □

A more general version of theorem 7.1 can be demonstrated (proposition 2.18 in [181]) and the result has recently been set in a more general context of elliptic problems[188]. However the relative simplicity of the proof for compact operators is sufficient here.

We now introduce as test function the kernel of the operator T;

$$g_z(y) = \frac{\partial G(y, z)}{\partial y_3}, \quad y \in \partial H \tag{7.2}$$

Theorem 7.2. *For $z \in H$ the equation $\Lambda_D^{1/2} f = g_z$ has a solution $f \in L^2(\partial H)$ if and only if $z \in D$.*

Proof. We first establish that $\tilde{S}^{1/2}(H^{-1/2}(\partial D)) \subset L^2(\partial D)$. Take $\phi, \varphi \in H^{-1/2}(\partial D)$,

then $\tilde{S}\phi, \tilde{S}\varphi \in H^{1/2}(\partial D)$. We have

$$\left(\tilde{S}^{1/2}\phi, \tilde{S}^{1/2}\varphi\right)_{L^2} = \left(\phi, \tilde{S}\varphi\right)_{L^2} < \infty,$$

because L^2 is the pivot space for $H^{-1/2}$ and $H^{1/2}$. For $z \in D$, we have $g_z = \mathcal{G}(G(\cdot, z)|_{\partial D})$, $\tilde{S}$ has an inverse and there is a $\phi \in H^{-1/2}(\partial D)$ such that

$$G(\cdot, z)|_{\partial D} = \tilde{S}\phi = \tilde{S}^{1/2}\ \left(\tilde{S}^{1/2}\phi\right) \in S^{1/2}(L^2(\partial D)).$$

Hence we have that $g_z \in \mathcal{G}S^{1/2}(L^2(\partial D))$. Theorem 7.1 now implies that also $g_z \in \Lambda_D^{1/2}(L^2(\partial H))$ and a solution $f \in L^2(\partial H)$ to $\Lambda_D^{1/2} f = g_z$ must exist.

For $z \in H \backslash \overline{D}$, we show that $g_z \notin \mathcal{R}(\mathcal{G}S^{1/2}(L^2(\partial H)))$ by using a contradiction. Suppose that there is a $\psi \in H^{1/2}(\partial D)$ such that $\mathcal{G}\psi = g_z$. Call u_δ the solution of the external Dirichlet problem (problem II of page 41) with boundary data ψ on ∂D. We have $u_\delta|_{\partial H} = 0$ and $G(\cdot, z)|_{\partial H} = 0$ by definition, and $g_z = \partial G(\cdot, z)/\partial n|_{\partial H} = \partial u_\delta/\partial n|_{\partial H}$ by assumption. This means that $G(\cdot, z)$ and u_δ have identical Cauchy data on ∂H, but $G(\cdot, z)$ contains a singularity in $H \backslash D$ and u_δ does not. Consider the space $H_{\overline{z}} : (H \backslash \overline{D}) \backslash B(z, \epsilon)$, in which $B(z, \epsilon)$ is a ball enclosing z of radius ϵ. In this space u_δ and $G(\cdot, z)$ are both harmonic and must coincide because the uniqueness of the Cauchy problem on ∂H. But u_δ and $G(\cdot, z)$ can not have the same Neumann boundary data on ∂B because B contains a singularity for $G(\cdot, z)$ and not for u_δ. We have therefore a contradiction, $g_z \neq \partial u_\delta/\partial n|_{\partial H}$ and there is no ψ for which $\mathcal{G}\psi = g_z$ Hence there is no $\phi = \tilde{S}^{-1/2}\psi$ and $g_z \notin \mathcal{R}(\mathcal{G}S^{1/2}(L^2(\partial H)))$.

□

The following corollary now follows immediately.

Corollary 7.3.

$$z \in D \iff \sum_n \frac{(g_z, f_n)^2}{\lambda_n} \approx ||\Lambda_D^{-1/2} g_z||^2 < \infty$$

7.2 Algorithm

The Factorisation Method provides an algorithm in which the object is reconstructed as points for which the function g_z is consistent with the measured data Λ_D, in which 'consistent' must be interpreted as being in the range of $\Lambda_D^{1/2}$. That is the object

is characterised by points for which $\Lambda_D^{1/2} f = g_z$ is solvable for some f or for which $||\Lambda_D^{-1/2} g_z||^2$ is finite. In practice Λ_D will be obtained experimentally from multiple measurements with limited precision of $c(x)$. If one $c(x)$ is obtained for each transmitter and if transmitters and receivers are co-located, each $c(x)$ forms a column of the matrix that represents Λ_D. On the other hand, if, as illustrated in the situation in figure 7.1a, the transmitter and receiver electrodes are not co-located or if more complex drive patterns are used for the transmitter electrode, an appropriate transformation needs to be performed to obtain an estimate of Λ_D. In the numerical studies below we have used co-located electrodes.

$||\Lambda_D^{-1/2} g_z||^2$ can only be calculated approximately and we introduce the function:

$$P(z) = \sum_n q^2(\lambda_n, \alpha) \frac{(g_z, f_n)^2}{\lambda_n} \approx ||\Lambda_D^{-1/2} g_z||^2. \tag{7.3}$$

in which $q(\lambda,\alpha)$ is a regularisation filter. Well known examples of regularisation filters are truncated SVD for which $q = 1$ for $\lambda \leq \alpha$ and $q = 0$ for $\lambda > \alpha$, or Tikhonov regularisation for which

$$q(\lambda_n, \alpha) = \frac{\sqrt{\lambda_n}}{\sqrt{\alpha} + \sqrt{\lambda_n}}. \tag{7.4}$$

Calculating $P(z)$ corresponds to solving a regularised version of $\mathcal{G}S^{1/2}\varphi_z = g_z$ because, remembering that $\mathcal{G}S^{1/2} : L^2(\partial D) \rightarrow L^2(\partial H)$ can be expressed by a singular system $(\sqrt{\lambda_n}, \varphi_n, f_n)$,

$$\varphi_z = \sum_n q(\lambda,\alpha) \frac{(g_z, f_n)\varphi_n}{\sqrt{\lambda_n}}. \tag{7.5}$$

Theorem 4.1 in Arens[138] shows that a regularisation filter for $\Lambda_D^{-1/2}$ does indeed generate a suitable regularisation filter for $\mathcal{G}S^{1/2}$. For an un-regularised system ($q(\lambda,\alpha) \equiv 1$), the function $||\varphi_z||^2$ blows up as z approaches ∂D from inside the domain and $z \notin D$, we have, $||\varphi_z||^2 \rightarrow \infty$. However, for a regularised system, $||\varphi_z||^2$ remains finite everywhere. Indeed, because of the regularisation filter, $P(z) = ||\varphi_z||^2$ is differentiable and can be interpreted as a Level Set from which the optimal object is found by equating it to the level for $P(z)$ that minimizes the cost function.

We note that this algorithm does not preserve the binary nature of theorem 7.2 on which it is based. A point either is, or is not, inside the object and the function g_z either lies, or does not lie, in the range of $\Lambda_D^{1/2}$. $P(z)$ on the other hand is finite

everywhere. It appears to measure distance to the object edge on a continuous scale and requires an a-posteriori identification of one level with the edge itself.

To preserve a binary nature, alternative algorithms have been explored in applications of the Factorisation Method. In his original paper on the Factorisation Method, Kirsch[116] used the equivalent of $P(z)$. As an alternative, he also used the value of the Tikhonov regularisation parameter α used to solve $\Lambda_D^{1/2} f = g_z$ subject to the Morozov discrepancy principle. A very large value of α indicates that the equation $\Lambda_D^{1/2} f = g_z$ is essentially insolvable and the test point must lie outside the object. Papers on Linear Sampling[189, 190], which test the equation $\Lambda_D f = g_z$, have also used this approach. In practice however α is a continuous parameter rather than a binary one and using $\alpha(z)$ as level set does not appear to have advantages over using $P(z)$.

In studies for EIT using Neumann to Dirichlet data in 2D[191, 119], the trend in the terms of the sum for $P(z)$ was used as criterion. A decreasing trend was used to indicate a convergent sum implying an interior point, while an increasing trend indicates an exterior point. In the presence of noise, the authors used an extrapolation from the terms for small n. In their case the low order terms appeared in odd-even pairs which, when averaged pair wise, provided a good prediction of the behaviour at larger n. In our case however, as shown in section 7.4 below, there is no such pairing. For limited aperture, noisy data there appears to be no heuristic strategy to recover or predict a trend in the terms of $P(z)$.

7.3 Numerical Results

The situation we are interested in concerns an experimental 90mm × 70mm display illustrated in figure 7.2. The aim of this system is to detect hand and finger movements in front of the display, but for the purpose of these numerical experiments we use a 'phantom' of a simple three pronged object located roughly 30mm above the display surface. Figure 7.2 also illustrates the cross section of the object in two planes. These cross sections of the 'original' object are to be compared with the reconstructions presented later. We are interested in two different configurations of sensor electrodes. In one it is assumed that a large number of co-located transmitter electrodes (voltage drive, or Dirichlet data) and receiver electrodes (capacitance measurements or Neumann data) are distributed over the full area of the display. Hence a relatively 'good'

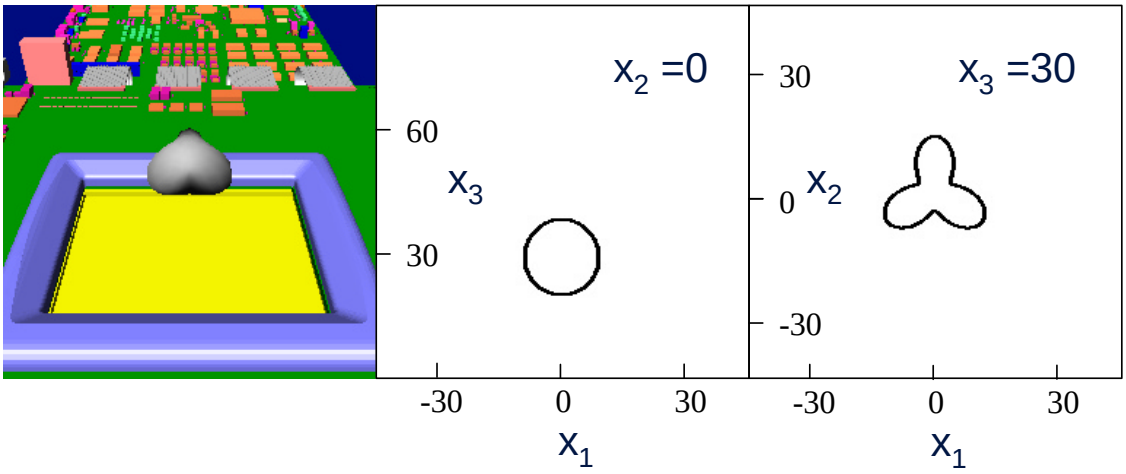

Fig. 7.2: Illustration of the test object used for the numerical simulations. The left image illustrates the object located 30mm above the sensor/display. The other two images give cross sections in the $x_2 = 0$ plane and the $x_3 = 30mm$ planes respectively.

measurement is made of the Dirichlet to Neumann map Λ_D. However the provision of a dense array of sensing electrodes over a display surface may not be feasible and it is important to also consider the situation in which electrodes are placed around the edge of the display. Hence in the second configuration a relatively small number of electrodes is located around the edge of the display and a 'poor' measurement is made of the Dirichlet to Neumann map Λ_D.

The numerical experiments were performed on a personal computer using C++ with Open Inventor support for the 3D visualisation. The Dirichlet to Neumann map Λ_D was calculated using factorisation (3.25). Piecewise constant quadrature was used to calculate the integral operators and the test objects were discretized to typically 420 triangles. Using a 4 or even 10 times denser discretization made no difference to the results presented below. NAG support using routine F01ABF was used to calculate $\tilde{S}^{-1}$ and F02WEF for the SVD. Pseudo random, normally distributed noise with standard deviation $\delta||\Lambda_D||$ was added to each element. Typically $\delta = 10^{-3}$ in the results below, this corresponded to 1% of the maximum value in Λ_D. The addition of noise introduces random variations in the elements of Λ_D which is no longer self adjoint. One way to deal with this would be to replace elements on opposite sides of the diagonal with average values but this makes certain assumptions about the noise model. Instead, in the calculation of (7.3), the left, (range) singular vectors were used corresponding to the

(approximate) calculation of $||(\Lambda_D^*\Lambda_D)^{-1/4}g_z||$. On a computer system with a 1.5GHz Pentium Processor with 500 MB RAM, the solution of the forward problem for the case of electrodes around the edge of the display, takes 3s, the spectral decomposition 0.3s, while the calculation of $P(z)$ typically takes 15ns per sampling point z.

The results are illustrated in figure 7.3. In 7.3a, 'good' measurements are made on a 16×16 array of 256 transmitter electrodes and the same number of co-located receiver electrodes. The Dirichlet to Neumann map Λ_D is schematically illustrated in the second image on each row in figure 7.3. Each column represents the capacitance (Neumann) data for a single given voltage (Dirichlet) drive. The map is clearly symmetric. The third and fourth images in figure 7.3a illustrate a map of $P(z)$ in (7.3) in the z_1, z_3 plane through the centre of the object ($z_2 = 0$) and in the z_1, z_2 plane ($z_3 = 30mm$). The greyscale is a logarithmic representation of the value of $P(z)$. No noise was added to the measurements and no regularisation was applied other than the regularisation that is inherent in the finite precision of the software, the discretization and the finite aperture of the measurements. As can be seen, the object is clearly represented by low values for $P(z)$ and the norm increases strongly for test points away from the object. The white outline in the cross section represents the level of $P(z)$ that minimizes the cost function and represents our best estimate of the reconstructed object. It compares reasonably well with the outline of the original shown in figure 7.2.

Unfortunately, measurements over the whole 2D surface are not feasible in practice as the necessary electronics takes up space and will cause unavoidable impairment of the display function itself. Restricting the electrodes to the edge of the display is more realistic. Figure 7.3b shows results when 'poor' measurements are made with only with 32 transmitter electrodes around the edge of the display and again on the same number of co-located receiver electrodes. Noise is now also added to the Dirichlet to Neumann map Λ_D which is now just a 32x32 array. The results in the map of $P(z)$ in figure 7.3b show that the noise distorts the reconstructed shape of the object. Also, the limited aperture makes good reconstruction in the z_3 direction difficult. Nevertheless, it is obvious *where* the object is and what its orientation is. If the object is moved and rotated above the display, the maps continue to reproduce location and orientation correctly.

The result in figure 7.3b shows that object reconstruction from capacitance measurements around the edge of the display area can work. The phantom in these numerical

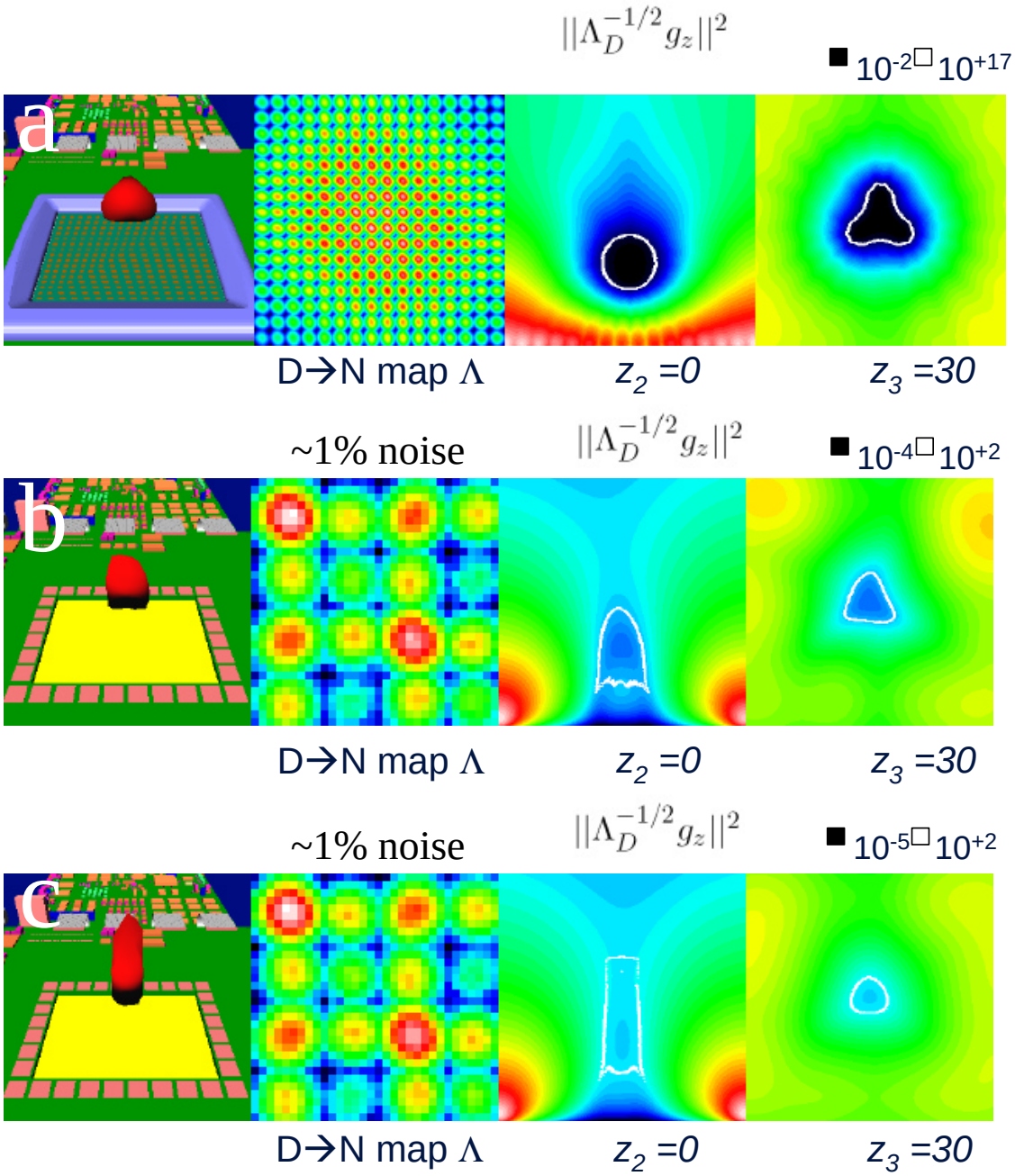

Fig. 7.3: Reconstruction results of the test object above a 90mm × 70mm test display using the Factorisation Method. a) 256 sampling points in a 16 × 16 array, b) 32 measurement points in a linear edge array, c) 32 measurement points using truncated SVD regularisation (half the singular values).

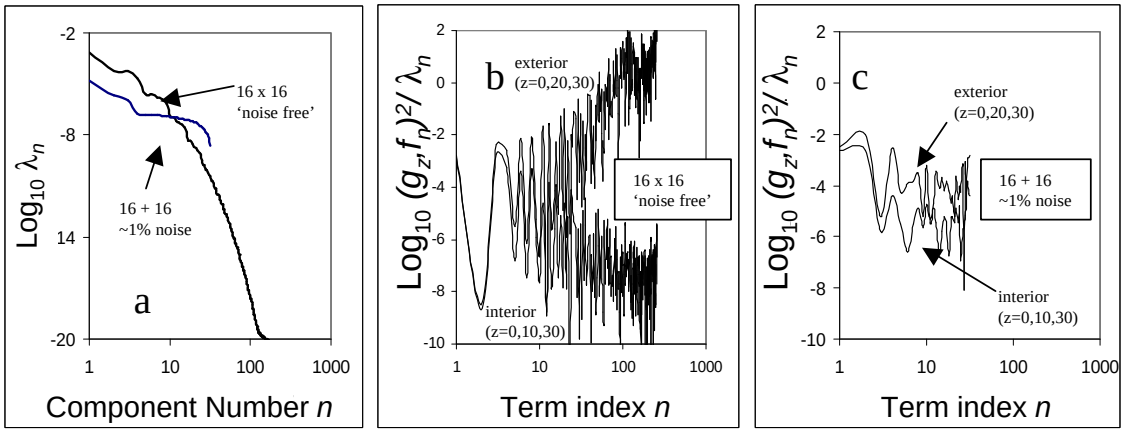

Fig. 7.4: a) Singular values of Λ_D for the two simulated experiments. Terms of the $P(z)$ sum for typical internal and external points. b) for the 'noise free' 16×16 2D electrode array and c) for 32 edge electrodes with 1% noise.

studies was chosen to illustrate both what can, and cannot be reconstructed. Clearly the object shape is at the edge of detectability at these noise levels but the fact that orientation and location are reconstructed successfully, opens up many interaction possibilities. For instance, the result suggests that the rotation by three fingers of a virtual knob in front of the display could be detected. Large gestures, that is larger objects, or multiple objects representing well spaced fingers or separate hands are much easier to reconstruct and are tolerant to higher noise levels.

The effect of the noise, which skews the maps of $P(z)$ in random directions, can be reduced through explicit regularisation. Figure 7.3c shows an example of fixed truncated SVD in which a fixed number of N terms are used in (7.3), $N = 16$. In this example, the regularisation obscures most shape information and deteriorates the vertical positioning of the object. A-posteriori criteria can be used to select a better SVD truncation in some cases, we return to this point later.

7.4 Discussion

In this section we investigate the numerical experiments further. First we look at the terms of the sum (7.3) for $P(z)$. Figure 7.4a shows the singular values that appear as the denominator in the terms. Shown are the eigenvalues for the noise free 16×16 electrode

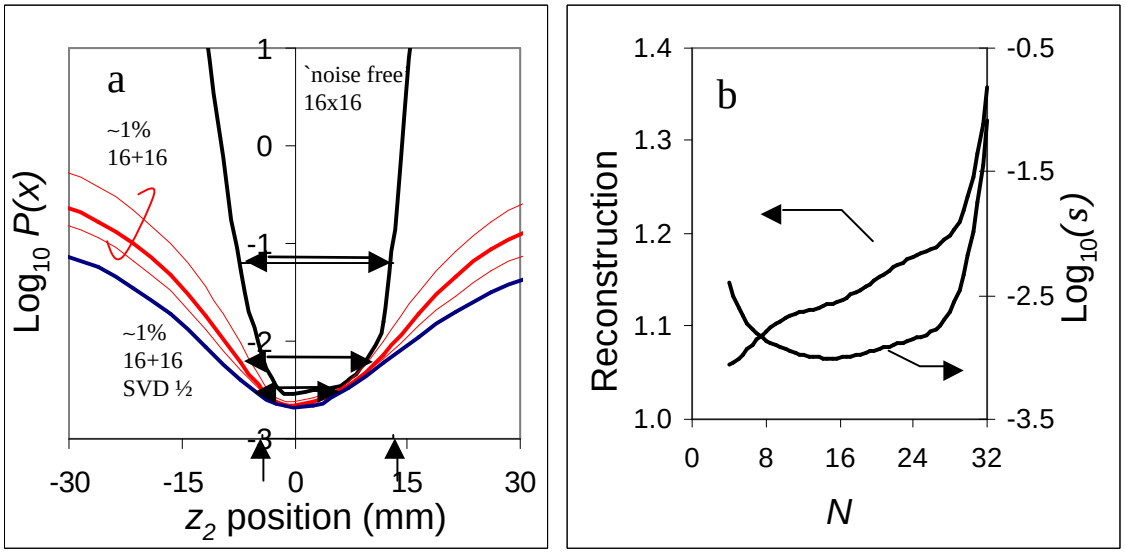

Fig. 7.5: a) Cross section of $P(z)$ for $z_1 = 0$, $z_3 = 30$ for the three experiments in figure 7.3, b) Quality of the reconstruction as measured by the asymmetry (left hand ordinate) and standard deviation (right hand ordinate) under different noise draws as function of SVD truncation N.

array reconstruction, that is for the Dirichlet to Neumann map shown in figure 7.3a. Also shown are the singular values for the noisy 32 edge electrode arrangement of figure 7.4b. In both cases the values decay sharply, although the decay is fairly featureless, the level and shape varies with electrode layout and object shape. Clearly the decay levels off sooner for the low resolution, noisy data, than for the higher resolution 'noise free' data. For test points inside the object, the (squared) numerators of the terms in (7.3) have to beat this decay in order to provide a convergent sum for $P(z)$. Figure 7.4b shows that indeed for large n, there is a trend for the sum terms to decease for interior points and increase for exterior points. However this trend itself is noisy and cannot be predicted easily from the terms at small n. The evidence for the trend disappears with the addition of noise as illustrated in figure 7.4c. There appears to be little structure in the terms although it is clear that the terms for the interior point remain, on the whole, below those of the exterior point. This supports the use of $P(z)$ as a good way to differentiate between interior and exterior points because it represents a (weighed) sum of the terms.

Figure 7.5a explores $P(z)$ further by plotting a cross section along the z_2 axis through the centre of the test object. The vertical arrows on the z_2 axis indicate the

edges of the original test object. For the 'noise free' case of figure 7.3a, $P(z)$ is fairly level inside the object and increases towards the object edge. As noted from equation (7.5), $P(z)$ provides a measure of distance to the object edge. However, although $P(z)$ increases sharply outside the object, it remains finite and smooth and there is no a-priori value for $P(z)$ that identifies the edge ∂D. Minimization of the cost function provides an acceptable a-posteriori value indicated by the white outline in figure 7.3a and here by the horizontal double arrow in figure 7.5. As can be seen, the double arrow indicating the reconstructed object corresponds fairly well to the vertical arrows of the original object. In particular the asymmetry of the object along the z_2 axis is reproduced.

It must be pointed out at this stage that the shape of the 'noise free' plot does not appear to be limited by the discretization or finite aperture of the sensor array. For instance, increasing both to simulate $64 \times 64 = 4096$ co-located sensors on a 180mm $\times$ 140mm area, provides virtually identical results. As discussed with reference to figure 4.3, increasing the triangulation density of the object from 420 to 1820 does not changes the eigenvalues of the forward map substantially. We suspect that the main factor determining the shape of the plot is in fact the 15 digit floating point arithmetic. The finite precision could be regareded as noise and hence our use of quotes in the label 'noise free'.

Adding deliberate noise to the simulated measurements changes the shape of $P(z)$ substantially. The group of curves labelled "$\sim$ 1%, 16+16" shown in figure 7.5, illustrate cross sections for this case. The thicker central curve represents the result from figure 7.3b. The thin outer curves represent plus and minus two standard deviations from an average established over many noise draws. Even in the presence of noise, $P(z)$ remains smooth. The reconstructed object indicated by the double arrows matches the original object quite well, although less so than the 'noise free' reconstruction.

Truncating the sum for $P(z)$ reduces the effect of noise as illustrated by the curve "$\sim$ 1%, 16+16 SVD 1/2", in which only the first 16 terms are used. The variability of this curve between noise draws is negligible, but the curve is also much more symmetric. The double arrows are near symmetric recalling the near circular shape of the white outline in figure 7.3c.

In fact we can use the asymmetry in the cross section as an indication of the quality of the reconstruction and correlate this with the sensitivity of $P(z)$ to noise (the

standard deviation) and the level of explicit regularisation. This is done in figure 7.5b. The curve plotted with respect to the left hand axis is a measure of the asymmetry (the ratio $\overline{P}(z_2 = 15)/\overline{P}(z_2 = -15)$) as function of the number of terms used in the sum for $P(z)$. The curve plotted with respect to the right hand axis is the standard deviation in $P(z_2 = \pm 15)$. These curves illustrate the conventional trade-off between detail in the reconstruction and sensitivity to noise that is common in regularisation of inverse problems. As regularisation is increased (N is decreased from 32), sensitivity to noise decreases with some decrease in the quality of the reconstruction. Too much regularisation causes further loss in quality of reconstruction without much improvement in stability. This behaviour is typically discussed in terms of an L-curve and the results in figure 7.5b could have been presented in that way. The results suggest an optimum regularisation at $N = 26$, and reconstructions with this parameter do work well. However, the asymmetry criterion is based on prior knowledge of the original object and would not be available for arbitrary reconstructions. Figure 7.5b is presented here to illustrate the regularisation behaviour of $P(z)$. In practice other criteria and prior knowledge may be available to measure the quality of the object reconstruction to aid on-the-fly optimization of the regularisation parameter and this the subject of further research.

7.5 Linear Sampling

In the discussion above we have investigated how well the Factorisation Method performs in the reconstruction of the shape of a test object. It is interesting to explore how Linear Sampling performs in comparison. We recall the discussion in the introduction that in a finite, regularised setting, both test against the same range (the space spanned by the singular vectors of Λ_D), but may do so with a different test. The Linear Sampling level set function is

$$||\Lambda_D^{-1} g_z||^2 \approx P_{LS}(z) = \sum_n q^2(\lambda, \alpha) \frac{(g_z, f_n)^2}{\lambda_n^2} \tag{7.6}$$

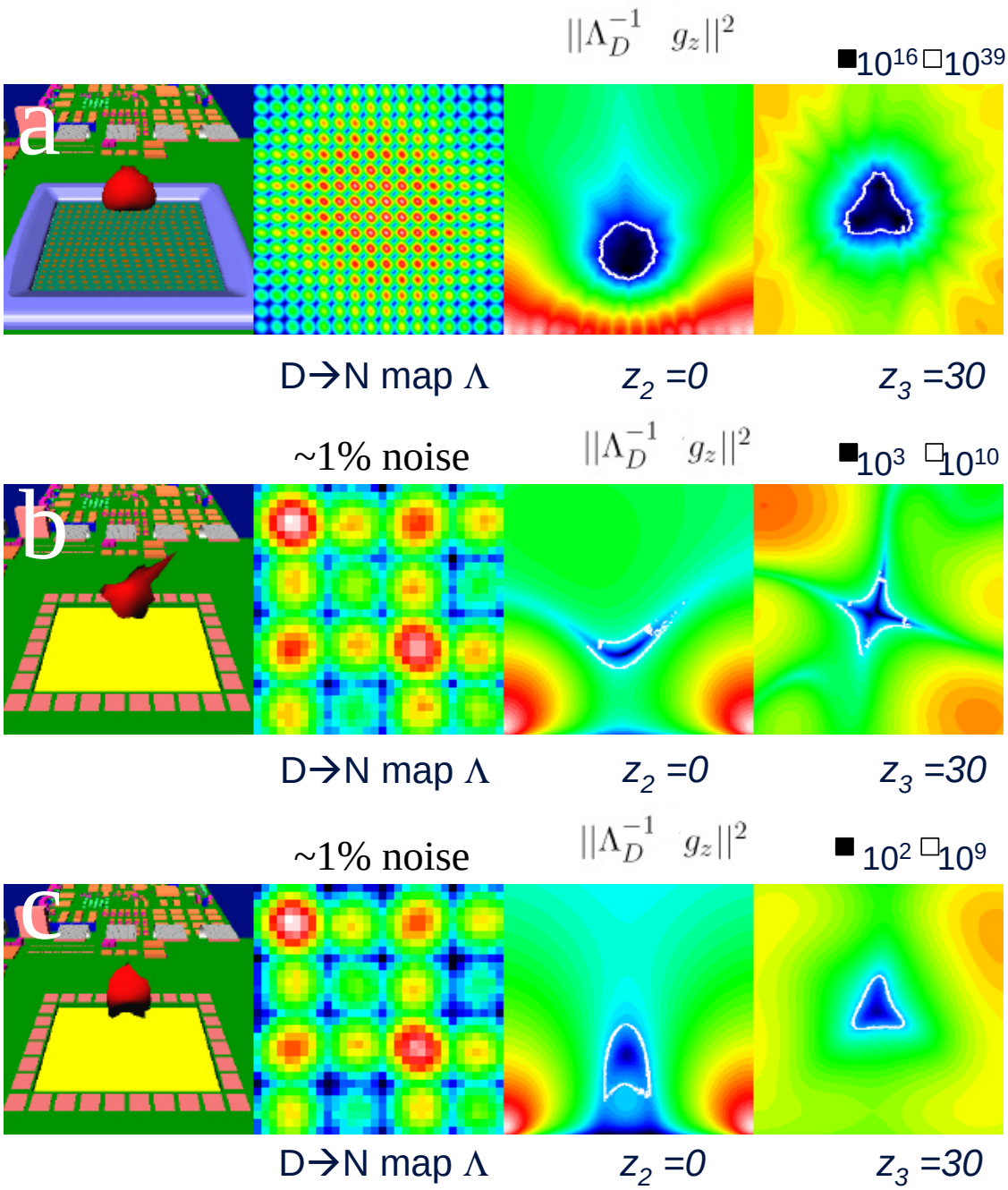

Fig. 7.6: Reconstruction results of the test object above a 90mm × 70mm test display using the Linear Sampling Method. a) 256 sampling points in a 16 × 16 array, b) 32 measurement points in a linear edge array, c) 32 measurement points using truncated SVD regularisation (half the singular values).

This can be interpreted as a differently regularised version of (7.3) in which the regularisation filter is replaced by

$$q(\lambda, \alpha) \rightarrow \left(\frac{\alpha}{\lambda}\right)^{\beta} q(\lambda, \alpha)$$

in which β is another regularisation parameter, $\beta = 1/2$ in this case.

Figure 7.6 gives the results for electrode configurations and noise levels that were used in the factorisation results in figure 7.3. As can be seen, the noise reconstruction for the 2D sensor array with 'noise free' measurements (7.6a) is very similar to the corresponding result in figure 7.3a. For the edge sensor layout with $\sim 1\%$ noise, the unregularised result in figure 7.6b is not a good reconstruction, but the truncated SVD (half the components) provides a good, and stable reconstruction. It appears that, from a practical point of view, Linear Sampling and the Factorisation Method can both be used to successfully reconstruct the object. Indeed, one can be regarded as a further regularisation of the other. Though formally substantially different, in practice the choice between them is likely to be a matter of experimental detail and personal preference.

7.6 Conclusions

In this chapter we have shown that it is possible to detect the position, orientation and shape of hands and fingers from capacitance measurements made between electrodes embedded in or around a display surface. Formally the problem is solved with the Factorisation Method applied to this case with a suitable test function g_z appropriate to the boundary conditions present here. The Factorisation Method provides an algorithm for the regularised calculation of a level set function $P(z)$ which is finite and differentiable everywhere and permits the use of minimization of the cost function for the final object reconstruction. The method works well even with limited and noisy data. Moreover the application of explicit regularisation can be used to trade-off noise and shape resolution and fine-tune the reconstruction in practice.

8. MONOTONICITY

Theorem 3.11 states that for two objects D_i and D_e, such that $D_i \subset D_e$, the operator $\Lambda_{D_e} - \Lambda_{D_i}$ is positive definite. This establishes a monotonicity principle for cross capacitance sensing. In EIT monotonicity has been the basis of a reconstruction method[192] which shares a number of interesting features with the sampling methods of the previous chapter, but may also have some of the advantages of the iterative method in chapter 6. In this chapter a monotonicity method for cross capacitance sensing is explored. First an introduction to the method in EIT is presented and then applied to cross capacitance sensing in a number of numerical examples. Finally we explore the relation with the Factorisation Method of the previous chapter.

8.1 Monotonicity in EIT

Tamburrino's monotonicity method[192] for EIT or ERT (Electrical Resistance Tomography) is based on the monotonicity of the resistivity matrix R.

$$\rho_1(r) \geq \rho_2(r) \text{ for } r \in \Omega \Rightarrow R_{\rho_1} \geq R_{\rho_2} \tag{8.1}$$

Take $\tilde{R}$ to denote the measured resistivity matrix corresponding to the unknown object D. R_k denotes the calculated resistivity matrix corresponding to a (small) probe domain D_k. We assume, for simplicity of exposition, that the resistivity inside objects (D or D_k) is larger than the background resistivity. Now,

$$D_k \subset D \Rightarrow \rho_D(r) \geq \rho_k(r), \quad r \in \Omega \tag{8.2}$$

In which $\rho_k(r)$ is the entire resistivity distribution over the domain Ω in the presence of object D_k and ρ_D the resistivity distribution over the whole domain in the presence of the original object D. Hence $\tilde{R} - R_k$ will be positive semi definite. Unfortunately

if the probe domain is outside the real object ($D_k \subset \Omega \backslash \overline{D}$), then ρ_k will in places be larger and in places be smaller than ρ_D and no statement can me made about whether $\tilde{R} - R_k$ is positive semi definite or not.

Thus by taking different probe domains at different positions, calculating R_k, and collecting probe domains for which $\tilde{R} - R_k$ is semi positive definite, a first estimate of the object can be made. Note that, the method also works for the case when the object is more conductive than the medium in which it sits. In that case the positive definite test collects points where $\tilde{R} - R_k$ is not positive semi definite.

Aykroyd *et al*[193] find the eigenvalues $\lambda_{k,j}$ of $\tilde{R} - R_k$ for each D_k and construct a sign index s_k

$$s_k = \frac{\sum_j \lambda_{k,j}}{\sum_j |\lambda_{k,j}|} \tag{8.3}$$

with this the first estimate of the object is defined as

$$\begin{aligned} D_{Ext} &:= \{D_k : s_k = 1\} \\ &= \{z \in \Omega : z \in D_k \text{ for some } D_k \subset \Omega \text{ for which } s_k = 1\} \end{aligned} \tag{8.4}$$

There is no guarantee that D_{Ext} will not contain sub domains outside the object, only that D_{Ext} will contain all of D. Because of this, we have $\rho_{D_{Ext}}(r) \geq \rho_D(r)$, that is $R_{D_{Ext}} - \tilde{R}$ is positive semi definite. If now a small probe domain D_k is removed from D_{Ext}, then $R_{D_{Ext}\backslash D_k} - \tilde{R}$ is guaranteed positive semi definite only if $D_k \subset D_{Ext}\backslash D$. Aykroyd *et al*[193] find the eigenvalues $\sigma_{k,j}$ of $R_{D_{Ext}\backslash D_k} - \tilde{R}$ and construct a second sign index t_k

$$t_k = \frac{\sum_j \sigma_{k,j}}{\sum_j |\sigma_{k,j}|} \tag{8.5}$$

with this, a second estimate of the object is defined as

$$D_{Int} := \{D_k : s_k = 1,\ t_k \neq 1\}$$

The probe domains $D_{Ext} - D_{Int}$ remain uncertain and Aykroyd *et al* treat this in a Bayesian frame work.

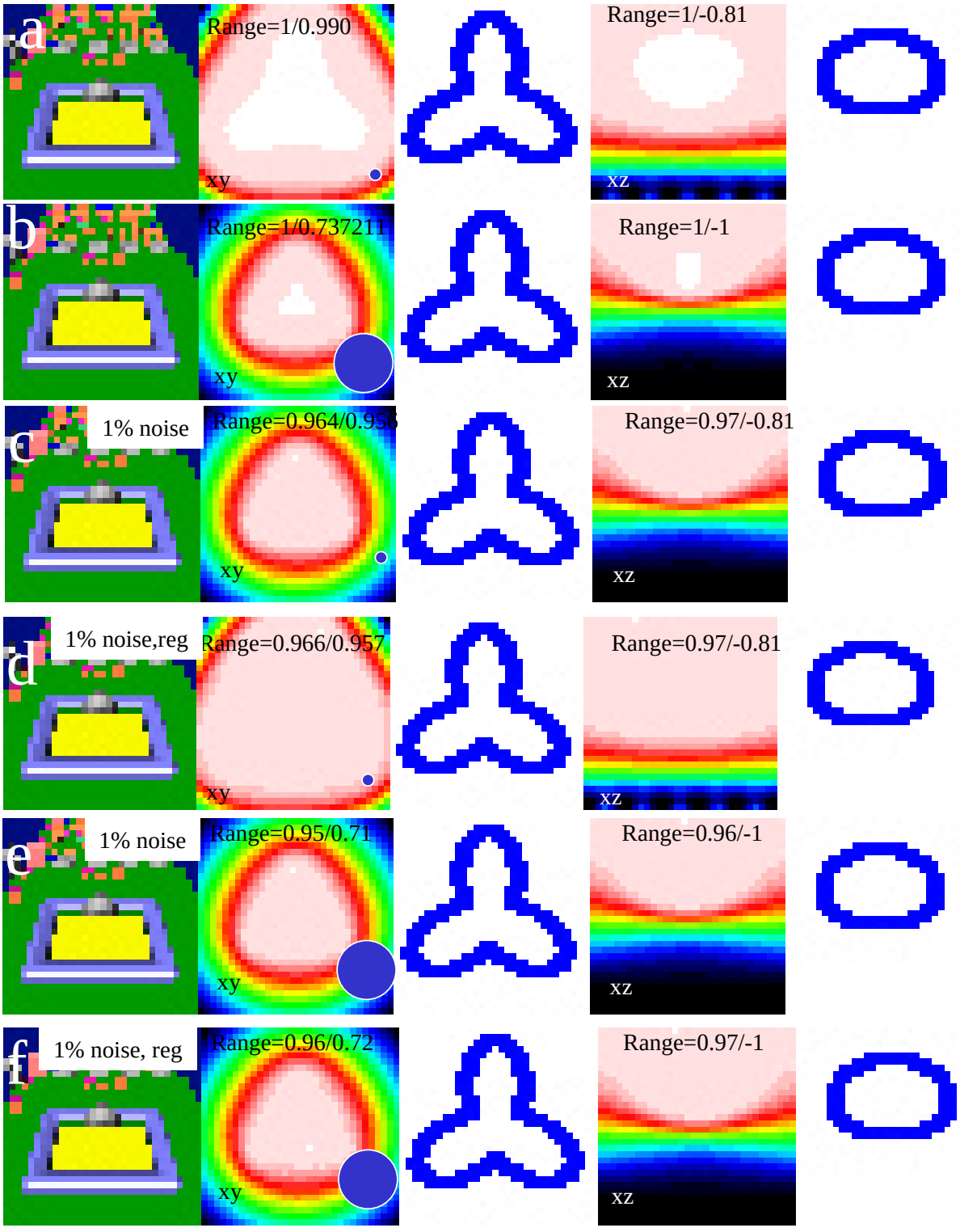

Fig. 8.1: Monotonicity maps

8.2 Monotonicity for CCS

We are interested in applying the monotonicity method to cross capacitance sensing. Because a boundary element method is used to solve the forward problem, rather than a probe domain, a probe object is used. A spherical object with certain radius, that is placed at different test positions. Figure 8.1 illustrates maps of the sign index s_k in much the same way as we have presented plots of $||\Lambda_D^{-1/2} g_z||^2$ in chapter 7. Cross sections in both a horizontal and vertical plane are shown along side cross sections of the original object for comparison. In the first row (a) the radius of the probe object is 1mm, which is to be compared with the 30mm $\times$ 30mm plane of the x,y cross section. A representative disk is illustrated in the bottom right of the xy cross section in the first row. In the plots, white corresponds to the maximum value of $s_k = 1$, while black corresponds to the minimum value occurring in that plot. This is $s_k = 0.990402$ for the xy cross section and $s_k = -0.815311$ for the xz cross section. Clearly the reconstruction of the object on the s_k test alone would depend very critically on the choice of the threshold $(1 - \epsilon)$ below which values of s_k would be regarded as not representing positive semi definite matrices. Following Aykroyd *et al* and the method in chapter 7, it is possible to find the optimum value for ϵ by minimizing $||\Lambda_{D_{Ext}} - \Lambda_D||^2$. This has not yet been implemented in the software, but doing this by inspection from figure 8.1 suggests that this would work. The major issue is the fact that the calculation of the maps is very time consuming, about 1000 times slower than the linear sampling, although it must be noted that in a practical application, many of the forward maps could be pre-calculated or evaluated in paralell.

The second row (b) of figure 8.1 tests the effect of increasing the size of the test object. As expected, the 'contrast ratio' of s_k increases, that is the minimum value of s_k decreases. However, the object reconstruction is clearly blurred in comparison with the result of the first row of figure 8.1.

The remaining rows in figure 8.1 explore the effect of adding 1% noise to the input forward map and the effectiveness of using regularisation by using only the first 100 eigenvalues for the calculation of s_k instead of the full 256 eigenvalues.

8.3 Monotonicity and Factorisation

To explore the method further, consider a small spherical probe object D_δ at z of small radius δ.

We first recall the standard integral mean value theorem (e.g Haggarty[194] 7.1.12)

Theorem 8.1. *Let f and g be contineous on $[a,b]$ with $g(x) > 0$ for $x \in [a,b]$. Then there exisits a c between a and b with*

$$\int_a^b f(x)g(x)\,\mathrm{d}x = f(c)\int_a^b g(x)\,\mathrm{d}x$$

This also holds if $g(x)$ is weakly singular because what is required in the proof is that $\int g$ exisits. We will also use the mean value theorem for harmonic functions.

The operator expression for S on a small spherical object is.

$$(Sg_\delta)(x) = \int_{\partial D_\delta} G(x,y)g_D(y)\,\mathrm{d}s(y), \qquad x \in \partial D$$

Using the integral mean value theorem (or the constant charge approximation in chapter 2). That is taking g_δ out of the integral, we get

$$= g_\delta \int_{\partial D_\delta} G(x,y)(y)\,\mathrm{d}s(y) \qquad x \in \partial D$$

Using proposition C.2, and we then get

$$= g_\delta\delta - g_\delta\frac{\delta^2}{2z}$$

and dropping the image charge term for small δ.

$$\approx \delta g_\delta$$

Note that we have made an approximation only in the last step and that the accuracy to that approximation can be made arbitraily small by choosing δ small.

Similarly for the operator T

$$\begin{aligned}(Tg_\delta)(x) &= \int_{\partial D_\delta} \frac{y_3}{2\pi|x-y|^3} g_D(y)\,\mathrm{d}s(y) \quad x \in \partial H \\ &= g'_\delta \int_{\partial D_\delta} \frac{y_3}{2\pi|x-y|^3}\,\mathrm{d}s(y) \quad x \in \partial H\end{aligned}$$

Where again integral mean value theorem was used. Note that the values of g_δ and g'_δ are not guaranteed to be the same. Recalling the definition (7.2) of the test function from chapter 7

$$g_z(x) = \frac{\partial G(x,z)}{\partial x_3} = \frac{z_3}{2\pi|x-z|^3}, \quad x \in \partial H$$

we are reminded that the kernel of T is harmonic and with the mean value theorem for harmonic functions

$$= 4\pi\delta^2 g'_\delta g_z(x)$$

Finally we have, simply by definition

$$\begin{aligned}(T^* f_H)(x) &= \int_{\partial H} \frac{x_3}{2\pi|x-y|^3} f_H(y)\,\mathrm{d}s(y) \quad x \in \partial D \\ &= (g_x, f_H)\end{aligned} \tag{8.6}$$

We now have (cf equation (3.25))

$$\Lambda_\delta f_H = T\tilde{S}^{-1}T^* f_H \approx 4\pi\delta \frac{g'_\delta}{g_\delta}\,(g_z, f_H) g_z \tag{8.7}$$

We recall that Λ_D generates a complete orthonormal basis of eigenfunction f_n and we can write $g_z = \sum_n (g_z, f_n) f_n$. Now consider

$$(\Lambda_D - \Lambda_\delta)\, f_i = \lambda_i f_i - \sum_n 4\pi\delta \frac{g'_\delta}{g_\delta}\,(g_z, f_n)(g_z, f_i) f_n \tag{8.8}$$

and

$$(f_i, (\Lambda_D - \Lambda_\delta) f_i) = \lambda_i - 4\pi\delta \frac{g'_\delta}{g_\delta}\,(g_z, f_i)^2 \tag{8.9}$$

From this we conclude that for $(\Lambda_D - \Lambda_\delta)$ to be positive definite, we have the requirement

$$\frac{(g_z, f_i)^2}{\lambda_i} < \frac{1}{4\pi\delta\frac{g'_\delta}{g_\delta}} \tag{8.10}$$

for all i.

Therefore, the Monotonicity method amounts to checking that a sequence is bounded, whereas the Factorisation Method seeks to check that the sequence is convergent. As we have already seen, in the finite dimensional space imposed by the sensor configuration, there is little fundamental difference between these two requirements. Whether we must now regard the Monotonicity method as an awkward disguise for the Factorisation method, or alternatively, because Monotonicity allows a variety of probe object shapes, the Factorisation method as merely a limiting case for the richer Monotonicity method, will be a interesting topic for future research.

9. CONCLUSIONS

This thesis has explored the theory and algorithms of recovering object location and shape from the electrostatic measurement of Cauchy data on a plane.

The forward problem was solved using Green's theorem and a Green's function appropriate for the half space. This led to an operator factorisation of the forward Dirichlet to Neumann map, $\Lambda_D : L^2(\partial H) \to L^2(\partial H)$ which has been demonstrated to be compact and injective. It is also uniquely dependent on the object and a monotonicity result applies.

The inverse problem of recovering the object from the map is ill-posed, not in the sense the solution may not exist, but in the sense that the solution is unstable, and requires regularisation.

The inverse problem can be solved in different ways depending on the amount of prior knowledge available and the level of detail required. This thesis has explored finding an initial guess, shape optimization, iterative reconstruction through analytic continuation, linear sampling and finally a monotonicity method.

These have different characteristics, advantages and drawbacks. The initial guess is fast and robust. The regularisation is implicit through the low number of degrees of freedom of the initial guess. The trade-off between regularisation and stability is illustrated by the fact that as more degrees of freedom are introduced in the shape optimization to include orientation and some shape information, the noise tolerance of the inverse method decreases.

Ultimately however, the shape optimization has limited scope because the trade-off becomes too unfavourable when meaningful shape reconstruction is required. It is necessary to make better use of the knowledge of the solution of the forward problem to design inverse algorithms.

The iterative reconstruction method is precise with well characterized regularisation. The method also works well when only relatively limited data on the forward

map is available, for instance the charge image corresponding to one transmitter only. However, the method is slow and restricted to predetermined topologies. That is prior information is required on whether there are one or more objects and whether these objects are star shaped, convex, doughnut or have other topologies.

The factorisation method has stronger demands on the data of the forward map than the iterative reconstruction method. That is, it requires multiple charge images from multiple transmitters to establish an estimate for the range of the forward map. However, if that is available, the method is fast and independent of topology. It is the preferred method provided that suitable data are available.

The monotonicity method provides an interesting link between a method developed for EIT and the factorisation method, but is, at this stage, not yet a practical reconstruction method. It may be possible however in future to combine aspects of the iterative reconstruction with the factorisation method through the monotonicity method.

The results in this thesis have used computer-simulated experimental data because real experimental data have not yet been available. However, once these data become available, the theory and numerical experiments in this thesis indicate that a demonstration of shape reconstruction from capacitance measurements made between electrodes embedded in or around a display surface will be be possible to create unobtrusive 3D gesture input for interactive displays, for what has been called “touchless interaction”.

BIBLIOGRAPHY

[1] P A Gough. Wearable technology. pp 158-163, New Everyday, 010 Publisher, ISBN 90-6450-502-0, 2003.

[2] Gerhard von der Emde and Horst Bleckmann. Finding food: Senses involved in foraging for insect larvae in the electric fish gnathonemus petersii. The Journal of Experimental Biology 201, 969-980, 1998.

[3] Gerhard von der Emde. Active electrolocation of objects in weakly electric fish. J Exp Bio 202 (10) 1205-1215, 1999.

[4] Neil Gershenfeld and J R Smith. Method for three-dimensional position, orientation and mass distribution. US pat 5844415, 1998.

[5] John David Jackson. *Classical Electrodynamics*. John Wiley Sons, New York, 2nd edition, 1975.

[6] Kasper van Wijk. *Answers to a selection of problems from Jackson's Classical Electrodynamics*. Samizdat Press, Golden, Colorado, http://samizdat.mines.edu, 1996.

[7] Richard A. Bolt. Put-that-there voice and gesture at the graphics interface. Proceedings of the 7th annual conference on Computer graphics and interactive techniques, p.262-270, July 14-18, 1980 Seattle, Washington, United States, 1980.

[8] Ron Stephen Schulz. Touch screen applications. Electronic Engineering April 1999 pp41-44, 1999.

[9] Jun Rekimoto. Smartskin: An infrastructure for freehand manipulation on interactive surfaces. CHI2002, 1991.

[10] Jefferson Y. Han. Low-cost multi-touch sensing through frustrated total internal reflection. Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology pp115-118, 2005.

[11] Edward Tse, Saul Greenberg, Chia Shen, and Clifton Forlines. Multimodal multiplayer tabletop gaming. Third International Workshop on Pervasive Gaming Applications - PerGames 2006, May 07th 2006, Dublin, Ireland, 2006.

[12] Mike Wu and Ravin Balakrishnan. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. ACM CHI Letters vol5 pp192-201, 2003.

[13] M Goldstein, O Baez, and Peter Danielsson. Employing electrical field sensing for detecting static thumb position using the finger-joint gesture keypad input pardigm. The fourth international conference on wearable computers, 2000.

[14] John M Vranish. Three dimensional interactive display. US Patent 2002/0000977, priority 13/3/2001, 2002.

[15] Marco Tartagni, Bhusan Gupta, and Alan Kramer. Capacitive distance sensor. US patent US66114862, 2000.

[16] David Manners. Fingerprint recognition too costly for use in mobile phone security. Electronics Weekly no 2185, 9 March, 2005.

[17] Shahzad Malik and Joe Laszlo. Visual touchpad: A two-handed gestural input device. Proceedings of the ACM International Conference on Multimodal Interfaces pp289-296, 2004.

[18] Shahzad Malik and Joe Laszlo. Interacting with large displays from a distance with vision-tracked multi-finger gestural input. Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology pp43-5, 2005.

[19] Jakub Segen and Senthill Kumar. Gesturevr: Vision-based 3d hand interface for spatial interaction. Proceedings of the Sixth ACM International Conference on Multimedia pp455-464, 1998.

[20] M Gleicher and N Ferrier. Evaluating video-based motion capture. IEEE Proc Computer Animation pp75-80, 2002.

[21] Daniel Vogel and Ravin Balakrishnan. Distant freehand pointing and clicking on very large, high resolution displays. Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology pp33-42, 2005.

[22] Andrew D Wilson. Touchlight: An imaging touch screen and display for gesture-based interaction. ICMI04, October 13-15, 2004, State College, Pennsylvania, USA, 2006.

[23] Hal Philipp. Charge transfer sensing. Sensor-Review (UK), vol.19 pp96-105, 1999.

[24] N D Young, G Harkin, R M Bunn, D J McCulloch, R W Wilks, and A G Knapp. Novel fingerpring scanning arrays using polysilicon tfts on glass and polymer substrates. IEEE Electron Devices vol 14 pp19-20, 1997.

[25] Thomas William Bach. Capacative proximity or profile detector. UK Patent application GB231514, 1997.

[26] J G Stanley and G W White. Electric field sensing airbag danger zone sensor. US Patent 5964478, 1999.

[27] Stephan Ohr. Non-contact sensor discerns passenger size for airbags. EE Times 6 Feb 2003, available at http://www.eetimes.com/story/OEG20030206S0035, 2003.

[28] T. Bach, C.R. Chatwin, and R.C.D. Young. Capacitive sensor arrays - digitising objects in three dimensions. Proceedings Proc. Int. Conf. in Sensors and Transducers (MTEC 1999), NEC Birmingham UK, 17th-18th February, 1999.

[29] Thomas G Zimmerman, Joshua R Smith, Joseph A Paradiso, David Allport, and Neil Gershenfeld. Applying electric field sensing to human-computer interfaces. CHI95, Human Factors in Computing Systems, ACM Press pp280-287, 1995.

[30] W L Watton and M E Pemberton. A direct-capacitance aircraft altimeter. Proc IEE vol 96 Part III pp203-213, 1949.

[31] T S McLeod and C M Deavin. Developement of capacitance altimeter. Proc IEE vol 110 (9) pp1505-1513, 1963.

[32] Richard D Brown. Virtual unreality and dynamic form: an exploration of space, time and energy. Leonardo vol 33 pp21-25, 2000.

[33] Programmable capacitance-to-digital converter with environmental compensation. Analog Devices Data Sheet AD7142 Rev 0 available at www.analog.com, 2006.

[34] Cees van Berkel. Electrostatics of cross capacitance sensing. Technical Note 4199, Philips Research Labs, Redhill, 2002.

[35] Maria A Stuchly and Trevor W Dawson. Interaction of low-frequency electric and magnetic fields with the human body. Proc IEEE vol 88 pp643-, 2000.

[36] Michael McGuffin and Ravin Balakrishnan. Acquisition of expanding targets. Proceedings of ACM CHI 2002 Conference on Human Factors in Computing Systems, 2002.

[37] Cees van Berkel. Image preparation for 3d-lcd. Proc SPIE vol 3639 pp84-91, 1999.

[38] C C Barber and BH Brown. Imaging spatial distrbutions of resistivity using applied potential tomography. Electron Lett vol 19 (22) pp943-935, 1983.

[39] Liliana Borcea. Electrical impedance tomography. Inverse Problems vol 18 R99-R136, 2002.

[40] William Lionheart, Nicholas Polydorides, and Andrea Borsic. The reconstruction problem. in Holder D.S., Electrical Impedance Tomography: Methods, History and Applications, Institute of Physics, 2004.

[41] J G Webster and P Hua. Recent north american progress in electrical impedance tomography. Recent North American progress in electrical impedance tomography" J G Webster and P Hua [Madison] Clinical Physics and Physiological Measurement vol 9 pp127-130, 1988.

[42] C J Kotre. Electrical impedance tomography. British J of Radiology vol 70 pp S200-S205, 1997.

[43] D S Holder. *Electrical Impedance Tomography: Methods, History and Applications*. Institute of Physics, London, 2004.

[44] N D Harris, A J Suggett, D C Barber, and H B Brown 1987 Clin Phys Physiol Meas A vol 8 pp155 165 (1987). Applications of applied potential tomography (ap) in respiratory medicine. Electron Lett vol 19 (22) pp943-935, 1983.

[45] M R Akbarzadeh, W J Tompkins, and J G Webster. Multichannel impedance pneumography for apnea monitoring. Proc Annu Int Conf IEEE Eng Med Biol Soc vol 12 pp1048-1049, 1990.

[46] M Cheney, D Isaacson, and J C Newell. Electrical impedance tomography. SIAM Rev vol 41 pp85-101, 1999.

[47] P Colli-Franzone, L Guerri, S Tentoni, C Viganotti, S Baruffi, S Spaggiari, and B Taccardi. Mathematical procedure for solving the inverse potential problem of electrocardiography. analysis of the time-space accuracy from in vitro experimental data. Mathematical Biosciences vol 77, pp353-396, 1985.

[48] D Isaacson, M Cheney Inverse Problems in Partial Differential Equations ed D Colton, R Ewing, W Rundell, (Philadelphia, and PA: SIAM) pp141 149 (1990). Current problems in impedance imaging. Int J imag Syst Technol vol 2 pp66-75, 1990.

[49] Rosalind J Sadleir and Richard A Fox. Detection and quantification of intraperitoneal fluid using electrical impedance tomography. IEEE Trans Biomed Eng vol 48 pp484-491, 2001.

[50] Habib Ammari, Ohin Kwon, Jin Keun Seo, and Eung Je Woo. T-scan electrical impedance imaging system for anomaly detection. SIAM J Appl Math vol 65 pp252-266, 2004.

[51] A E Luedtke, H W Woo, and Y Kim. Electrical impedance tomography for non-invasive detection of dvt in human calf proc. Ann Conf IEEE Eng Med Biol Soc, 13-16 Nov., 1987.

[52] A P Calderón. On an inverse boundary value problem. *Seminar on Numerical Analysis and its Applications to Continuum Physics (Rio de Janeiro, 1980)* pp65-73 Soc Brasil Mat Rio de Janeiro, 1980.

[53] Michael Lukaschewitsch, Peter Maass, and Michael Pidcock. Tikhonov regularization for electrical impedance tomography on unbounded domains. Inverse Problems vol 19 pp585-610, 2003.

[54] R L Parker. The inverse problem of resistivity sounding. Geophysics vol 142 pp2143-2158, 1984.

[55] A Ramirez, W Daily, A Binley, D LaBrecque, and D Roelant D. Detection of leaks in underground storage tanks using electrical resistance methods. J Env Eng Geophys vol 1 pp189-203, 1996.

[56] A Ramirez, W Daily, D LaBrecque, E Owen, and D Chestnut. Monitoring an underground steam injection process using electrical resistance tomography. Water Resources Res vol 29 pp73-87, 1993.

[57] F Santosa, P Kaup, and M Vogelius. A method for imaging corrosion damage in thin plates from electrostatic data. Inverse Problems vol 12 pp279-293, 1996.

[58] A N Tikhonov and V Y Arsenin. *Solutions of ill-posed problems*. Winston and Sons, New York, 1977.

[59] C G Xie, S M Huang, B S Hoyle, R Thorn, C Lenn, D Snowden, and M S Beck. Electrical capacitance tomography for flow imaging: system model for development of image reconstruction algorithms and design of primary sensors. IEEE Proceedings-G vol 139 pp89-98, 1992.

[60] M S Beck, M Byars, T Dyaowski, R Waterfall, R He, S J Wang, and W Q Yang. Principles and industrial applications of electrical capacitance tomography. Measurement+Control vol 30 pp197-200, 1997.

[61] Robert West. In industry, seeing is believing. Physics World June 2003 pp27-30, 2003.

[62] Jamil El-Ali, Peter K Sorger, and F Jensen. Cells on chips. Nature vol 442 pp403-411, 2006.

[63] S Gawad, L Schild, and P Renaud. Micromachined impedance spectroscopy flow cytometer for cell analysis and particle sizing. Lab on a Chip vol 1 pp76-82, 2001.

[64] Solly Kelli, Wang Xiaobo, Xu Xiao, Strulovici Berta, , and Zheng Wei. Application of real-time cell electronic sensing (rt-ces) technology to cell-based assays. ASSAY and Drug Development Technologies 2 (4) pp363-372, 2004.

[65] J D Pettigrew. Electroreception in monotremes. The Journal of Experimental Biology vol 202 pp1447-1454, 1999.

[66] Thodore Holmes Bullock and Walter Heiligenberg (eds). Electroreception. John Wiley Sons, New York, 1986.

[67] T H Bullock. The future of research on electroreception and electrocommunication. The Journal of Experimental Biology vol 202, pp1455-1458, 1999.

[68] Renata Freitas, GuangJun Zhang, James S. Albert, David H. Evans, and Martin J. Cohna. Developmental origin of shark electrosensory organs. Evolution and Development vol 8 pp74-80, 2006.

[69] Harold H Zakon and Graciela A Unguez. Development and regeneration of the electric organ. J of Exp Biology vol 202 pp1427-1434, 1999.

[70] Stephan Frings and Jonathan Bradley (eds). *Transduction Channels in Sensory Cells*. Wiley-VCH, New York, 2004.

[71] Gerhard von der Emde, S Schwartz, L Gomez, R Budelli, and R Grant. Electric fish measure distance in the dark. Nature vol 395 pp890-894, 1998.

[72] Chris Assad, Brias Rasnow, and Philips K Stoddard. Electric organ discharges and electric images during electrolocation. J Exp Bio 202 1185-1193, 1999.

[73] Brian Borchers Richard C Asher and Clifford H Thurber. *Parameter Estimation and Inverse Problems*. Elsevier, Amsterdam, 2005.

[74] M Bertero and P Boaccaci. *Inverse Problems in Imaging*. IOP, London, 1998.

[75] Per Christian Hansen. Regularisation tools, a mathlab package for analysis and solution of discrete ill-posed problems. available on the web http://www.imm.dtu.d, 1993.

[76] J Hadamard. *Lectures on Cauchy's Problem in Linear Partial Differential Equations*. Yale University Press, New Haven CT, 1923.

[77] Richard Mankiewicz. *The story of mathematics*. Cassel and Co, London, 2000.

[78] Peter L Bernstein. *Against the Gods, The remarkable story of risk*. John Wiley and Sons Inc, New York, 1996.

[79] A Tikhonov. On the solution of incorrectly formulated problems and the regularisation method. Soviet Math Dokl vol 4 pp1035-1038, 1963.

[80] D L Phillips. A technique for the numerical solution of certain integral equations of the first kind. J Assoc Comput Mach vol 9 pp84-97, 1962.

[81] Manus Foster. An application of the wiener-kolmogorov smoothing theory to matrix inversion. J SIAM vol 9 pp387-392, 1961.

[82] A E Hoerl. Application of ridge analysis to regression problems. Chemical Engineering Progress vol 58 pp54-59, 1962.

[83] Albert Tarantola. *Inverse Problem Theorey, Methods for Data Fitting and Model Parameter Estimation*. Elsevier, Amsterdam, 1987.

[84] S M Huang, C G Xie, R Thorn, D Snowden, and M S Beck. Design of sensor electronics for electrical capacitance tomography. IEEE Proceedings-G vol 139 pp83-88, 1992.

[85] S B Colak, M B van der Mark, G W 't Hooft, J H Hoogenraad, E S van der Linden, and F A Kuijpers. Clinical optical tomography and nir spectroscopy for breast cancer detection. IEEE Journal of Selected Topics in Quantum Electronics vol 5 pp1143-1158, 1999.

[86] M Cheny, D Isaacson, J C Newell, S Simske, and J C Goble. Noser: An algorithm for solving the inverse conductivity problem. Int J imag Syst Technol vol 2 pp66-75, 1990.

[87] V Kolehmainen, S R Arridge, W R B Lionheart, M Vauhkonen, and J P Kaipio. Recovery of region boundaries of piecewise constant coefficients of an elliptic pde from boundary data. Inv Problems vol 15 pp1375-1391, 1999.

[88] C W Groetsch. The theory of tikhonov regularization for fredholm equations of the first kind. Pitman, Boston, 1984.

[89] AA Aleksandrov, J Bourgain, M Giesecke, V Havin, and Y Vymenets. Uniqueness and free interpolation for logarithmic potentials and the cauchy-problem for the laplace equation in r(2). Geom. Funct. Anal. vol 5 529-571, 1995.

[90] O-P Tossavainen and M Vauhkonen. Estimating shapes and free surfaces with electrical impedance tomography. Proc 3rd World Congress on Industrial Process Tomography, Banff, Canada, 2003.

[91] M P Connolly, D H Michael, and R Collins. The inversion of surface potential measurements to determine crack size and shape. J Appl Phys vol 64 pp2638-2647, 1988.

[92] M McIver. Characterisation of surface breaking cracks in metal sheets by using ac electric fields. Proc R Soc A vol 421 179-194, 1989.

[93] K Bryan and M Vogelius. A computational algorithm to determine crack locations form electrostatic boundary measurements. the case of multiple cracks. Int J Eng Sci vol 32 pp579-603, 1994.

[94] N D Aparicio and M K Pidcock. The boundary inverse problem for the laplace equation in two dimensions. Inverse Problems vol 12 pp565-577, 1996.

[95] G Alessandrini, L Del Piero, and L Rondi. Stable determination of corrosion by a single electrostatic boundary measurement. Inverse Problems vol 19 pp973-984, 2003.

[96] R Siegel. Boundary perturbation method for free boundary problem in convectively cooled continuous casting. Trans ASME vol 108 pp203-235, 1986.

[97] Ibrahim Akduman and Rainer Kress. Electrostatic imaging via conformal mapping. Inverse Problems vol 18 pp1659-1672, 2002.

[98] E Beretta and S Vessella. Stable determination of boundaries from cauchy data. SIAM J Math Anal vol 30 pp220-232, 1998.

[99] G Alessandrini and L Rondi. Optimal stability for the inverse problem of multiple cavities. Journal of Differential Equations vol 176 pp356-386, 2001.

[100] J Cheng, Y C Hon, and M Yamamoto. Conditional stability estimation for an inverse boundary problem with non-smooth boundary in r^3. Trans. Amer. Math. Soc. vol 353 pp4124-4138, 2001.

[101] M A Golberg and C S Chen. The method of fundamental solutions for potential, helmholtz and diusion problems, in boundary integral methods - numerical and mathematical aspects. ed. M.A. Golberg, WIT Press, Computational Mechanics Publications, Ashurst, 103-176, 1998.

[102] Victor Isakov. *Inverse Problems for Partial Differential Equations*, volume 127 of *Applied Mathematical Sciences*. Springer Verlag, New York, 1998.

[103] David Colton and Rainer Kress. *Inverse Acoustic and Electromagnetic Scattering Theory*. Springer-Verlag, New York, 1998.

[104] David Colton, Joe Coyle, and Peter Monk. Recent developments in inverse acoustic scattering theory. SIAM Review vol 42 pp369-414, 2000.

[105] Rainer Kress. *Linear Integral Equations*. Springer Verlag, New York, 2nd edition, 1999.

[106] A Kirsch, R Kress, P Monk, and A Zinn. Two methods for solving the inverse acoustic scattering problem. Inverse Problems vol 4 pp749-770, 1988.

[107] R Potthast. A point source method for inverse acoustic and electromagnetic obstacle scattering problems. IMA J Appl Math vol 61 pp119-140, 1998.

[108] Claire D Lines and Simon N Chandler-Wilde. A point source method for inverse scattering by rough surfaces. Brunel University, Department of Mathematical Sciences Technical Report TR01/02, 2002.

[109] C D Lines and S N Chandler-Wilde. A time domain point source method for inverse scattering by rough surfaces. Computing vol 75 pp157-180 springerlink.com, 2005.

[110] Rainer Kress and Lassi Paivarinta. On the far field in obstacle scattering. SIAM J Appl Math vol 59 pp1413-1426, 1999.

[111] Klaus Erhard and Roland Potthast. The point source method for reconstructing an inclusion from boundary measurements in electrical impedance tomography and acoustic scattering. Inverse Problems vol 19 pp1139-1157, 2003.

[112] A Litman, D Lesselier, and F Santosa. Reconstruction of a two-dimensional binary obstacle by controlled evolution of a level-set. Inverse Problems vol 14 pp685-706, 1998.

[113] F Santosa. A level-set approach for inverse problems involving obstacles. ESAI: COCV vol 1 17-33, 1996.

[114] S Osher and J A Sethian. Fronts propagating with curvature dependent speed: algorithms based on the hamilton-jacobi formulation. Journal of Computational Physics vol 79 12-49, 1988.

[115] J A Sethian. Numerical algorithms for propagating interfaces: Hamilton-jacobi equations and conservation laws. Journal of Differential Geometry vol 31 131-161, 1990.

[116] Andreas Kirsch. Characterisation of the shape of a scattering obstacle using the spectral data of the far field operator. Inverse Problems vol 14 pp1489-1512, 1998.

[117] Peter Hähner. An inverse problem in electrostatics. Inverse Problems vol 15 pp961-975, 1999.

[118] Martin Brühl. Explicit characterisation of inclusions in electrical impedance tomography. SIAM J Math Anal vol 32 pp1327-1341, 2001.

[119] Birgit Schappel. Electrical impedance tomography in the half space: Locating obstacles by electrostatic measurements on the boundary. 3rd World Congress on

Industrial Process Tomography, Alberta, Canada. Dates: 2nd - 5th September 2003, pp788-73, 2003.

[120] Roland Potthast. A survey on smapling and probe methods for inverse problems. Inverse Problems vol 22 ppR1-R47, 2006.

[121] Margaret Cheney. The linear sampling method and the music algorithm. Inverse Problems 17 591-595, 2001.

[122] Andreas Kirsch. The music algorithm and the factorisation method in inverse scattering theory for inhomogeneous media. Inverse Problems vol 18 pp1025-1040, 2002.

[123] R Kress and L Kühn. Linear sampling methods for inverse boundary value problems in potential theory. Applied Numerical Mathematics vol 43 pp161-173, 2002.

[124] Michael Brandfass, Aaron D Lanterman, and Karl F Warnick. A comparison of the colton-kirsch inverse scattering methods with linearized tomographic inverse scattering. Inverse Problems vol 17 pp1797-1816, 2001.

[125] A.G. Ramm and Semion Gutman. Analysis of a linear sampling method for identification of obstacles. Acta Mathematicae Applicatae Sinica vol 21 pp399-404, 2005.

[126] Yves Meyer and Ronald Coifman. *Wavelets, Calderon-Zygmund and multilinear operators*. Cambridge University Press, Cambridge, 1997.

[127] Gerald B Folland. *Introduction to Partial Differential Equations*. Princeton University Press, Princeton, 1995.

[128] Mitchell H Taibleson. On the theory of lipschitz spaces of distributions on euclidean n-space i. principal properties. J. Math. Mech. 13 1964 407479, 1964.

[129] Gregory Verchota. Layer potentials and regularity for the dirichlet problem for laplace's equation in lipschitz domains. Journal of Functional Analysis vol 59 pp572-611, 1984.

[130] S N Chandler-Wilde. The impedance boundary value problem for the helmholtz equation in a half-plane. Mathematical Methods in the Applied Sciences vol 20 pp813-840, 1997.

[131] Simon N Chandler-Wilde and Peter Monk. Existence, uniqueness, and variational methods for scattering by unbounded rough surfaces. SIAM J Math Anal. vol 37 pp598618, 2005.

[132] Mark D Preston, Peter G Chamberlain, and Simon N Chandler-Wilde. An integral equation method for a boundary value problem arising in unsteady water wave problems. Submitted, 2006.

[133] G Krishnasamy, F J Ripo, and T J Rudolphi. Hypersingular boundary integral equations: their occurrence, interpretation, regularization and computation. in P K Banerjee and S Kobayashi (eds.), Developments in Boundary Element Methods, Elsevier, Amsterdam, pp207-252, 1992.

[134] G Monegato. The numerical evaluation of hypersingular integrals. Journal of Computational and Applied Mathematics vol 50 pp9-31, 1994.

[135] P A Martin and F J Rizzo. Hypersingular integrals: how smooth must the density be? International Journal for Numerical Methods in Engineering, vol 39 pp687-704, 1996.

[136] M N J Moore, L J Gray, , and T Kaplan. Evaluation of supersingular integrals: Second-order boundary derivatives. International Journal for Numerical Methods In Engineering, vol 69 pp19301947, 2007.

[137] C A Brebbia and S Walker. *Boundary Element Techniques in Engineering*. Butterworth, London, 1st edition, 1980.

[138] Tilo Arens. Why linear sampling works. Inverse Problems vol 20 pp163-173, 2004.

[139] Lord Kelvin (Sir W Thomson). *Reprints of Papers on Electrostatics and Magnetism*. Macmillan, London, 2nd edition, 1884.

[140] J W Dauben. *The history of mathematics from antiquity to the present: a selective bibliography.* Garland, New york and London, 1985.

[141] David Colton and Rainer Kress. *Integral Equation Methods in Scattering Theory.* Wiley Interscience, New York, 1983.

[142] L Greengard and V Rokhlin. A fast algorithm for particle simulations. J Comput Phys vol 73 pp325-348, 1987.

[143] Xiaobai Sun and Nikos P Pitsianis. A matrix version of the fast multipole method. Siam Review vol 43 pp289-300, 2001.

[144] Numerical Algorithms Group. Nag fortran library, mark 18. Current manual (mark 20) available on the web, www.nag.co.uk, 1997.

[145] Yousef Saad and Henk A van der Vorst. Iterative solution of linear systems in the 20th century. J of Comp and Appl Mathematics 123 pp1-33, 2000.

[146] Gene H Golub and Dianne P O'Leary. Some history of the conjugate gradient and lanczos algorithms: 1948-1976. Siam Review vol 31 pp50-102, 1989.

[147] G W Stewart. On the early history of the singular value decomposition. SIAM Review vol 35 551, 1993.

[148] Philip E Gill, Walter Murray, and Margaret H Wright. *Practical Optimization.* Academic Press, London, 1981.

[149] Martin Hanke and Per Christian Hansen. Regularization methods for large-scale problems. Surveys on Mathematics for Industry vol 3 pp253-315, 1993.

[150] Josie Wernecke. *The Inventor Mentor.* Addison-Wesley Publishing Company, Reading Massachusetts, 1994.

[151] A J Kinderman and J F Monahan. Computer generation of random variables using the ratio of uniform deviates. ACM Transactions on Mathematical Software vol 3 pp257-260, 1977.

[152] Olivier Pironneau. Optimal shape design for elliptic systems. Springer, Berlin, 1984.

[153] Frank Hettlich and William Rundell. Iterative methods for the reconstruction of an inverse potential problem. Inverse Problems vol 13 pp251-266, 1996.

[154] A Charalambopoulos. On the frechet-differentiability of boundary integral operators in the inverse elastic scattering problem. Inverse Problems vol 11 1137-62, 1995.

[155] Roland Potthast. Frechet differentiability of boundary integral operators and its applications to an inverse acoustic scattering problem. Inverse Problems vol 10 pp431-447, 1994.

[156] A Kirsch. The domain derivative and two applications in inverse scattering theory. Inverse Problems vol 9 pp81-96, 1993.

[157] Eldad Haber, Uri M Ascher, and Doug Oldenburg. On optimization techniques for solving nonlinear inverse problems. Inverse Problems vol 16 pp1263-1280, 2000.

[158] M Hanke. A regularizing levenburg-marquardt scheme, with applications to inverse ground water flitration problems. Inverse Problems vol 13 pp79-95, 1997.

[159] Barbara Kaltenbacher. Some newton-type methods for the regularization of nonlinear ill-posed problems. Inverse Problems vol 13 pp729-753, 1997.

[160] Thorsten Hohage. Logarithmic convergence rates of the iteratively regularized gaussnewton method for an inverse potential and an inverse scattering problem. Inverse Problems vol 13 pp12791299, 1997.

[161] B Kaltenbacher, A Neubauer, and A G Ramm. Convergence rates of the continuous regularized gaussnewton method. J Inv Ill-Posed Problems vol 10 pp261-280, 2002.

[162] George Arfken. *Mathematical Methods for Physicists*. Academic Press, New York, 2nd edition, 1970.

[163] Helmut Harbrecht and Thorsten Hohage. Fast methods for three-dimensional inverse obstacle scattering problems. to appear in J Int Eq Appl, 2007.

[164] Athanasios D Zacharopoulos, Simon R Arridge, Oliver Dorn, Ville Kolehmainen, and Jan Sikora. Three-dimensional reconstruction of shape and piecewise constant region values for optical tomography using spherical harmonic parametrization and a boundary element method. Inverse Problems vol 22 pp1509-1532, 2006.

[165] H Tortel. Electromagnetic imaging of a three-dimensional perfectly conducting object using a boundary integral formulation. Inverse Problems vol 20 pp385-398, 2004.

[166] Thorsten Hohage. Fast numerical solution of the electromagnetic medium scattering problem and applications to the inverse problem. Journal of Computational Physics vol 214 pp224-238, 2004.

[167] Saeed Babaeizadeh and Dana H Brooks. Spherical harmonics for shape-based inverse problems as applied to electrical impedance tomography. Proc SPIE Computational Imaging IV, Charles A Bouman, Eric L Miller, Ilya Pollak, Eds, pp, 2006.

[168] Sungwhan Kim, Ohin Kwon, and Jin Keun Seo. Location search techniques for a grounded conductor. SIAM J Appl Math vol 62 pp1383-1393, 2002.

[169] A Kirsch and R Kress. An optimisation method in inverse acoustic scattering. in Boundary Elements IX, C Brebbia, W Wendland and G Kuhn eds, Springer-Verlag, Heidelberg pp3-18, 1987.

[170] Ulrich Tautenhahn. On the asymptotical regularisation of nonlinear ill-posed problems. Inverse Problems vol 10 pp1405-1418, 1994.

[171] L Marin, L Elliott, D B Ingham, and D Lesnic. Boundary element method for the cauchy problem in linear elasticity. Engineering Analysis with Boundary Elements vol 25 pp 783-793, 2001.

[172] Gottfried Bruckner and Sergei V Pereverzev. Self-regularization of projection methods with a posteriori discretization level choice for severely ill-posed problems. Inverse Problems vol 19 pp147-156, 2003.

[173] Patricia K Lamm. Variable-smoothing local regularization methods for first-kind integral equations. Inverse Problems vol 19 pp195-216, 2003.

[174] A Cimetiere, F Delvare, M Jaoua, and F Pons. Solution of the cauchy problem using iterated tikhonov regularization. Inverse Problems vol 17 pp553-570, 2001.

[175] O Scherzer and C W Croetsch. Scale-space and morphology in computer vision. Springer Lecture Notes in Computer Science, Berlin: Springer vol 2106 pp 317-25, 2001.

[176] J Kaipio, V Kolehmainen, E Somersalo, and M Vauhkonen. Statistical inversion and monte carlo sampling methods in electrical impedance tomography. Inverse Problems vol 16 pp1487-1522, 2000.

[177] Oliver Dorn, Eric L Miller, and Carey M Rappaport. A shape reconstruction method for electromagnetic tomography using adjoint fields and level sets. Inverse Problems vol 16 pp1119-1156, 2000.

[178] A Leitao and O Scherzer. On the relation between constraint regularisation, level sets, and shape optimization. Inverse Problems vol 19 ppL1-L11, 2003.

[179] P C Hansen and Dianne P O'Leary. The use of the l-curve in the regularization of ill-posed problems. SIAM Journal on Scientific Computing vol 14 pp1487-1503, 1993.

[180] U Tautenhahn and U Hamarike. The use of monotonicity for choosing the regularization parameter in ill-posed problems. Inverse Problems vol 15 pp1487-1505, 1999.

[181] Heinz W Engl, Martin Hanke, and Andreas Neubauer. *Regularization of Inverse Problems*. Kluwer Academic Publishers, Dordrecht, 1996.

[182] A B Bakushinskii. Remarks on choosing a regularization parameter using the quasi-optimality and ratio criterion. USSR Comput Math Math Phys vol 24 pp181-182, 1984.

[183] Federico París and José Canas. *Boundary Element Method, Fundamentals and Applications*. Oxford University Press, Oxford, 1st edition, 1997.

[184] Andreas Karageorghis and Graeme Fairweather. The method of fundametal solutions for axisymmetric potential problems. Int J Numer Meth Engng vol 44 pp1653-1669, 1999.

[185] Murat Belge, Misha E Kilmer, and Eric L Miller. Efficient determination of multiple regularization parameters in a generalized l-curve framework. Inverse Problems vol 18 pp1161-1183, 2002.

[186] Frank Natterer. *The Mathematics of Computerized Tomography.* John Wiley Sons, Stuttgart, 1986.

[187] Martin Hanke and Birgit Schappel. The factorisation method for electrical impedance tomography in the half space. submitted to Siam Journal on Applied Mathematics, 2006.

[188] Z Anal Anwendungen vol 25 pp81 102 (2006). The factorisation method for real elliptic problems. Quarterly Journal of Medicine vol 95, pp557-558, 2002.

[189] David Colton, Michele Piana, and Roland Potthast. A simple method using morozovs discrepancy principle for solving inverse scattering problems. Inverse Problems vol 13 pp1477-1493, 1997.

[190] Andrea Tacchino, Joe Coyle, and Michele Piana. Numerical validation of the linear sampling method. Inverse Problems vol 18 pp511527, 2002.

[191] Martin Hanke and Martin Brühl. Recent progress in electrical impedance tomography. Inverse Problems vol 19 pp1-26, 2003.

[192] A Tamburrino and G Rubinacci. A new non-iterative inversion method for electrical resistance tomography. Inverse Problems vol 18 pp1809-1829, 2003.

[193] Robert G Aykroyd, Manuchehr Soleimani, and William R B Lionheart. Full shape reconstruction from partial ert monotonicity information using a bayes-mcmc approach. 4th World Congress on Industrial Process Tomography, Aizu, Japan, 1999.

[194] R Haggarty. *Fundamentals of mathematical Analysis.* Prentice Hall, Harlow, 1992.

[195] Gerald B Folland. *Real Analysis.* John Wiley, New York, 1984.

[196] William McLean. *Strongly Elliptic Systems and Boundary Integral Equations.* Cambridge University Press, Cambridge, 2000.

Appendix A

ADAM PROGRAM

The majority of the numerical experiments reported in this thesis have been obtained with the program `DrAdam.exe`, the program is available and can be used to reproduce the results or perform additional tests.

A.1 Installation and Operation

The program runs on Microsoft Windows, principally 98, 2000 and XP. In addition to the standard run-time libraries (DLLs), a number of additional libraries are required for the NAG and Open Inventor support. These are provided with the program. A license is required for the open inventor libraries, details of which are provided in the warning dialog box that appears on program start-up.

- Copy the `DrAdam_run` directory to the location of choice.
- Copy all the files in `DrAdam_run/DLLs` to a directory in the run path, for instance to `Windows/System32`, or add the name `xxx/DrAdam_run/DLLs` to the run path, in which `xxx` is the drive and directory that the `DrAdam_run` directory has been created in.
- Create an environment variable `OIVHOME=xxx/DrAdam_run`, How this is done depends on the operating system. In Windows XP, right click My Computer, then select Properties, Advanced tab and finally Environment Variables.
- Run the program simply by clicking on the `DrAdam.exe` in the directory.

Upon startup, the program presents the OnNew dialog box shown in figure A.1 which sets all the parameters for program use. When `OK` is selected on this dialog box, the

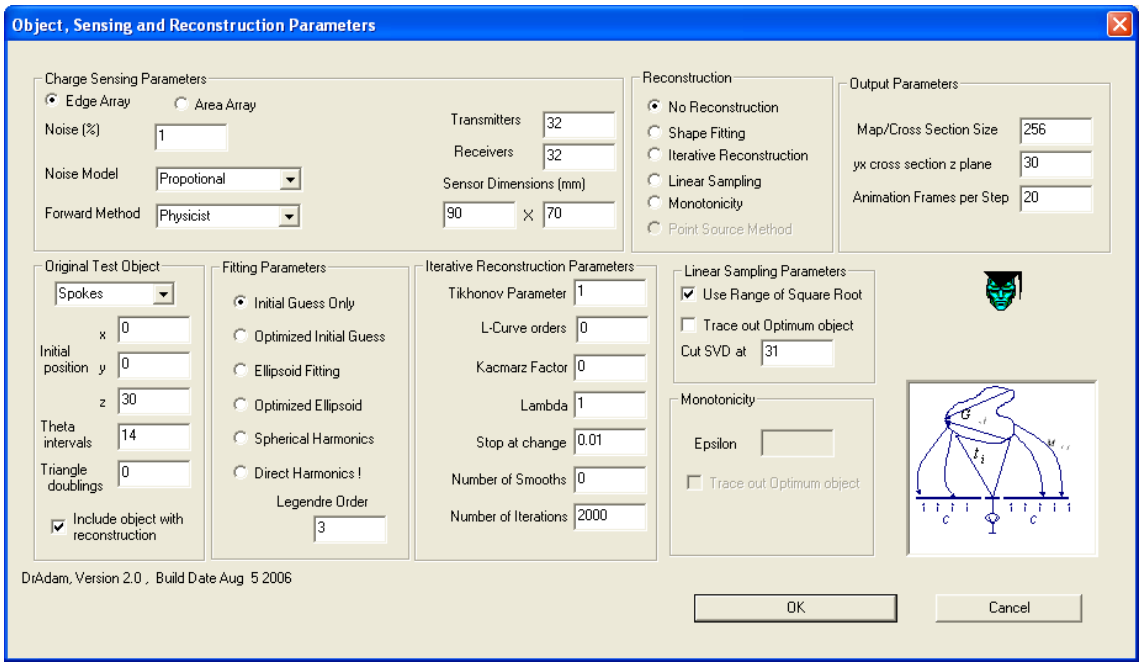

Fig. A.1: Adam's File:OnNew Dialog box.

main program window opens up with an Open Inventor Examiner window showing a model of the fingermouse cross capacitance sensing system with a test object located above the display / sensing area.

When `Calculate` is then selected on the menu bar, the calculation specified in the Dialog box is started. The result are presented in an output strip discussed in section 4.5. The output strip is also saved to disk as a .bmp file. The program also produces a `Monitor.txt` file with useful output information. If the `Calculate` menu item is selected a second time, the program executes a series of reconstructions with the object in different positions above the sensor.

A.2 The Parameter Options

A.2.1 Charge Sensing Parameters

The two radio buttons chose between an linear array of electrodes around the edge of the display (`Edge Array`) or a 2D array spread over the surface of the display itself (`Area Array`). In the latter case the number of `Transmitters` or `Receivers` needs to be a square. For n electrodes (transmitters or receivers) the system lays out $\sqrt{n} \times \sqrt{n}$

electrodes. The sensor dimensions determine the area of over which the 2D array is spread or the rectangle along whose perimeter the edge array is laid out. This parameter allows experimentation with the physical aperture of the measurements.

Random noise can be added to the simulated data. The noise is normally distributed with a Ratio of Uniforms method, an algorithm due to Kindermann *et al*[151] and augmented with quadratic bounding curves. The standard deviation is entered as a percentage under `Noise`. How that percentage is used depends on the `Noise Model` drop down box. The options are `Proportional` in which case noise is added to the signal of each transmitter and receiver as a percentage proportion of that signal. This assumes that the system adjusts its gain in response to the signal strength on each receiver. In the case of the `Uniform` noise model, noise is added to all measurements as a percentage of the norm of the forward matrix $||\Lambda_{D\,i,j}||$. This essentially assumes that the receiver gain is set one value for all receivers for all measurements.

Four different `Forward Method`s can be chosen with the drop down box. The `Physicist` method uses the $TS^{-1}T^*$ factorisation (3.25). The `Experimentalist` method can be written as $T(S\backslash t)$ because is solves (3.20) and applies (3.21) for each transmitter independently. The `Mathematician` method uses factorisation $\tilde{M}(\frac{1}{2}I + \tilde{K})^{-1}T^*$ of (3.30). Finally the `Pseudo` method uses a $TS^{\ddagger}T^*$ factorisation in which the pseudo inverse $S^{\ddagger}$ is created using

$$\begin{aligned} S^{\ddagger}_{i,j} &= 0 && \text{for } i \neq j \\ &= \frac{\partial D_i}{\sum_k \partial D_k S_{i,k}} && \text{for } i = j \end{aligned}$$

The calculation of this inverse is very fast which can make this option useful. The pseudo inverse is also useful because in the execution of particularly the optimization algorithms, it can happen that the forward map is evaluated for non-physical objects and the matrix $S_{i,j}$ is not invertible. In particular, during the spherical harmonics optimization in section 5.3 it may be that a parameter set occurs for which the expansion in equation (5.6) becomes negative for some (θ, ϕ). This results in an object for which S^{-1} in (3.25) becomes impossible to evaluate. To allow the Gauss-Newton iteration to proceed, the program automatically calculates the pseudo forward map instead and this provides a convenient method to allow the algorithm to proceed without interruption.

A.2.2 Reconstruction

These parameters control the reconstruction algorithm that will be used when the menu item `Calculate` is selected later. The option `No Reconstruction` is self explanatory, if this option is selected the program produces an output strip as illustrated in figures and 4.4 and 4.6. The `Shape Fitting` reconstruction algorithms are described in chapter 5, the options are detailed below in subsection A.2.5. The `Iterative Reconstruction` alogrithm is described in chapter 6 and the options below in subsection A.2.6. The `Linear Sampling` method is discussed in chapter 7 and the options below in subsection A.2.7. Finally the `Monotonicity` method is discussed in chapter 8 and the options below in subsection A.2.8. The `Point Source` method is not currently active but may be included in a future version of the program.

A.2.3 Output Parameters

The output of the program can be controlled only to a relatively minor degree. The `Map/Cross Section Size` controls the size of the panels in the output strip. The parameter is particularly important for the Linear Sampling and Monotonicity algorithms because the size of the cross section images is also the resolution of the sampling grid. The `yx cross section z-plane` value controls the z value at which the cross section in the fourth image in the output strip is taken. The y value in the xz cross section of the third image is always zero. The reconstruction animation, which is started with the menu item `Calculate` is selected a second time, is a sequence of 10 positions for the original object. The `Animation Frames per Step` controls how many intermediate positions are created between the 10 positions. The animation frames are stored to disk.

A.2.4 Original Test Object

This section controls the shape and triangulation density of the original object. A number of test objects can be chosen from the drop down box. Some of these are internally generated star shapes (Sphere, Ellipsoid, Spokes and Kite) and some are loaded from .iv or .wrl file (Bishop, Knight, Body). For the latter group the triangulation density is determined by the inventor or vrml code in the file. Additionally it is possible to double this triangulation once or more times to make a more dense triangle grid. For the

internally generated objects, the triangle density is controlled by setting the number intervals that the range $[0, \pi]$ of the altitude angle θ is divided into. 14 intervals create an object of 420 triangles, while 28 intervals cause the object to have 1826 triangles. This parameter also determines the triangle density of the reconstructed object. If the option `Include object with reconstruction` is ticked, both the original and reconstructed object are shown in the first, camera view of the output strip. This can be useful for direct comparison although it sometimes hides details in the reconstructed object that may be of interest.

A.2.5 Fitting Parameters

The `Fitting Parameters` determine the kind of object that is being fitted if the `Shape Fitting` algorithm is selected (cf section A.2.2). The `Initial Guess Only` option creates only the sphere that is used as initial guess for all other shape optimization routines and the iterative procedure in Chapter 6. Details of this are given in section 5.1. The `Optimized Initial Guess` options improves the initial guess by optimizing position and size of the sphere against the simulated data with a full forward calculation. The `Ellipsoid Fitting` option optimizes the orientation of an ellipsoid with fixed axis ratio and fixed position of the initial guess, as described in section 5.2. The option `Optimized Ellipsoid` takes this a stage further by simultaneously optimizing orientation and axis of the ellipsoid. The `Spherical Harmonics` option fits the coefficients of a spherical harmonics description up to the `Legendre Order`. The code in the program to calculate the Legendre polynomials is based on bobvlieg2.cpp by Thomas Germer[1]. The `Direct Harmonics` option also creates a spherical harmonics fit to the object but does this directly from the object as described in appendix B.

A.2.6 Iterative Reconstruction Parameters

The parameters in the `Iterative Reconstruction Parameters` determine the execution of the reconstruction algorithm described in chapter 6. The `Tikhonov Parameter` sets $\hat{\alpha}$ defined in section 6.4. The parameter is used at the start of an iterative reconstruction, and the iterative reconstruction animation, to calculate α_q using operator norms based on the initial guess. If `L-Curve orders` is set to zero, then this value is

[1] http://physics.nist.gov/Divisions/Div844/facilities/scatmech/code/bobvlieg2.cpp

used for the regularisation of the functionals in section 6.1. If the `L-Curve orders` is set to non zero, then the program uses an L-curve criterion to find the best regularisation parameter within a log10 orders range around α_q.

The `Kacmarz Factor` must be set between 0 and 1. A value of 0 means that the results from multiple transmitters are averaged as illustrated in figures 6.9A and 6.9B, whereas a value of 1 means that the results are treated in a Kacmarz like approach as illustrated in figures 6.9C and 6.9D. Intermediate values for `Kacmarz Factor` result in a mix of the two. The `Lambda` parameter sets the value of $\hat{\lambda}$ defined in section 6.4, like $\hat{\alpha}$ this value is used once, at the beginning to calculate λ. The `Stop at Change` parameter sets the threshold E for the stopping criterion as described in equation (6.17). If set to zero the iteration will run for a maximum of iterations set by `Number of Iterations`.

If the inverse charge imaging is poorly regulated, the object reconstruction described in section 6.2 can cause and amplify irregularities in the reconstructed shape. One way to avoid this is to increase `Tikhonov Parameter` or increase the range over which the L-curve criterion is used to look for an optimum value. As an alternative it is possible to perform one or more smoothing operations on the object after each iteration, by setting `Number of Smooths` to a non-zero positive value.

A.2.7 Linear Sampling Parameters

The section `Linear Sampling Parameters` sets options for the reconstructions with an indicator function as described in chapter 7. The option `Use Range of Square Root` chooses, if set, between the Factorisation method or the linear sampling method, if not set. The option `Trace out Optimum Object`, if set, uses the minimization of the cost function to determine the optimum level of $P(z)$ to find the object. To do this the program must sample $P(z)$ over a 3D volume which can make setting the option quite expensive in computation terms. If the option is not set the level set is sampled only $P(z)$ on the grid set by `Map/Cross Section Size`. Finally, the `Cut SVD at` parameter is the truncated SVD parameter at which the sum for $P(z)$ is cut to regularize its calculation.

A.2.8 Montonicity Method

None of the options for the Monotonicity method are currently active. Please note that the method is very computationally intensive and the user is advised to execute the method with `Map/Cross Section Size` set to a very low value, for instance 16.

Appendix B

DIRECT SPHERICAL HARMONICS

Section 5.3 experimented with limited success with the optimization reconstruction of objects using spherical harmonics. The expansion in spherical harmonics in equation (5.6) describes any star shaped object, but it is valid to ask how many orders are needed for a satisfactory shape description of a given object, regardless of the input data, the method used to find the set of parameters or indeed the sensing technology. Also, for shapes that fail to be a star shape, how well does a finite expansion of low order spherical harmonics approximate the object. A few numerical experiments that explore these questions are presented here

B.1 Star Shapes

Given an object with surface $r(\theta, \phi)$ the coefficients can be calculated using the orthogonality integrals for Legendre functions, cosines and sines. This gives;

$$C_l^m = \frac{2l+1}{2\pi}\frac{(l-m)!}{(l+m)!}\int_0^{2\pi}\int_0^{\pi} r(\theta,\phi)P_l^m(\cos\theta)\cos(m\phi)\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi \qquad \text{(B.1)}$$

$$S_l^m = \frac{2l+1}{2\pi}\frac{(l-m)!}{(l+m)!}\int_0^{2\pi}\int_0^{\pi} r(\theta,\phi)P_l^m(\cos\theta)\sin(m\phi)\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi \qquad \text{(B.2)}$$

The first cosine term (C_0^0) represents the radius of the object and does not depend on choice of origin. All of the first sine terms (S_l^0) are zero. All other coefficients do depend on the choice of origin We chose as our origin of the object the center of gravity in surface integral terms.

$$z = \frac{\int_{\partial D} y\,\mathrm{d}s(y)}{\int_{\partial D}\mathrm{d}s(y)} \qquad \text{(B.3)}$$

The program uses NAG routine `D01EAF` on $[0,\pi]\times[0,2\pi]$ to perform the multi-

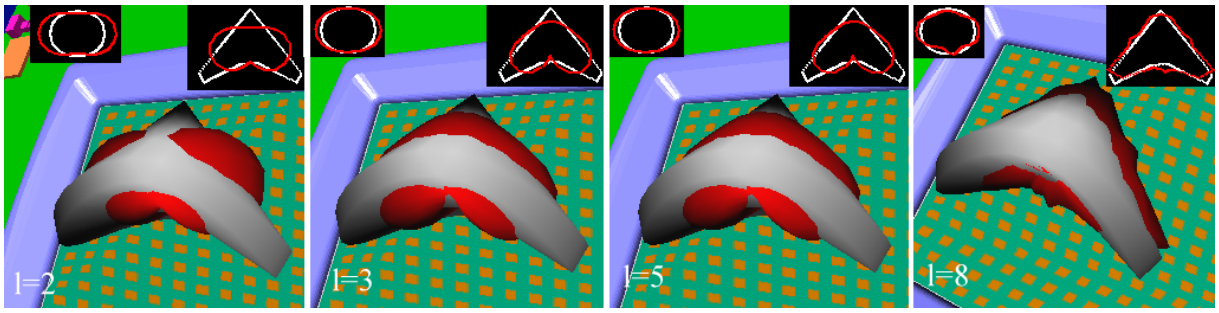

Fig. B.1: Direct calculations of Spherical Harmonics of a kite shaped object for different Legendre orders.

ple integrals for the required coefficients simultaneously. In the integrands of (B.1) and (B.2), the function $r(\theta, \phi)$ is most costly to evaluate. The program uses a piecewise linear approximation with linear interpolation of the vertex values of the triangle containing (θ, ϕ).

Figure B.1 shows the approximation of the kite shaped object used in figure 5.4 by increasing partial sum of (B.1) and (B.2). As can be seen increasing orders provide and increasingly good approximation but the description is never quite snug, even at $l = 8$.

B.2 The Bishop

It is also interesting to ask how good spherical harmonics are for objects that are not star shapes? As case study we take the bishop chess piece used in chapter 5. It is now not possible to use (B.1) and (B.2) to calculate the coefficients but instead need to define an optimization problem to find a 'best' star shape to describe the bishop. It is natural to define the best star shape in terms of shared volume. That is we wish to minimize the cost function.

$$F(z, C_m^l, S_m^l) = \int_H \left[\Phi_B(x) - \Phi_S(x)\right]^2 \, \mathrm{d}x \tag{B.4}$$

In which $\Phi_B(x)$ and $\Phi_S(x)$ are the flux calculated with equation (4.16) for the Bishop and Starshape object respectively. In our calculation, the volume integral is divided into solid angles subtended by the triangles making up the starshape. Inside each solid angle the integral is calculated using piecewise constant quadrature. Like the optimizations discussed in chapter 5, the optimization is performed using `E04FCF` for

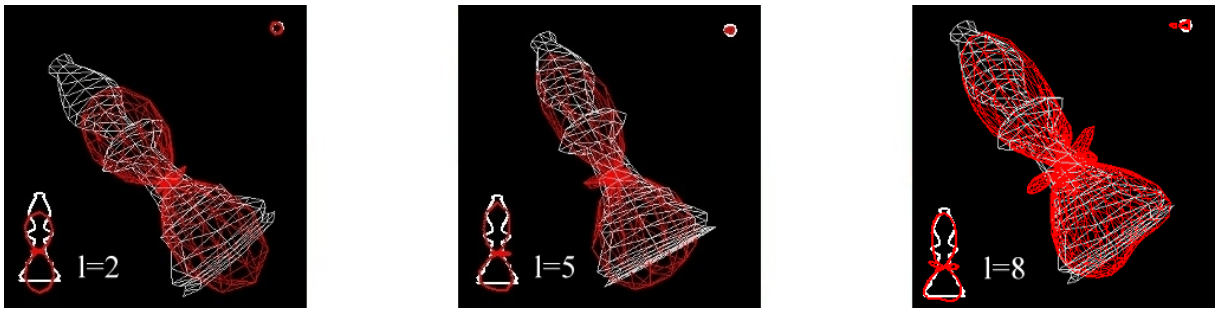

Fig. B.2: How good are harmonics at building the Bishop?, The l=2 shows the kind of fit we'd expect from familiar spherical harmonics as electron orbitals

the modified Gauss-Newton method.

Figure B.2 shows the results for the bishop object. As can be seen the spherical harmonics are quite good at approximating the shape but the fit is not as snug as we might wish, particularly for $l = 8$. This is due to the limits of a numerical optimization with a large number of parameters. Although object optimization with the cost function (B.4) it easier than with cost function (5.2), the condition number of the Jacobian still increases as the number of Legendre order increases.

Appendix C

SUPPORTING LEMMAS

In this chapter a number of theorems and lemmas are collected that are used in the main body of the thesis.

C.1 Half Space Green's Function

Proposition C.1. *The Green's function*

$$G(x,y) = \Phi(x,y) - \Phi(x,y') = \frac{1}{4\pi|x-y|} - \frac{1}{4\pi|x-y'|},$$

has asymptotic behaviour

$$G(x,y) = O\left(\frac{\cos\theta}{|x|^2}\right)$$

and

$$\frac{\partial G(x,y)}{\partial x_j} = \quad O(1/|x|^3)$$

as $|x| \to \infty$ *with* $x \in H$, *where* $\cos\theta = x_3/|x|$.

Proof. It is sufficient to write $y_1 = y_2 = 0, y_3 = d$ and $x_1 = r\sin\theta, x_2 = 0, x_3 = r\cos\theta$. Then

$$\frac{1}{|x-y|} = \frac{1}{\sqrt{r^2\sin^2\theta + (r\cos\theta - d)^2}}, \tag{C.1}$$

$$\approx \frac{1}{\sqrt{r^2 - 2dr\cos\theta}} \quad \text{for } r >> d, \tag{C.2}$$

$$\approx \frac{1}{r} + \frac{rd\cos\theta}{r^3}. \tag{C.3}$$

Similarly, using $y'_1 = y'_2 = 0, y'_3 = -d$,

$$\frac{1}{|x-y'|} =\approx \frac{1}{r} - \frac{rd\cos\theta}{r^3}. \tag{C.4}$$

We therefore have

$$\begin{aligned} G(x,y) &\approx \frac{2d\cos\theta}{4\pi r^2} \quad \text{for } r >> d, & \text{(C.5)}\\ &= O\left(\frac{\cos\theta}{|x|^2}\right). & \text{(C.6)} \end{aligned}$$

The property for $\partial G(x,y)/\partial y_j$ for $j = 1,2,3$ follows from fact that $G(x,y)$ is symmetric in x, y, and then differentiation of (C.5) with respect to x_j. This gives x_j

$$\frac{\partial}{\partial x_j}\frac{1}{x_1^2+x_2^2+x_3^2} = \frac{-2x_j}{(x_1^2+x_2^2+x_3^2)^2}.$$

At one extreme, if we consider an x_j =constant plane, this is $O(1/|x|^4)$, whereas at the other extreme along the x_j-axis it is $O(1/|x|^3)$.

Alternatively we can differentiate (C.1) with respect to d. This gives

$$\begin{aligned} \frac{\partial}{\partial d}\frac{1}{|x-y'|} &= \frac{r\cos\theta - d}{\left(r^2\sin^2\theta + (r\cos\theta - d)^2\right)^{3/2}}, & \text{(C.7)}\\ &\approx \frac{r\cos\theta - d}{\sqrt{r^2 - 2dr\cos\theta}} \quad \text{for } r >> d, & \text{(C.8)}\\ &\approx (r\cos\theta - d)\left[\frac{1}{r^3} + \frac{3d\cos\theta}{r^4}\right]. & \text{(C.9)} \end{aligned}$$

A similar expression for the negative term ($y'_3 = -d$) can be derived. We then obtain

$$\begin{aligned} \frac{\partial G(x,y)}{\partial d} &\approx (3\cos^2\theta - 1)\frac{d}{r^3}, & \text{(C.10)}\\ &= O\left(\frac{1}{|x|^3}\right). & \text{(C.11)} \end{aligned}$$

□

Lemma C.2. *Let D be sphere with radius r. Let $z \geq r$ measure distance from the*

centre of the sphere. The single layer potential with unity moment $\phi = 1$ is given by

$$u(z) = \frac{1}{4\pi} \int_{\partial D} \frac{1}{|z-y|} \, \mathrm{d}s(y) = \frac{r^2}{z} \tag{C.12}$$

Proof.

$$\int_{\partial D} \frac{\mathrm{d}s(y)}{|z-y|} = \int_0^{\pi} \int_0^{2\pi} \frac{r^2 \sin\theta \, \mathrm{d}\phi \, \mathrm{d}\theta}{\sqrt{r^2 \sin^2\theta + (z - r\cos\theta)^2}}$$

$$= \int_0^{\pi} \int_0^{2\pi} \frac{r^2 \sin\theta \, \mathrm{d}\phi \, \mathrm{d}\theta}{\sqrt{r^2 \sin^2\theta + z^2 - 2rz\cos\theta + r^2\cos^2\theta}} = \int_0^{\pi} \frac{2\pi r \sin\theta \, \mathrm{d}\theta}{\sqrt{1 + (z/r)^2 - 2(z/r)\cos\theta}}$$

$$= -\frac{\pi r^2}{z} \int_0^{\pi} \frac{\mathrm{d}2(z/r)\cos\theta}{\sqrt{1 + (z/r)^2 - 2(z/r)\cos\theta}} = \frac{\pi r^2}{z} \int_{2(z/r)}^{-2(z/r)} \frac{\mathrm{d}(1 + (z/r)^2 - x)}{\sqrt{1 + (z/r)^2 - x}}$$

$$= \frac{\pi r^2}{z} \int_{(z/r-1)^2}^{(z/r+1)^2} \frac{\mathrm{d}y}{\sqrt{y}} = 2\frac{\pi r^2}{z} \sqrt{y} \Big]_{(z/r-1)^2}^{(z/r+1)^2}$$

$$= \frac{4\pi r^2}{z}$$

□

The continuous extension to the boundary of the double layer potential with the Poisson kernel, can be demonstrated as a general result, see for instance claim 2.38 and theorem 2.44 in Folland[127]. Here we first want to demonstrate the identity on the boundary directly for a continuous bounded and square integrable function on the plane, and extend this to functions in L^2 in the subsequent lemma, following the arguments of Folland[127].

Lemma C.3. *Let $G(x,y) = \frac{1}{4\pi|x-y|} - \frac{1}{4\pi|x-y'|}$ and $f_c \in (C(\partial H) \cap L^{\infty}(\partial H) \cap L^2(\partial H))$ then*

$$\lim_{x_3 \downarrow 0} \int_{\partial H} \frac{\partial G(x,y)}{\partial y_3} f_c(y) \, \mathrm{d}s(y) = f_c(x), \quad x \in \partial H. \tag{C.13}$$

Proof. Consider

$$u(x)\big|_{\partial H} = \lim_{x_3 \downarrow 0} \int_{\partial H} f_c(y) \frac{x_3}{2\pi|x-y|^3} \, \mathrm{d}s(y)$$

Locate in ∂H a disk $D(x; \sqrt{x_3})$, centered at x_1, x_2 of radius $\sqrt{x_3}$. By the integral mean value theorem there is a $x_D \in D$ and $x_E \in \partial H \backslash D$, such that

$$\begin{aligned} u(x)\big|_{\partial H} &= \lim_{x_3\downarrow 0} f_c(x_D)\int_D \frac{x_3}{2\pi|x-y|^3}\,\mathrm{d}s(y) + \lim_{x_3\downarrow 0} f_c(x_E)\int_{\partial H\setminus D}\frac{x_3}{2\pi|x-y|^3}\,\mathrm{d}s(y) \\ &= \lim_{x_3\downarrow 00} f_c(x_D)\int_0^{\sqrt{x_3}}\frac{x_3}{(x_3^2+r^2)^{3/2}}r\,\mathrm{d}r + \lim_{x_3\downarrow 0} f_c(x_E)\int_{\sqrt{x_3}}^{\infty}\frac{x_3}{(x_3^2+r^2)^{3/2}}r\,\mathrm{d}r \\ &= \lim_{x_3\downarrow 0}\left[f_c(x_D) - \frac{f_c(x_D)}{\sqrt{1+1/x_3}} + \frac{f_c(x_E)}{\sqrt{1+1/x_3}}\right] \\ &= f_c(x). \end{aligned} \tag{C.14}$$

which is what we wanted to show □

To extend this proof to functions in $L^2(\partial H)$ we would need a regularity condition equivalent to the mean value theorem. This is given by the condition of a Lebesque point. A point x in the domain of f_H is a Lebesgue point if

$$\lim_{\epsilon\to 0^+}\frac{1}{|B(x,\epsilon)|}\int_{B(x,\epsilon)}|f_H(y)-f_H(x)|\,\mathrm{d}y = 0.$$

Here, $B(x,\epsilon)$ is the ball centered at x with radius ϵ, and $|B(x,\epsilon)|$ is the Lebesgue measure of that ball. So a Lebesque point is a point that is 'regular' enought for the integral mean value theorem to apply over the ball $|B(x,\epsilon)|$. It can be shown that, given any $f\in L^1$, almost every x is a Lebesgue point, and the proof above could be applied to functions in L^2. Equivalently, we make the observation that because the continuous functions are dense in L^2 and so for a $f_H\in L^2$ and $\epsilon>0$ we can choose an f_c such that $||f_H-f_c||^2<\epsilon$. By taking $\lim_{\epsilon\to 0}$, the idenity (C.13) becomes an approximate identity in the L^2 sense as in the next lemma:

Lemma C.4. *Let* $G(x,y)=\frac{1}{4\pi|x-y|}-\frac{1}{4\pi|x-y'|}$ *and* $f(x)\in L^2(\partial H)$ *then*

$$\lim_{x_3\to 0}\left|\left|\int_{\partial H}\frac{\partial G}{\partial x_3}f(y)\,\mathrm{d}s(y) - f(x)\right|\right| = 0 \tag{C.15}$$

Proof. We follow the arguments given by Folland[127]. First we note that for $x_3 \neq 0$

$$\begin{aligned}\int_{\partial H} \frac{\partial G}{\partial x_3}\,\mathrm{d}s(y) &= \int_{\partial H} \frac{x_3}{2\pi|x-y|^3}\,\mathrm{d}s(y) \\ &= \frac{1}{x_3^2}\int_{\partial H} \frac{1}{2\pi\left(\left(\frac{x_1-y_1}{x_3}\right)^2+\left(\frac{x_2-y_2}{x_3}\right)^2+1\right)^{3/2}}\,\mathrm{d}s(y) \\ &= 1, \end{aligned} \tag{C.16}$$

and so

$$\begin{aligned}\int_{\partial H} \frac{\partial G}{\partial x_3} f(y)\,\mathrm{d}s(y) - f(x) &= \int_{\partial H} \frac{x_3}{2\pi|x-y|^3}(f(y)-f(x))\,\mathrm{d}s(y) \\ &= \frac{1}{x_3^2}\int_{\partial H} \frac{f(y_1,y_2)-f(x_1,x_2)}{2\pi\left(\left(\frac{x_1-y_1}{x_3}\right)^2+\left(\frac{x_2-y_2}{x_3}\right)^2+1\right)^{3/2}}\,\mathrm{d}s(y)\end{aligned}$$

with a change of integration variables $y_1 \to x_3y_1 + x_1$ and $y_2 \to x_3y_2 + x_2$ we get

$$= \int_{\partial H} \frac{f(x_1+x_3y_1, x_2+x_3y_2)-f(x_1,x_2)}{2\pi\left(y_1^2+y_2^2+1\right)^{3/2}}\,\mathrm{d}s(y).$$

Using a generalization of the Minkowski or triangle inequality for integrals[195]

$$\left|\left|\int h(x,y)\,\mathrm{d}y\right|\right| \leq \int ||h(x,y)||\,\mathrm{d}y,$$

and writing f_{x_3} for $f(x_1 + x_3y_1, x_2 + x_3y_2)$,

$$\left|\left|\int_{\partial H} \frac{f_{x_3}-f(x_1,x_2)}{2\pi\left(y_1^2+y_2^2+1\right)^{3/2}}\,\mathrm{d}s(y)\right|\right| \leq \int_{\partial H} \left|\left|\frac{f_{x_3}-f(x_1,x_2)}{2\pi\left(y_1^2+y_2^2+1\right)^{3/2}}\right|\right| \mathrm{d}s(y)$$

and so,

$$\left|\left|\int_{\partial H} \frac{\partial G}{\partial x_3} f(y)\,\mathrm{d}s(y) - f(x)\right|\right| \leq \int_{\partial H} \frac{||f_{x_3}-f||}{2\pi\left(y_1^2+y_2^2+1\right)^{3/2}}\,\mathrm{d}s(y). \tag{C.17}$$

Choose a contineous and compactly supported function g such that $||f-g|| < \epsilon/3$. Then also $||f_{x_3} - g_{x_3}|| < \epsilon/3$. Because they are contineous and supported on a common

compact set $||g - g_{x_3}|| \to 0$ for $x_3 \to 0$ and so also $||g - g_{x_3}|| < \epsilon/3$ for x_3 sufficiently small and so

$$||f_{x_3} - f|| \leq ||f_{x_3} - g_{x_3}|| + ||g_{x_3} - g|| + ||g - f|| \leq \epsilon.$$

So $\lim_{x_3 \to 0} ||f_{x_3} - f|| \to 0$. Using this and the dominated convergence theorem in (C.17), we obtain the desired result. □

C.2 Standard Results

Most of these have been taken directly from text books and the literature with adjustments in notation. They are provided for convenience . The proofs have been omitted.

The double layer potential, for a constant moment $\phi = c$ is of considerable interest. It is proposition 3.19 in Folland[127] or example 6.16 of Kress[105]

Proposition C.5.

$$\int_{\partial D} \frac{\partial \Phi(x,y)}{\partial \nu(y)} \, \mathrm{d}s(y) = \begin{cases} -1 & \textit{for} \quad x \in D \\ -\frac{1}{2} & \textit{for} \quad x \in \partial D \\ 0 & \textit{for} \quad x \notin \overline{D} \end{cases}$$

Corollary C.6. *For $x \in \overline{H}$ it holds that*

$$\int_{\partial D} \frac{\partial G(x,y)}{\partial \nu(y)} \, \mathrm{d}s(y) = \begin{cases} -1 & \textit{for} \quad x \in D \\ -\frac{1}{2} & \textit{for} \quad x \in \partial D \\ 0 & \textit{for} \quad x \notin \overline{D} \end{cases}$$

The fact that a weakly singular kernel gives rise to a compact operator is of considerable importance in this thesis. The definition and theorem given here is from Kress[105], in particular his theorem 2.23 and discussion following 4.11 (Lax).

On a surface ∂D in $\mathbb{R}^m$, a kernel K is said to be *weakly singular* if it is defined and continuous for all $x, y \in \partial D, x \neq y$, and there exist positive constants M and $\alpha \in (0, m-1]$ such that

$$|K(x,y)| \leq M|x-y|^{\alpha - m + 1}, \quad x, y \in \partial D, \; x \neq y.$$

Lemma C.7. *The integral operator with continuous or weakly singular kernel is a compact operator on $L^2(\partial D)$ if ∂D is of class C^1*

Proof. (Sketch) Theorem 2.23 of Kress[105] provides the proof on $C(\partial D)$. In this, the compactness of operators with continuous kernels is established with the Arzelà-Ascoli theorem. Operators with weakly singular kernels are shown to be compact by constructing a sequence of operators with continuous kernels that is norm convergent to the operator with weakly singular kernel. To apply these results to operators on the surface ∂D, it is observed that since the surface ∂D is of class C^1, the normal vector is continuous on ∂D. This means that sufficiently small regions of ∂D can be projected bijectively onto the tangent plane to ∂D, this is then used to establish the existence of the integral operator as an improper integral.

Kress extends the case to $L^2(\partial D)$ on page 44 using the Lax theorem.

□

An operator $A : X \to Y$ is Fredholm if the im$A$ is closed in Y and the subspaces ker$A$ and $Y/$im$A$ are finite-dimensional. The difference between the dimension of the latter two subspaces is the index of the operator and is a measure how well the spaces of domain and image 'fit'. For instance any linear map between two finite spaces, $A : \mathbb{C}^n \to \mathbb{C}^m$ has index$(A) = n - m$. Fredholm operators with zero index are of interest because the Fredholm alternative (see for instance Mclean[196] theorem 2.27 or Kress[105] theorem 4.15) facilitates the study the solvability of the equation $A\phi = f$. Compact operators are famously *not* Fredholm because the image of a compact operator is not closed. On the other hand, the compact modification of the identity is Fredholm with index zero. Here is the version of Mclean[196] theorem 2.22.

Lemma C.8. *If $A = I + K$, where $K : X \to X$ is compact, then $A : X \to X$ is Fredholm with index zero.*

In the proof of theorem 3.1 we rely on theorem 2 on page 260 of [126]. There this theorem is initially stated for bounded domains but its validity is extended to the case of an unbounded half space, the epigraph of a Lipschitz function, in subsequent sections of the book.

The domain $\Omega \subset R^{n+1}$ is defined by $\Omega = \{(x,t) \in R^{n+1} : t \in R, x \in R^n, t > \phi(x)\}$, and where $\phi(x)$ is a real-valued Lipschitz function on the whole of $\mathbb{R}^n$, satisfying

$||\nabla\phi||_\infty < L < \infty$. With that, the operator K is defined here as the integral operator on $\mathbb{R}^n$ with the kernel

$$k(x,y) = \frac{\phi(x) - \phi(y) - (x-y)\cdot\nabla\phi(y)}{\omega_n \left(|x-y|^2 + (\phi(x)-\phi(y))^2\right)^{(n+1)/2}}.$$

Using the notation, $X = (x,t)$, $Y = (y,t)$, the operator $\mathcal{K}$ is defined as the double layer potential operator (cf (3.15)),

$$\mathcal{K}f(X) = \frac{1}{\omega_n}\int_{\partial\Omega} \frac{(Y-X)\cdot N(Y)}{|Y-X|^{n+1}} f(Y) d\sigma(Y), \quad X \in \Omega, Y \in \partial\Omega.$$

The cone $\Gamma(x)$ and maximal function $u^\star(x)$ are as defined for $x \in \partial\Omega$ by the formulae on page 40 with $M > L/\sqrt{1+L^2}$ in the definition of $\Gamma(x)$. With that the theorem is stated as:

Lemma C.9. *Let Ω be a Lipschitz domain. Then the "boundary" operator*

$$\frac{1}{2}I + K \ : \ L^2(\partial\Omega, \mathrm{d}\sigma) \to L^2(\partial\Omega, \mathrm{d}\sigma) \tag{C.18}$$

is an isomorphism. If g lies in $L^2(\partial\Omega, \mathrm{d}\sigma)$, then the harmonic function $u_0 = \mathcal{K}(I/2 + K)^{-1}g$ is the unique solution of the following Dirichlet problem: find $u_0 \in C^2(\Omega)$ such that $u^* \in L^2(\partial\Omega)$,

$$\nabla^2 u_0 = 0 \qquad \text{in } \Omega \tag{C.19}$$

and

$$\lim_{y\to x, y\in\Gamma(x)} u_0(y) = g(x) \qquad \text{for almost all } x \in \partial\Omega. \tag{C.20}$$

Remark: We note that for the case of interest in this thesis in general and theorem 3.1 in particular, we have, trivially, $K = 0$.