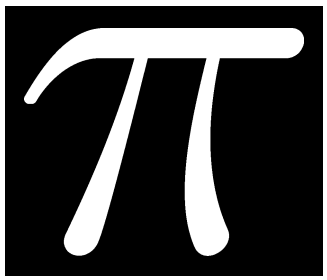


The
Pacific
Institute
for the Mathematical Sciences

http://www.pims.math.ca
pims@pims.math.ca

Proceedings of the Second PIMS Industrial Problem Solving Workshop

PIMS IPS 2

Co-sponsored by:

The National Science and Engineering Council of Canada

The Alberta Science and Research Authority

The British Columbia Information, Science and Technology Agency

Editor: M. Lamoureuax, University of Calgary

Foreword by the PIMS Director

In its second year of operation, the Pacific Institute for the Mathematical Sciences (PIMS) has evolved into a dynamic, flexible, and creative organization that constantly identifies new challenges and opportunities for its scientists and for its industrial partners. As part of its mandate, PIMS is committed to connecting the considerable expertise in mathematical sciences research at the seven Alberta and BC universities with the public and private sectors. In consultations with the private sector, and thanks to the financial support of NSERC, through its Research Partnership Program, the Alberta government through the Alberta Science and Research Agency and the British Columbia government through the Information, Science and Technology Agency, PIMS has designed a number of programs to facilitate university–industry interaction. They are:

- The Industrial Problem Solving Program.
- The Thematic Outreach Program.
- The Collaborative Industrial Program.

This series of initiatives is designed to bring together academic scientists, graduate students, and industrial researchers in the mathematical sciences to investigate mathematical problems arising in the industrial sector. In all its activities and programs[1], PIMS places a high priority on the development of young researchers, therefore each one of these programs has a component that contributes in its own way to:

- The Industrial Mathematics Training Program.

At the heart of the Industrial Problem Solving Program are the intensive week-long workshops initiated with the aim of creating a mutually beneficial link between researchers in industry and academic mathematicians. Researchers with industrial and commercial concerns are invited to present one of their current technical problems for study in working sessions with leading specialists from the academic community. Problems may come from a wide variety of subject areas, but should be amenable to mathematical modeling and analysis.

In August 1997, PIMS held its first Industrial Problem Solving Workshop at UBC. About 80 scientists came together to work on problems presented by *Powertech Inc.*, *Petro-Canada*, *McMillan Bloedel Inc.*, *Kinetic Sciences*, and *The BC Cancer Agency*. These problems ranged from devising techniques for vertical seismic profiling to more efficient fingerprinting identification and to modeling epithelial lesions in lung cancer. A very high degree of satisfaction by both the industrial and academic partners in this program was expressed.

Following the success of that first workshop, PIMS decided to make it an annual event and a part of a larger *Annual Forum on Industrial Mathematics*. Every year the forum will start with an extensive training camp in the modern methods of applied mathematics. In particular, topics like mathematical modeling, numerical methods and computer simulations would help in preparing the participants for the IPS workshop. The first *PIMS Graduate Industrial Mathematics Modeling Camp* was held at Simon Fraser University, May 24 - May 29, 1998. Forty Canadian graduate and near-graduate students came to Simon Fraser University to work with five mentors on various techniques for modeling industrial problems. The students came from 12 universities across Canada: SFU, UBC, UCalgary, UAlberta, UVictoria, UToronto, McGill, UWO, UWaterloo, UQuébec at Sherbrooke, UManitoba and Queen's University. Following the camp, PIMS arranged for all the students and most mentors to join the second *PIMS Industrial Problem Solving Workshop* (PIMS-IPS) which was held this year in Calgary.

From June 1 to June 5, 1998, about 100 scientists came together at the University of Calgary for the second PIMS-IPS. They worked on problems presented by *The Boeing Corporation*, *The Computer Modeling Group*, *The Geomech Project*, *ITRES Research Ltd*, *Powertech Labs* and *VisionSmart*. The scientists came away from the workshop with the feeling that this year's event was

[1] A detailed discussion of the PIMS activities can be found in the PIMS annual report for 1997–1998 or on the webpage http://www.pims.math.ca

even better than last year's. The following report describes their work. We are indebted to Dr. M. Lamoureux from the Department of Mathematics and Statistics at the University of Calgary for collecting and editing these proceedings.

Workshops like this one are only possible through the very hard and selfless work of many PIMS scientists. The success of this year's workshop is due to the efforts and determination of Drs. R. Westbrook and D. Calistrate from the Mathematics and Statistics Department of the University of Calgary as well as M. Paulhus and Dr. H. Huang, the PIMS industrial facilitators for Alberta and British Columbia, respectively. I want to express my appreciation to them and to all the colleagues who have helped along the way.

The next Industrial Mathematics Modeling Camp will be held in the Spring of 1999 at the University of Alberta, to be followed by the 3rd Industrial Problem Solving Workshop to be held at the University of Victoria. We will be looking forward for the continuation of this story of commitment, dedication and success.

Dr. Nassif Ghoussoub, Director
Pacific Institute for the Mathematical Sciences

Editor's Preface

In June 1998, the second Industrial Problem Solving Workshop of the Pacific Institute for the Mathematical Sciences was held at the University of Calgary, under the moniker PIMSIPS 2. The objective was to build on the success of the first PIMSIPS Workshop held in 1997, bringing together mathematical scientists to collaborate on six problems posed by local industries. PIMSIPS 2 involved some 100 participants from North America and Europe, including seven industrial representatives and about 40 graduate students.

Format of the workshop

The format of this second workshop followed closely the organization of the first PIMSIPS Workshop in Vancouver, which was based on a formula developed by the Oxford Study group for a five–day workshop. On the morning of the first day of the workshop, the participants are introduced to the industrial problems by a sequence of six short lectures presented by industry scientists who describe in some detail their research problems. In the afternoon, the participants break into groups to work on the individual problems, beginning with further questions to the industrial scientists, collection of notes and data, and exchange of ideas for possible directions of study. The research then continues over the course of the workshop. The groups are self-selected, usually based on the particular interests and skills of the participants and how these relate to the given problems. The dynamics of working in these groups were considerably enhanced by the presence of three "Workshop Specialists," who had been invited to assist in the workshop because of their considerable experience in this kind of activity. They provided the drive and focus to keep the participants on track as they applied their mathematical skills to the problems at hand. On the final day of the workshop, the results of the research efforts were presented in lecture format to the assembled participants, followed by a brief response from the industry scientists.

A couple of changes were instituted in this year's activities. First, we had no expository lectures scheduled, as is commonly done in the Oxford Study Groups; the work generated by the industrial problems was more than enough to keep the participants active and occupied. Second, in the week before the workshop, the PIMS Graduate Industrial Mathematics Modeling Camp was held at Simon Fraser University, to prepare the graduate student participants of PIMSIPS 2 for the activities of the workshop. Conveniently, many of these graduate students were able to attend the annual meeting of the Canadian Applied Mathematics Society in Vancouver, held in the weekend between the training camp and the PIMSIPS Workshop. This training camp was a very successful addition to the workshop, providing us a strong group of highly motivated students primed for action.

Acknowledgments

An operation of this magnitude can be successfully undertaken only with the efforts of a large number of dedicated, hardworking individuals and the support a variety of organizations. In addition to the roughly 100 research participants, the PIMSIPS Workshop owes a particular debt of gratitude to the following groups for their contributions to the activities of the meeting:

- Claude Laflamme, PIMS Site Director for Calgary, and Rex Westbrook, Workshop Organizer, for their work in bringing this workshop to fruition at the University of Calgary, and ensuring its smooth running over the course of the week. Without their tireless efforts, this workshop could not have taken place;

- Dan Calistrate, Huaxiong Huang, and Marc Paulhus, the PIMS Industrial Facilitators for Calgary and Vancouver, who developed a variety of industrial connections into concrete proposals for study at the workshop, facilitated the operations of the workshop, and followed up with the companies to ensure a proper recording of their research problems in these proceedings;

- Gordon Sick, of the Faculty of Management at the University of Calgary, who led the Industrial Coordinators to a key contact for one of the industrial problems;

- the three Industrial Problem Solving Workshop Specialists, Chris Budd of the University of Bath, Alistair Fitt of the University of Southampton, and Rachel Kuske of the Institute for Applied Mathematics at the University of Minnesota, who added their considerable expertise to the activities of the workshop;
- the eight principal authors of the workshop reports in these proceeding, who devoted considerable time in compiling the various research efforts of the week and delivering them to the editor in a coherent form. They are:
  - Daya Gaur – the Boeing problem #1,
  - Marc Paulhus – the Boeing problem #2,
  - Brian Seymour – the CMG problem,
  - Paul Webster – the Geomechanics Project problem,
  - Chris Budd and John Stockie – the ITRES problem,
  - Rex Westbrook – the Powertech problem,
  - Jonathan Samuel – the VisionSmart problem;
- Nassif Ghoussoub and Marc Paulhus, for their assistance in editing the workshop proceedings and the final orchestration of the printing process;
- Mike Boorman, Dean of the Faculty of Science at the University of Calgary, for his support for both this large workshop and more generally for the activities of PIMS, and for his concluding address to the participants at the end of the workshop week;
- the Department of Mathematics and Statistics at the University of Calgary, and in particular the Head and Assistant Head, Ernest Enns and Marguerite Fenyvesi, for ensuring the space and facilities to conduct such a large operation. As well, we thank the support staff and graduate students of the Department, especially Jan Cerny, Joanne Longworth, Arunas Salkauskas, Satoshi Tomoda and Mark Wrubleski, who provided a variety of logistic support and computer services;
- a special thanks to our industrial contributors
  - The Boeing Corporation
  - The Computer Modeling Group Limited
  - The Geomechanics Project
    (Baker Altas, PanCanadian Petroleum, Petro-Canada Oil and Gas, and Talisman Energy)
  - ITRES Research Limited
  - Powertech Labs Inc.
  - VisionSmart Inc.

Structure of this monograph

This monograph records the activities of PIMSIPS 2. It begins with a listing of the six problems brought forward by the industrial contributors, each culminating in a "posed problem" for study over the course of the workshop. Following this listing is a report for each of the problems, including two reports on the Boeing problem, detailing the results obtained during the workshop. Using the conventions of the first PIMSIPS Workshop, we have organized these seven reports by the alphabetical order of the names of the industrial contributors. The final report is a record on the details of how to run an industrial study group, originally prepared by Tim Myers and Barbera van de Fliert for the PIMSIPS 1 proceedings, and updated here based on our experience with these workshops. The monograph concludes with a reasonably complete list of participants for the workshop and their coordinates – my apologies for the inevitable errors and omissions there within.

Michael Lamoureux, Editor
Dept. of Mathematics and Statistics
University of Calgary

Contents

Foreword by the PIMS Director . . . iii
Editor's Preface . . . v
Format of the workshop . . . v
Acknowledgments . . . v
Structure of this monograph . . . vi
PIMSLIPS . . . 1
Problem summaries and contact people . . . 2

1 Maintaining Excess Capacity **5**
1.1 Abstract . . . 5
1.2 Introduction . . . 5
1.3 Modeling the future demand . . . 6
1.4 Queuing Model . . . 8
1.5 Using option pricing . . . 9
1.5.1 Real options . . . 10
1.5.2 Valuation of the plant using the Brennan-Schwartz model . . . 12
1.6 Conclusion . . . 13
1.7 References . . . 13

2 A Renewal Model for Sales **15**
2.1 Introduction . . . 15
2.2 Renewal Models for Future Sales . . . 15
2.3 Analytical Solution to a Simpler Problem . . . 20
2.4 The Inventory Problem . . . 21
2.5 References . . . 22

3 Petroleum Reservoir Simulation **23**
3.1 Introduction . . . 23
3.2 The Physical Problem . . . 23
3.3 Potential Problems . . . 24
3.4 Model Equations . . . 24
3.5 Approaches for Finding Exact Solutions . . . 25
3.6 Radially Symmetric Flow . . . 26
3.7 Results and Discussion . . . 28
3.8 Stability and Well-posedness . . . 28
3.8.1 Stability of a Paradigm Shock Solution . . . 30
3.8.2 Stability for the Oil recovery problem . . . 32
3.9 References . . . 37

4 On Seismic Imaging **39**
4.1 Introduction . . . 39
4.2 Background Of The Problem . . . 40
4.2.1 Mathematical foundations . . . 40
4.2.2 A physical model of propagation of a seismic pulse . . . 41

4.2.3 Single horizontal reflector . . . 41
4.2.4 A single dipping reflector . . . 42
4.2.5 Vertical seismic profiling . . . 42
4.3 Theoretical Approach . . . 43
4.3.1 Constant velocity field . . . 43
4.3.2 Linear velocity field . . . 44
4.3.3 Direct signal travel time . . . 44
4.4 Numerical Approach . . . 45
4.4.1 Method of solution . . . 45
4.5 Conclusions . . . 46
4.6 References . . . 47

5 Trip Wire Detection **49**
5.1 Introduction. . . . 49
5.2 Detailed Problem Description. . . . 50
5.3 The Radon Transform . . . 51
5.3.1 What is a wire? . . . 51
5.3.2 The Radon Transform. . . . 52
5.3.3 The algorithm. . . . 53
5.3.4 Approaches to thresholding. . . . 54
5.3.5 Results. . . . 55
5.3.6 Speeding up the algorithm. . . . 59
5.4 Other Methods. . . . 60
5.5 Summary and Future Directions. . . . 60
5.5.1 Main program. . . . 61
5.6 References . . . 62

6 Torsion in Multistrand Cables **63**
6.1 The Problem . . . 63
6.2 Solution Strategy . . . 64
6.3 Changes in the Theoretical Model . . . 65
6.4 Numerical experiments . . . 66
6.5 Conclusions . . . 67
6.6 References . . . 67

7 Detection of Cracks **69**
7.1 Introduction . . . 69
7.2 Data Analysis . . . 70
7.3 Suggested approaches . . . 73
7.3.1 Beam width . . . 73
7.3.2 Registration . . . 73
7.3.3 Hairline cracks . . . 74
7.3.4 Vary beam wavelength . . . 74
7.3.5 Imaging techniques . . . 74
7.4 Literature Review . . . 74
7.5 References . . . 75

8 Organization of a Workshop **77**
8.1 Introduction . . . 77
8.2 How to contact industry . . . 78
8.3 Screening the problems . . . 78
8.4 Report/proceedings . . . 78
8.5 Organization . . . 78
8.6 References . . . 79
8.7 A selection of problems investigated at recent meetings at PIMS . . . 80

List of Figures

1.1 Time Series Plot of *log* of Demand 7
1.2 Time series plot of ϵ_t 8
1.3 ACF plot of ϵ_t 9
1.4 qq-plot of ϵ_t 10
1.5 The queuing model 11
1.6 Option price for the planes 14

2.1 Examples of Gamma density functions 17
2.2 Simulations of orders 18
2.3 Simulation of orders with drift 19
2.4 Tree of possible outcomes 21

3.1 Typical reservoir simulation geometry. 24
3.2 Rectangular grid orientations. 24
3.3 Dependence of key functions on M, α, and β. 25
3.4 Characteristic curves for the mixed convex/concave f case. 28
3.5 Saturation S versus radial variable r 29
3.6 Pressure p versus radial variable r 30
3.7 Pressure field p along a ray 31
3.8 Saturation field S along a ray 32
3.9 Pressure field p in the entire solution domain 33
3.10 Saturation field S in the entire solution domain 34
3.11 Darcy porous medium. 35
3.12 Hele-Shaw. 36

4.1 Isochronous curves for the three profiles. 46

5.1 Two trip wires 50
5.2 ρ and θ in the Radon transform. 52
5.3 Test image and transform 52
5.4 Sample inversion of an image 55
5.5 The four "idealised wire" images. 57
5.6 The top 200×461 portion of the image 58
5.7 Edge detection images 58
5.8 Inverse transform after edge detection 59

6.1 Cross-section of a multistrand cable 63
6.2 Single layer and angle of strands 64
6.3 Single layer of strands 66
6.4 Torque vs Force in gwire 18 67
6.5 Torque vs Force in Peace 68

7.1 Intensity data 71
7.2 Scale factors 72

7.3 Scale factors 73

List of Tables

2.1 Production levels of a hypothetical assembly plant. 22

4.1 Summary of numerical results. 47

5.1 Linear features in an image 56
5.2 Linear features in an image 56
5.3 Maximum scaled Radon transform 57

PIMSLIPS

While endevouring to sip from the Pierian spring, we occasionally slip in the surrounding mud. Below are some slips of the tongue, or otherwise interesting comments from the workshop that we record for posterity.

- "... we get a sollow-holid cylinder ..." (Rex Westbrook, PIMS-IPS Organizer)
- "Correction, a hollow-solid cylinder ..." (ditto)
- "How many feet in a Newton?" (Alistair Fitz, oil resevoir problem)
- "Those questions are beyond my pay code." (Greg Robel, Boeing Corporation)
- "This is how it works in a nutshell..." (David Lyder, eggshell problem)
- "This number is very suspicious, because it is half of π (Huaxiong Huang, PIMS)
- "Claim $z' = c$ iff $z'' = 0$" (Anonymous)
- "We found that 10 is a little bigger than 9." (Rachel Kushe, eggshell problem)
- "Going from 2D to 3D adds a whole new dimension to the problem." (Jeff Grossman, Geomech problem)
- "Never trust a man from industry." (Michele Titcombe)
  "How about a woman from industry?" (Alistair Fitt)
- "Something that is elliptical is not necessarily an ellipse." (Brian Seymour)
- "... since we live in a somewhat three dimensional world..." (Michael Slawinski, Geomech)

Problem Summaries and Contact People

- **Boeing Corporation** *Maintaining excess capacity*

  Boeing manufacture planes of high value in very low production volumes, and typically sees large fluctuations in demand. Massive changes in production capacity to match these changes in demand can be prohibitively expensive in this industry, so an alternate method of managing production levels is sought.

  Problem Posed
  Suppose that the current demand for a firm's output is low. What is an optimal strategy for maintaining excess capacity in order to be able to profit from an unexpected surge in demand?

  Contact person: Dr. Greg Robel, `gregory.f.robel@boeing.com`

- **The Computer Modeling Group Ltd.** *Petroleum reservoir models*

  Often only two partial differential equations are required to model the flow of two fluids in an underground petroleum reservoir. One of the equations is a parabolic evolution equation for the fluid pressures within the reservoir. The other is a nonlinear hyperbolic conservation law that describes the movement of the fluids. The equations are coupled through their coefficients, in that the pressure solution of one determines the velocity field needed for the other, while subsequent movement of the fluids, and the withdrawal or injection of fluids at boundaries, alters the pressure field. The solution of this problem in a single space dimension is well understood. However, there are analytical and numerical difficulties in solving these equations in higher dimensions, even in two dimensions.

  Problem Posed
  It would be very useful to have a way of obtaining high accuracy solutions to these problems so as to guide the development of numerical schemes for more complicated situations. In particular, problems due to grid orientation in the numerical schemes could be investigated. Either analytical solutions, or analytical solutions enhanced with auxiliary numerical computations, would be of great utility.

  Contact person: Dr. Peter Sammon, `peter@cmgroup.com`

- **The Geomechanics Project** *Seismic imaging*

  The most fundamental measurement in exploration seismology is the travel time between given sources and receivers. Consider an infinite medium of constant velocity, V, containing a single scattering point, Q. If the measured travel time between source and receiver of the scattered signal is T, the distance traveled is VT. Hence the locus of all possible scattering points is an ellipse with the source and receiver being the foci of the ellipse. If we possess several source-receiver configurations we can uniquely determine the location of the scattering point, Q, which would correspond to intersection points. This method constitutes a core of an element of seismic data processing known as Kirchhoff migration.

  In reality, speed is not constant for a given area and scattering/reflection is not limited to a single point. Important technical and economic decisions are based on locating sub surface structures, e.g.the apex of an anticline.

  Problem Posed
  The purpose of this project is to investigate the range of solutions based on various assumptions for speedfields.

  Contact person: Dr. Michael Slawinski, `mslawins@enme.ucalgary.ca`

- **ITRES Research Limited** *Landmine tripwire detection*

  Trip wires are strings, lines, or wires that when "tripped" assist the detonation of land mines. The trip wires are camouflaged in a wide variety of backgrounds such as vegetation, across roads, etc. Currently, there exists no reliable standoff means to automatically (and safely) detect trip wires.

Itres Research Limited, in conjunction with Defense Research Establishment Suffield (DRES), is developing passive and active imaging systems to investigate stand-off trip wire detection. The imaging systems produce digital images of scenes containing images of trip wires against a variety of backgrounds. The trip wire images are linear and span the image frame at random orientations.

The methodology recently employed by Itres to locate and identify linear features in digital images involves convoluting the digital image with a gradient operator to highlight edges. The processed image is then segmented and a Hough transform is applied to locate linear features. This methodology, however, is computationally inefficient and possesses a high false alarm rate.

Problem Posed
The challenge is to develop an innovative and robust approach for detecting linear features (both continuous and disjointed lines spanning the image frame at random orientations) in "cluttered" digital images. The approach must possess a low false alarm rate and can lend itself to be implementable in real-time. Application of the approach developed may go beyond field land mine detection to areas such as detecting power/telephone lines from helicopters; such lines are considered flying hazards.

Contact person: Dr. Avygdor Moise, `avy@itres.com`

- **PowerTech** *Tensile strength of multistrand cables*

  Multistrand cables are used widely in industry. Overhead electrical conductors, wire ropes and suspension cables in mine hoists are typical examples. The simplest construction involves a straight core wire surrounded by concentric layers of strands. All strands in a given layer remain twisted about the cable axis at a fixed angle known as the layer angle. Adjacent layers remain twisted in opposite directions to one another (clockwise or anticlockwise) in order to minimize torsion when the cable is loaded axially. Such a construction involves interaction between strands in contact.

  A simple model, developed recently by Lanteigne and Akhtar, predicts the maximum failing load and torsion of the cable using the geometry of the cable and the mechanical properties of the constituent wires as the input parameters. That model has made the simplifying assumption that no frictional interaction occurs between adjacent strands. Experiments carried out by Akhtar and Lanteigne, are in excellent agreement with the predictions of the model for multistrand conductors made with aluminum alloy strands. However, the model predicts torsion values for cables containing galvanized steel strands that deviate substantially from those measured experimentally. It has been concluded that interstrand frictional interaction does not occur between strands made of aluminum and aluminum alloys and that frictional interaction occurs when galvanized steel strands are used for the construction of the multistrand cable.

  Problem Posed
  Modify the existing model or develop an alternative model to predict torsion in multistrand cable involving interstrand friction.

  Contact person: Dr. Ainul Akhtar, `ainul.akhtar@bchydro.bc.ca`

- **VisionSmart** *Detection of cracks in egg shells*

  Candling eggs during packaging is the standard method for detecting cracks, cage marks, pinholes etc... This is a labour intensive process which can lead to unacceptably high rejection rates if the candlers are fatigued, the batch of eggs particularly poor, etc. To date, there has been considerable work done by Villetard Systems Group (Edmonton) on the problem of characterizing the surface features of eggs using the transmission qualities of the eggs when illuminated with a narrowly focused laser but no satisfactory algorithm with high detection rate and low false positive rate has been found.

  Problem Posed
  Some suggestions for further research in the characterization of egg shell structure by non-intrusive means include a purely statistical approach and the development of neural networks. In

all these methods the fundamental problem of dealing with highly complex, naturally occurring patters which possess considerable noise and can only be observed for a small amount of time must be addressed.

Contact person: Dr. David Lyder, `vsmart@planet.eon.com`

Chapter 1

An Optimal Strategy for Maintaining Excess Capacity

Yanez Ales,[1] Andreea Amarie,[2] John Anderies,[3] Brad Bart,[1] Daniel Chertok,[1] Daya Gaur,[1] Arvind Gupta,[1] Kelly Kwok,[3] Jian Liu,[3] Markus Orasch,[4] Greg Robel,[5] Gordon Sick,[6] Mohammadreza Simchi,[6] James Timourian,[2] Ryan W. Tse[3]

Report compiled by Daya Gaur

1.1 Abstract

Boeing is a manufacturing industry with very low production volumes of very large units. As such, they experience huge fluctuations in demand. A standard inventory model dictates massive changes in production capacity as demand varies. However all such models assume a continuous production stream. In this report we investigate the following question whether such a model is sensible in a problem of such large scale granularity. We describe a combination of stochastic, financial and simulation models to model the production of airplanes. A preliminary simulation of the model is also presented.

1.2 Introduction

After the merger with McDonnell Douglas and the acquisition of the defense and space units of Rockwell, Boeing is the largest airline manufacturer in the world, holding 60% of the market. Boeing traditionally experiences wide fluctuation in the plane demand. However, they have good models for long-term projection of future demand. This raises the following question: given projections about the future demand, what should be the schedule for building airplanes so as to maximize profit? The primary cost associated with any scheduling system is the cost of the inventory. Currently, Boeing follows the *Just In Time* production model (zero inventory), meaning that planes are only constructed as ordered.

The objective of this study is to make a case that by maintaining the inventory of partially built planes, both Boeing and the customer can profit: Boeing by minimizing the delivery schedules (which in turn translates into increased revenue) and the customer by not having to wait for up to two years (to start a new route, for example). By having an inventory Boeing would also not face

[1]Simon Fraser University
[2]University of Alberta
[3]University of British Columbia
[4]Carleton University
[5]Boeing Corporation
[6]University of Calgary

the problem of having to continuously hire and fire employees. This is a particularly aggravating problem for Boeing since they employ highly skilled labour force which itself is a scarce commodity.

Although Boeing is part of a manufacturing industry, its volume of production is insignificant compared to that in other manufacturing industries such as the automobile industry. Given the high level of granularity in Boeing's production, the standard models of job shop, flow lines, transfer lines, machining systems etc. Reference [3] cannot be applied directly to model this production of airplanes. The models cited above assume a high volume of production and a continuous demand for the product but neither of these assumptions is valid in Boeing's case.

In the airplane market, typically the time required to produce one plane is 24 months. Customers are willing to pay more for the planes delivered earlier as this gives them financial flexibility in decision making. Hence, were Boeing to keep an inventory of partially built planes they could sell these planes at a higher price and generate more revenue. For the customer the extra money is a strategic investment to have an edge over the competitors (for instance these competitors could not introduce routes in a shorter time frame). In financial terms, this extra worth can be viewed as an *option* to buy a plane by a specific date.

Since the demand for planes is unpredictable, stored planes permit Boeing to capitalize on large upswings in demand. In terms of practicality, this is an easy scheme to implement since the inexpensive, time-consuming steps of building planes are typically done before the expensive, faster steps. For example the hull construction, piping and wiring, are all done before the engines, cockpit and interior decoration are added. We will be analyzing a production system wherein Boeing stores partially built planes at different levels of completion.

The following questions were addressed in some detail at the Second PIMS Industrial Problem Solving Workshop held in Calgary from June 1-5, 1998.

1. What is a reasonable way to model the projected future demand of a plane model?
2. What kind of production model would be best suited to modeling an airplane assembly line? Is it possible to simulate such a system?
3. How should the planes be priced based on the delivery date?

Section 1.3 shows that the future demand of the planes could be modeled as a Wiener process. This conclusion was reached by doing a time series analysis of the projected demand data. Section 1.4 describes the queuing model which is used to maintain the inventory and simulate the airplane production process. It also comments on the inconclusiveness of the preliminary simulation results. Section 1.5 describes how the pricing of the planes is to be done based on the delivery date. The idea is to sell the customer an option to buy a plane instead of selling a plane. Though in reality the option would be a *real option*, during our simulation, for the sake of simplicity we modeled it as *European option*.

Note that the data within this report are hypothetical, or from public sources, and no statement made herein reflects an official policy of the Boeing Corporation.

1.3 Modeling the future demand

Traditionally, it is believed that plane orders follow a regular business cycle. However, in this section we show that in fact the number of airplane orders, denoted Y_t, follows a geometric Brownian motion:

$$Y_t = e^{X_t} \qquad (1)$$

where X_t is a simple Brownian motion.

A series of data obtained from 1955 to 1997 was used in our study. Taking logarithm on both sides of equation 1 we obtain $logY_t = X_t$. We will show that the logarithm of the data follows a simple Brownian motion. Before delving into the details, we explain why geometric Brownian motion is expected instead of simple Brownian motion. The usual assumptions when one analyzes data are:

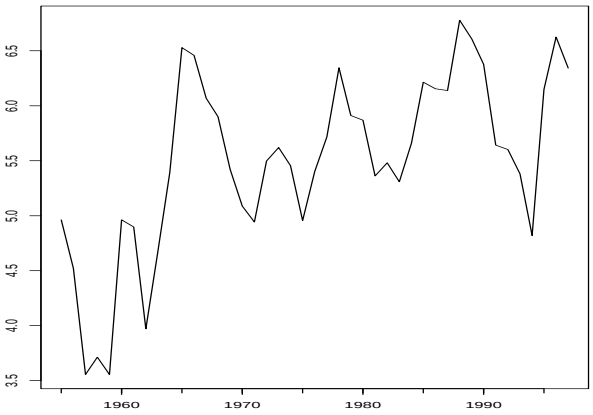


Figure 1.1: Time Series Plot of *log* of Demand

- Linearity of the model.
- Normality of the observations.
- The constancy of variance.
- Independence of observations.

The assumption about the normality of the data is violated by our data set. If any of these assumptions are not met by the data, then Tukey [6] suggests two alternatives: either a new analysis must be devised to meet the assumptions, or the data must be transformed to meet the usual assumptions. As it is usually easier to transform the data than to develop a new method of analysis, we transform the data of plane orders by a logarithm function, a common transformation used for count data. The time series plot of the transformed data is shown in Figure 1.3.

Brownian motion takes place in continuous time and continuous space. Our attempt to model the data proceeds by approximating it by a discrete process such as a random walk. We define Lag difference as $X_t - X_{t-1}$. The series of Lag difference on the 'log' data, $logY_t - logY_{t-1}$, shows randomness. We define $\epsilon_t = logY_t - logY_{t-1}$, and investigate the pattern formed. We argue that the time series plot of ϵ_t, which is shown in Figure1.3, shows a random pattern.

The sample autocorrelation function (ACF), which is the sample estimate of the autocorrelation function defined as $r(h) = Corr(\epsilon_{t+h}, \epsilon_t)$, shows that ϵ_{t+h} and ϵ_t are uncorrelated for $h = 1, 2, 3, ...$ This can be graphically seen from the ACF plot in Figure 1.3 with the data series $lnDdif = \epsilon_t$. All the sample estimates of autocorrelation, which are corresponding to the vertical bars in the graph, are not significantly different from zero (since all these bars are within the 2-standard-deviation horizontal lines) for $Lag\ (h)$ greater than zero.

The qq-plot in Figure 1.3, which plots the quantiles of a normal distribution against the quantiles of the data, suggests that the ϵ_t are normally distributed. This normality is further supported by the modified version of qq-plot for χ^2 distribution. The agreement of normality and zero correlation of ϵ_t implies independence. Now we reach the conclusion that X_t or $logY_t$ follows a Random Walk. The estimates of the mean and standard deviation of the ϵ_t are 0 and 0.3, respectively, where the zero mean is suggested by the t-test.

Finally, we point out that this by no means is conclusive evidence that the model of geometric Brownian motion is an optimal model for the series of data. Rather we have only established that

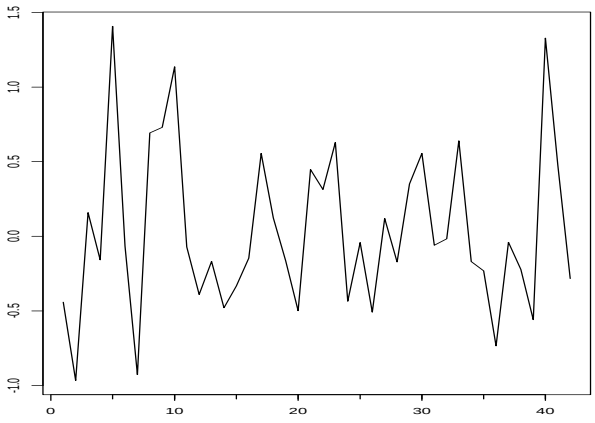


Figure 1.2: Time series plot of ϵ_t

there are no strong violations to the assumption that the input data follows the geometric Brownian motion.

1.4 Queuing Model

We argued in Section 1.2 that the production of planes cannot be modeled using job shop, assembly lines or flow line models of the manufacturing industry due to high level of granularity in the process. The model which most closely resembles our problem is that of multiple cell manufacturing system [3]. The idea is to replace a large job shop by specialized cells producing groups of job types that are identified as being similar in design and production process. Many issues arise in design, planning and operation of a cell model. In our simulation we restrict ourselves only to the operation of a cell model. Two important issues which arise are:

- *Buffer sizes:* Input and output of each cell is modeled using an inventory buffer. The maximum (and minimum) buffer sizes will have a drastic affect on the cost of the production. Our simulation tries to determine the revenue generated based on the Input rate and the Output rate. Input and the Output rates correspond to the buffer sizes.

- *Cell operation:* Each cell has to be designed so that it operates at maximum efficiency and maximum profit. In the simulation model we do not address this question. Standard treatment would be to model each cell as a Jackson queuing network [4].

We propose a combination of stochastic and simulation model to model Boeing's production process using multiple cells. Our model, shown in Figure 1.5, comprises two cells (specialized job shops) which produce planes that are 50% and 100% assembled respectively. Both the cells are producing planes at some specified rate. Planes which are produced by cell 1 are completed in cell 2. For the purpose of our simulation there is a server which keeps track of the incoming requests and services them based on the availability of the planes in cell 2.

The following facts were used in generating the simulation model.

- The demand is stochastic in nature that is geometric Brownian (see Section 1.3).

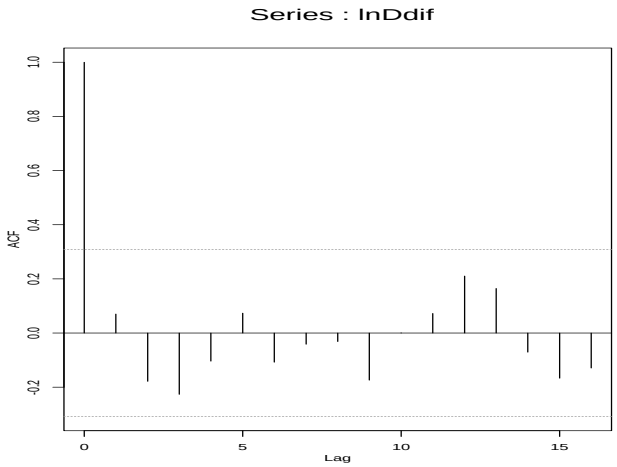


Figure 1.3: ACF plot of ϵ_t

- Costs associated with the production model are:
  - Cost of producing a plane to 50% or 100% completion state.
  - Cost of storing the plane in cell 1 and cell 2. These costs are dependent on the storage time.
- Revenue generated is *Sale Price* minus *Total Cost*, where the *Sale Price* is computed based on the delivery dates using the option pricing mechanism discussed in Section 1.5.
- Rate of production in the cells 1 and 2 has to closely follow the demand. Therefore the rates at which partially built planes are constructed was also modeled as an exponential function based on the input data.

An alpha version of a simulation was constructed to test our hypotheses. In order for our model to be successful, the simulation had to verify that our queuing strategy would generate more profit than Boeing's current strategy. Given the assumptions and the time frame, the simulation results were inconclusive. A better strategy for determining the rates at which planes are produced in cells 1 and 2 is needed for the simulation to be effective.

1.5 Using Option pricing to determine the worth of the airplane

In this section we will describe the option pricing mechanism used to determine the sale price of an airplane under the assumption that the option is a European option.

As was shown by Black, Scholes and Merton in 1973, the value c of a European call option with exercise price X and expiration date T is:

$$c = S_0 N(d_1) - X e^{-r_f T} N(d_2) \qquad (2)$$

where:

- S_0 is the spot price;

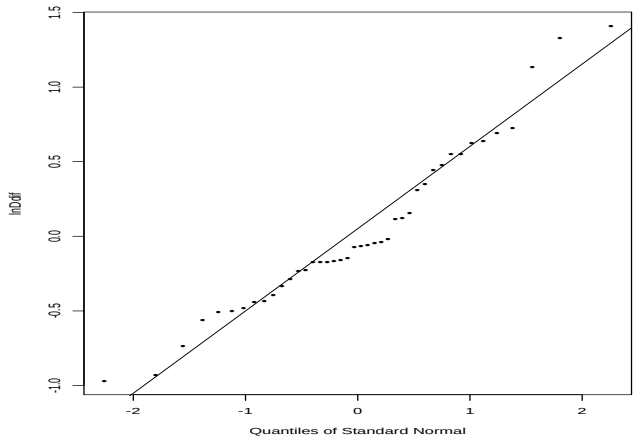


Figure 1.4: qq-plot of ϵ_t

- r_f is the interest rate;
- $d_1 = \frac{log\left(\frac{S_0}{X}\right)+\left(r_f+\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}$;
- $d_2 = d_1 - \sigma\sqrt{T}$;
- $N(d) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d} e^{-\frac{x^2}{2}} dx$.

The option price for an airplane based on equation 2 is shown in Figure 1.4, where the *x-axis* is the time line and the point (x, t) is the cost of buying a plane t days from the current day.

The assumptions made in plotting this graph are:

- The plant never closes. This means that there is a continuous production.
- No employees are laid off.
- The spot price of one plane is 75 million dollars.
- The exercise price of the option is 75 million dollars.
- The risk free interest rate is $r_f = 0.03$.
- The price volatility is $\sigma^2 = 0.3$.

1.5.1 Real options

In Boeing's case, the option of buying an airplane is in fact a *real option*. A real option is the flexibility to make decisions about real assets, decisions which can involve adoption, abandonment, exchange of one asset for another or modification of the operating characteristics.

In our specific problem we have the following four real options:

- to maintain or to add excess capacity (how much should we add?);

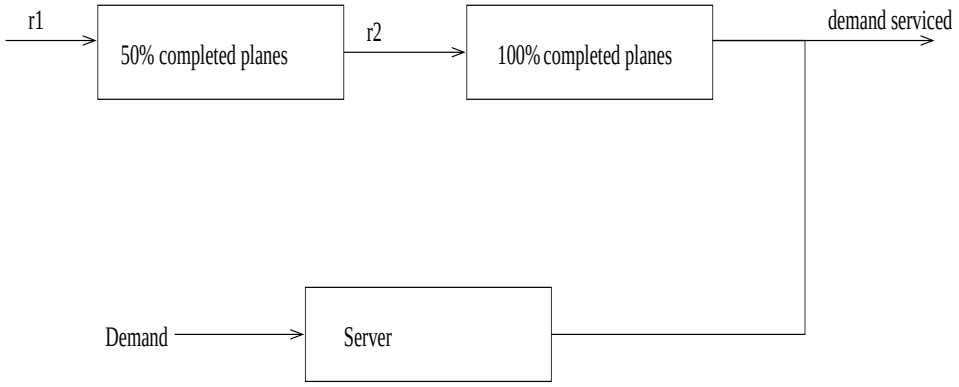


Figure 1.5: The queuing model

- to keep the plan operational;
- to mothball (we do not use the plant, but we still keep the people and machines inside it);
- to abandon, get rid of the plant.

For each of the real options outlined above we have to determine the option value and the exercise price. We adapt the technique of Brennan and Schwartz, outlined in the next section, to our domain.

The option price and the exercise date decisions depend on the structure of the stochastic demand. We will assume that the demand D follows the following diffusion process:

$$dD = \alpha\left(D,t\right)dt + \sigma\left(D,t\right)d\omega$$

where:

- α is the expected growth in demand, measured in orders per unit of time;
- σ is the annual standard deviation of underlying asset returns;
- dω is a Wiener process with zero drift and unit variance per unit time. That is, $\omega(t) - \omega(t-1)$ is normally distributed with expectation 0 and variance 1 and is independent of $\omega(\alpha) - \omega(\alpha - 1)$ for any time interval between $\alpha - 1$ and α that does not intersect the time interval between $t-1$ and t.

We first investigate the pure option to add capacity. Using the risk neutral interpretation (we look at expectation and ignore the risk premium without losing generality) and Itô's lemma we get the equation (similar to that in reference [1] obtained for the mine pricing model):

$$rW = W_D\left(\frac{E[dD]}{dt}\right) + \frac{1}{2}\sigma^2\left(D\right)W_{DD}$$

where:

- $W\left(D,t\right)$ is the value of the option to meet excess demand.

The left side of the above equation represents the required return per unit time of a risk-neutral investor for an investment in the option (see [5]). The right side represents the expected return to ownership of the option (per unit time), assuming that the underlying asset is priced by a risk-neutral investor (see the same as above)

The drift $\frac{1}{2}\sigma^2(D) W_{DD}$ is due to Itô's lemma.

The value of the option to abandon the plant plus the net cash flow will be now given by the equation:

$$\frac{1}{2}\sigma^2(D)\nu_{DD} + \left(\frac{E[dD]}{dt}\right)\nu_D - r\nu + pD = 0$$

where:

- pD represents the net profit (p is the net profit from the sale of one plane).

Using the information we obtained from Boeing's public website, and from our own time-series analysis, we actually found the values of $\sigma^2(D)$ and $\frac{E[dD]}{dt}$ and replaced them in the equations written above. In conclusion, these two second order Euler differential equations can now be solved.

For example, the general solution for the first equation is:

$$W(D) = \beta_1 D^{\gamma_1} + \beta_2 D^{\gamma_2}$$

where β_1 ,β_2 are easy to find and γ_1 ,γ_2 which will also appear in $\nu(D)$ can be determined from boundary conditions such as:

- value matching conditions:
$$W(D_L) = \nu(D_L) - E$$
$$\nu(D_H) = W(D_H) - I$$
where D_L is the lowest demand, D_H is the highest demand, E is the value of abandonment and I is the value of investment

- optimality conditions:
$$W(D_H) = \nu(D_H)$$
$$W(D_L) = \nu(D_L).$$

In Boeing's case, we think it is a good idea to use also the mothball decision.

1.5.2 Valuation of the plant using the Brennan-Schwartz model

The approach outlined in this section is described by Brennan and Schwartz in reference [1]. Consider the value of an airplane production plant H. We assume that the price of the airplane produced at this plant is not subject to fluctuations. The total value of the plant output depends on the number of incoming orders D, which is a random variable following a Wiener process, as described in Section 1.5. For simplicity, we neglect all dependencies on other macroeconomic variables. This can be justified by observing that the rate of output is implicitly dependent on other factors, e.g., the current instantaneous riskless interest rate.

By Itô's lemma (as in [2]), the change in the value of the plant is given by

$$dH = H_D dD + H_t dt = \frac{1}{2}H_{DD}(dD)^2. \tag{1.1}$$

We neglect the impact of taxation, as well as the possibility of catastrophic macroeconomic events (crash of the stock market, war etc.). Therefore, the revenue delivered by the plant is given by $q(D - A)$, where q is the rate of production and A is the marginal cost of production incorporating

the fixed costs. Applying the arbitrage argument as in [1], we arrive at the following equation for H:

$$\frac{1}{2}\sigma^2 D^2 H_{DD} + q(D - A) + H_t + (\rho S - C)H_D - \rho H = 0, \tag{1.2}$$

where σ is the volatility of demand, ρ the riskless interest rate, S the price of the airplane, C the convenience yield.

The boundary conditions can be set as usual, i.e.,

$$H(0) = F, H_{DDD}(P) = 0, \tag{1.3}$$

where F is the fixed cost of maintaining the plant and P is its production capacity. The terminal condition at time T when the order is filled (planes delivered) can be set as

$$H(T) = q(D - A). \tag{1.4}$$

Equation (1.2), subject to boundary conditions (1.3) and the terminal condition (1.4), can be solved by conventional numerical methods.

1.6 Conclusion

We showed that the nature of demand follows a geometric Brownian motion. This conclusion was reached by doing a time series analysis of the data. For simulation purposes the selling price of an airplane was determined by the Black Scholes model for the European options. We used the approach of Brennan and Schwartz to determine the valuation of the plant when the options are real.

In conclusion we state that, maintaining inventory might be a good idea to increase the revenue. Modeling of the projected demand and the option pricing for airplanes needs to be studied in more detail for the simulation to be effective. Valuations of the plant based on real options need to be explored further.

Preliminary results suggest that a detailed study is needed to evaluate the efficacy of the approach.

1.7 References

[1] M.S. Brennan and E.S. Schwartz, Evaluating Natural Resource Investments, The Journal of Business, 58 (2), 1985, 135 - 158.

[2] P. Wilmott, S. Howison, J. Dewynne, The mathematics of Financial Derivatives: A student introduction, Cambridge University Press, Cambridge, 1998.

[3] J. A. Buzacott and J. G. Shantikumar, Stochastic models of manufacturing systems, Prentice Hall, 1993.

[4] J. A. Buzacott and D. D. Yao, On queuing network models of flexible manufacturing systems, *Queuing Systems: Theory and Applications*, 1:5-27 1986.

[5] R. A. Jarrow, Finance Theory, Prentice Hall, 1986.

[6] J. W. Tukey, Exploratory Data Analysis, Addison Wesley, 1977.

```
> with(finance):T:=1; exercise:=75; rate:=0.03;
> sdev:=0.07;
>
```

$$T := 1$$

$$exercise := 75$$

$$rate := .03$$

$$sdev := .07$$

```
> for i from 1 by 1 to 364 do t[i]:=i:nperiods:=1-i/365:
  amount:=exercise*exp(rate*nperiods):
> price[i]:=evalf(blackscholes(exercise, exercise, rate, nperiods,
  sdev)):
> od:
> l:=[[n,price[n]] $n=1..364]:
> plot(l,x=1..364);
```

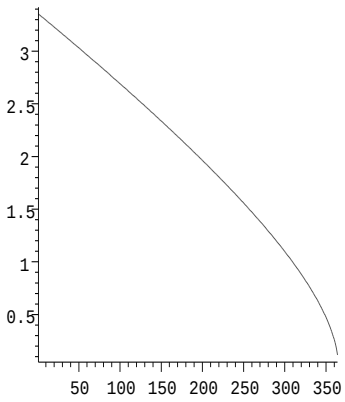


Figure 1.6: Option price for the planes

Chapter 2

Inventory Optimization using a Renewal Model for Sales

Dan Calistrate[1], Marc Paulhus[2], Miro Powojowski[3], Gordon Sick[4],

Report written by Marc Paulhus

2.1 Introduction

The problem posed by Dr. Greg Robel from The Boeing Company was to optimize inventory control and production rates of a certain item in order to take advantage of unexpected surges in demand. Certainly any optimization will require a realistic and simple model to predict future sales. The problem presenter had suggested that sales might arrive according to a Poisson distribution.

We suggest that the company look at *renewal theory* for models of future sales orders. They have some very distinct advantages. They are powerful and reasonably easy to use, in fact the Poisson distribution is a special case. Renewal models are flexible enough to incorporate a variety of characteristics, such as clustering or regularity, upward or downward drift, and mean reversion. Also, some arguments can be made to justify the renewal models based on financial intuition.

The next section will define and discuss the renewal models that we are suggesting. In Section 2.3 we set up a realistic (albeit naive) problem for which we have an analytic solution. The last section will more directly address the problem presented and will discuss possible solution methods.

2.2 Renewal Models for Future Sales

An informal definition of the renewal process is now given. For a more rigorous treatment of the renewal theory the reader is referred to any standard text on the subject, e.g., Cox [1962]. Some events - in this case sales of a product - will be assumed to occur at point times in the interval $[0, \infty)$. Thus any sales history can be summarized in a (finite or infinite) sequence $\{T_i\}_{i \in J}$ (e.g., $\{10, 14, 31\}$ is a short sales history with three sales at times 10, 14 and 31). A probability density function $f(t)$ will be specified and it will be assumed that

Condition 1

$$P(T_1 \leq t) = \int_0^t f(\tau) d\tau$$

[1]University of Calgary
[2]University of Calgary
[3]Université de Montréal
[4]University of Calgary

as well as for any $k > 0$

$$P(T_{k+1} - T_k \leq t) = \int_0^t f(\tau)d\tau$$

Simply put, the first sale arrives randomly at T_1, following the density $f(t)$, and given that a sale occurs at some time T_k, the additional time till the next sale $T_{k+1} - T_k$ is also random with density $f(t)$. At each sale, the system is renewed and the history prior to the last sale is of no consequence.

For example, look at the probability density function given in Figure 2.2. Their precise definitions will be given later. For now notice that, with the middle density, most of the mass under the curve lies near the origin. Hence, it is most likely that another sale will occur soon after a sale has been made. If a sale is not made in a long period of time we can expect to have to wait even longer before the next sale occurs. This should produce a clustering effect in the sales predictions.

With the bottom density, there is very little mass near the origin and hence it is unlikely that a sale will occur immediately following another sale. This should cause the sale predictions to be somewhat regular.

The top density corresponds to the Poisson distribution.

Figure 2.2 displays a single sample realization of future sales under each of the corresponding density functions. You will notice that the middle density did produce clustered sale predictions whereas the bottom density predicted sales on a quite regular basis. The top density produced sale predictions somewhere in between.

These remarks can be given a precise formulation as follows: let

$$p(t, \Delta t) = P(T_{k+1} - T_k \in (t, t + \Delta t] \,|\, T_{k+1} - T_k > t)$$

This corresponds to the probability of a sale arriving at some time in the interval $(t, t + \Delta t]$ after the last sale, conditional on the event that there was no sale in the interval $(0, t]$ after the last sale. Let Δt be an arbitrary, but fixed, small positive number. In the case of the density $f(t)$ as in the middle figure, one finds that $p(t, \Delta t)$ is a decreasing function of t, so as the time since the last sale goes by, the probability of a sale arriving in a fixed time step Δt decreases. For the bottom density, $p(t, \Delta t)$ is initially an increasing function and it reaches a peak for some positive value of t. For the Poisson process, the function $p(t, \Delta t)$ is constant in t, hence this process is sometimes described as memoryless.

The mean of $f(t)$, μ, is the expected inter-sale waiting time. The amount of clustering is determined by the shape of the density function and is separate from the mean. Although any density function leads to a renewal process, it is proposed here that the Gamma distribution family be used. The Gamma family can be parametrized with just two parameters as follows:

$$\Gamma(\alpha, \rho; x) = \frac{(\frac{x}{\rho})^{\alpha-1} e^{-\frac{x}{\rho}}}{\rho\Gamma(\alpha)}$$

with $\rho > 0$, $\alpha > 0$. The mean of this distribution is $\rho\alpha$ and the variance is $\rho^2\alpha$. The parameter ρ is the scale parameter, controlling the mean, whereas α is the shape parameter, controlling the amount of clustering, regardless of the value of ρ. The Gamma family is convenient since it contains members exhibiting the clustered ($\alpha < 1$) and the regular ($\alpha > 1$) behaviours, as well as the Poisson process ($\alpha = 1$). Finally, the family is well known and many estimators for its parameters have been proposed and investigated.

If historical sales data is available then μ should be easy to estimate. Otherwise, μ should be determined by management based on future sales expectations. The choice of α should be made based on sales history tempered by managerial intuition.

We can model drift by including it in the mean of the renewal density function. That is, if we expect that demand should increase by 8% a year, then rather than renewing with an identical function after a sale, we renew with a density function with a correspondingly higher mean. Note that the mean of the renewal density function will depend on the absolute time, not on the time passed since the last sale. An example of a realization of sales with increasing drift is given in

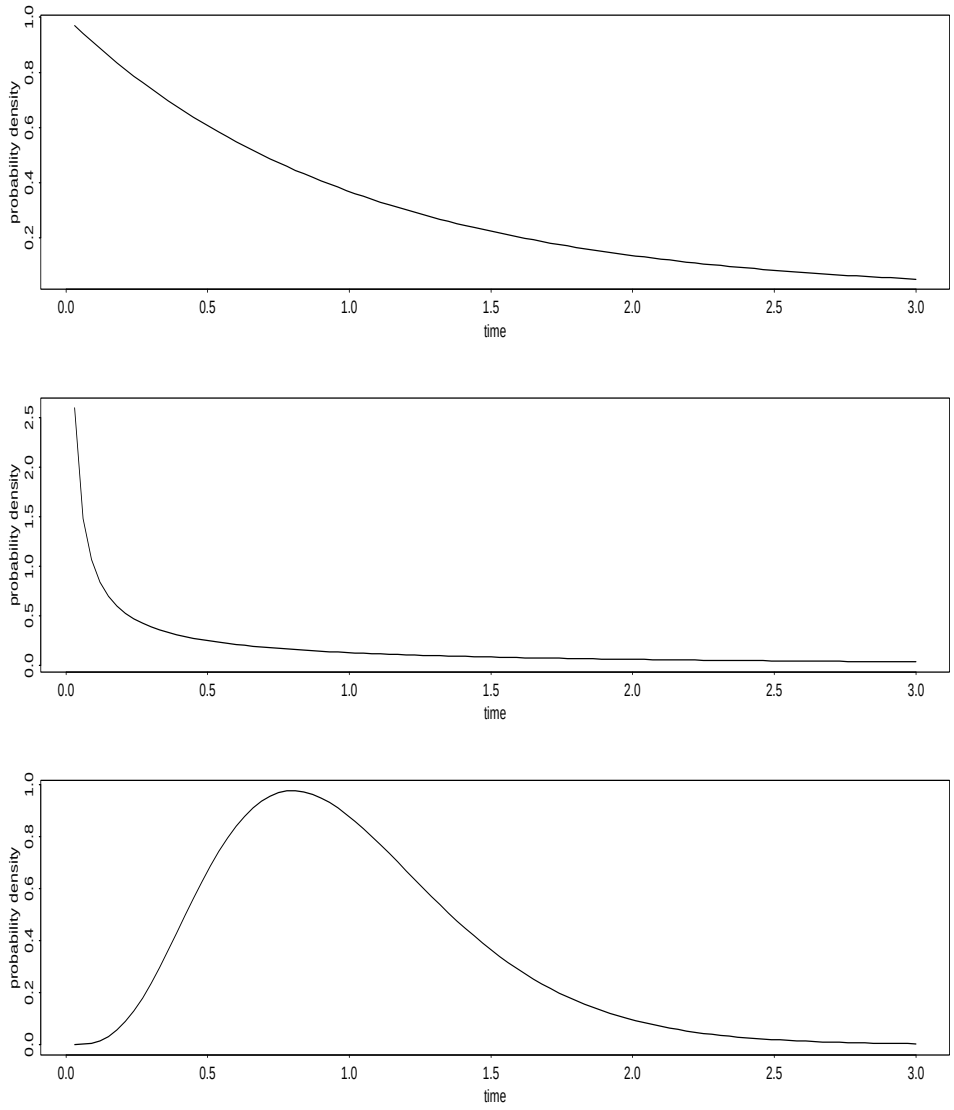


Figure 2.1: Examples of Gamma density functions. The top corresponds to the Poisson Distribution i.e. $\alpha = 1$. The middle corresponds to $\alpha = 1/5$ and will produce clustered sales predictions. The bottom corresponds to $\alpha = 5$ and will produce regular sales predictions.

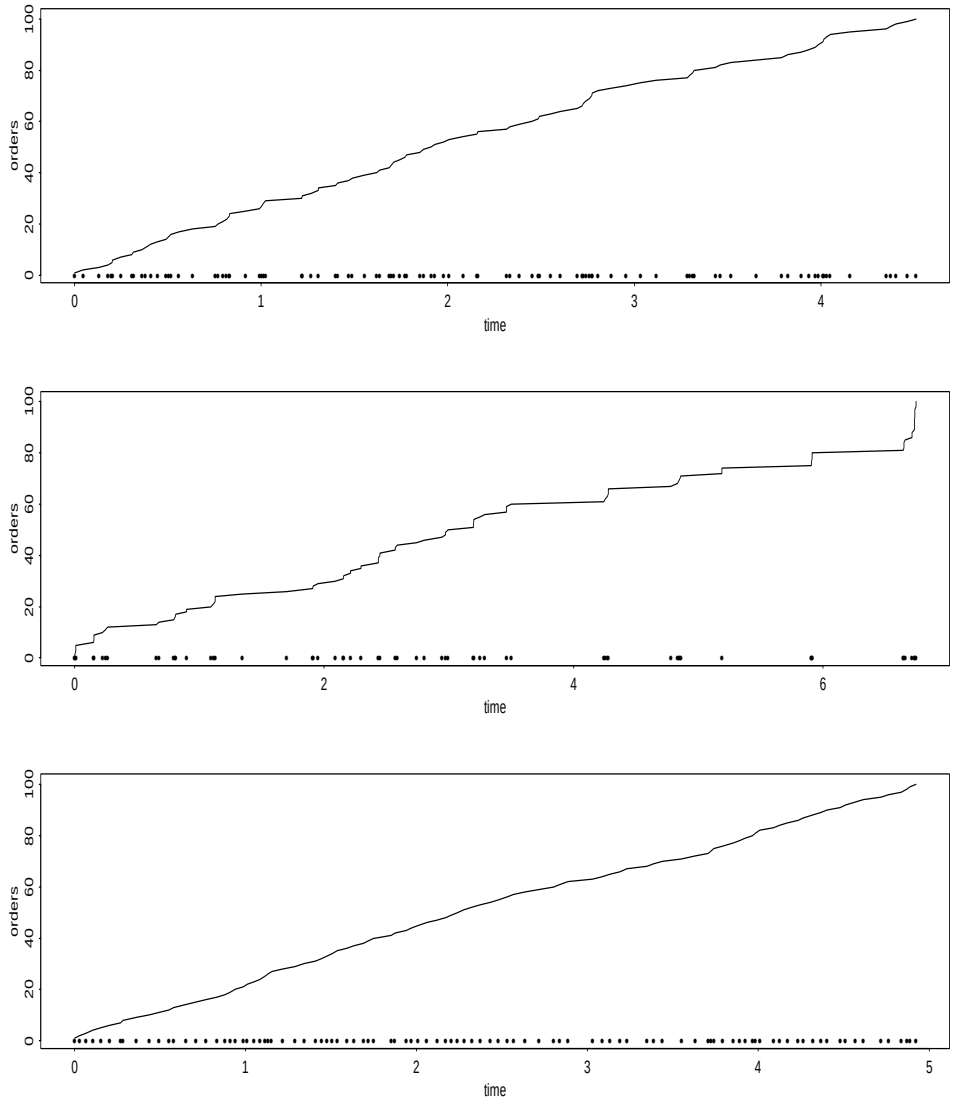


Figure 2.2: A simulated set of orders under the corresponding density functions with shape parameters as in Figure 2.2 and identical means. The dots along the bottom represent the simulated sales events. The line indicates the cumulative sales.

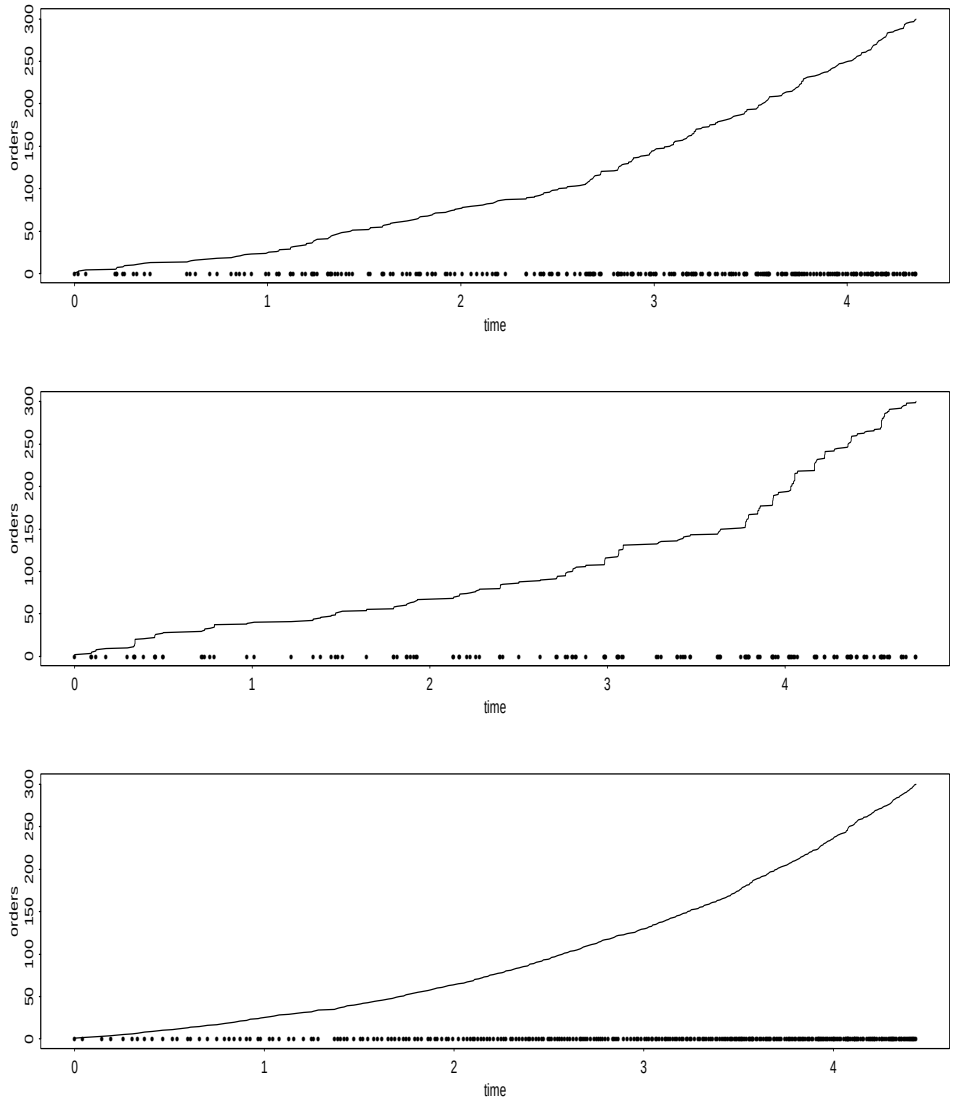


Figure 2.3: A simulated set of orders under the corresponding density functions with shape parameters as in Figure 2.2 with drift included. The dots along the bottom represent the simulated sales events. The line indicates the cumulative sales.

Figure 2.2. The density functions used to produce the figures are the same as those in Figure 2.2 except we included an upward drift in μ.

Mean reversion can be modeled in a similar way. If, in our simulated sales predictions, we are above/below the predicted mean then we can decrease/increase μ accordingly. One could also make α a function of the distance from the expected mean.

The process can become more complex if desired. For example, the management may have identified two independent customer types, one of which buys the product on a regular basis and one of which is more sporadic. Then, the sales could be modeled by two simultaneous and independent renewal processes, each with a different μ and α.

More complex models are easy to state but will complicate an analytical approach. If we are using simulation to answer our questions then the more complicated renewal processes should be no more difficult to use.

2.3 Analytical Solution to a Simpler Problem

To help the reader understand the procedure, we propose a hypothetical scenario which uses the renewal process and can be solved analytically. This scenario is much simpler than the problem posed to the workshop but serves as a useful example.

Suppose there is a software company that has developed a piece of software which we will call *product A*. Copies of the software are produced instantaneously and are sold to the customer for P dollars. Note that it is the lack of need for inventory control which makes this problem simpler than the one presented to the workshop. The company has already invested money in the development of product A and must now concentrate on optimizing the profits from the sale of the product. We will discretize the problem by defining a unit step of time. For the company to keep the product line alive for another unit of time will cost c dollars. This might reflect cost of advertising and the cost of a sales team.

Furthermore, the company has assigned the development team to another project and has no interest in developing further versions of product A. Like all software they know that the product will have a limited life span and hence we must answer the question, when is the optimal time to abandon the product (stop advertising and layoff the sales team). The residual value at abandonment will be K dollars which can be positive or negative. We are also interested in pricing the product line for the case that we have the opportunity to sell it.

Suppose management, based on past sale histories of similar products, concluded that the sales of the product will be clustered. This may be because a single buyer is likely to buy more than one copy or maybe if one corporation upgrades its software then it puts pressure on their competitors to do the same or maybe company B is copying the strategy of company A since company A has devoted some effort in researching the available products. In any case, μ and $\alpha < 1$ are chosen. Realistically we should include a downward drift but we will omit it for simplicity.

We will say that the product line is worth A_0 dollars at the start, which is an unknown that we are interested in computing. Start by looking at the tree of possible outcomes, Figure 2.3. Each level of the tree represents the next unit of time forward in our discrete model. Starting at the root of the tree, move right if a sale is made in the next time step and left if a sale is not made. The nodes of the tree are labeled with the value of the product line at that state. Due to the properties of the renewal process, if the product line was worth A_0 at time zero, then the product line will be worth A_0 just after any sale. We denote by A_t the value of the product line under the assumption that we have not made a sale in the last t time steps.

We will assume a risk-neutral investor and a riskless interest rate of r. Fix an arbitrary time, T, such that if a sale is not made in the last T time steps then we will abandon the product line. Hence we have assumed $A_T = K$. We now have the following set of linear equations:

$$\begin{aligned} A_t &= \frac{\pi_t(A_0 + P) + (1 - \pi_t)A_{t+1} - c}{1 + r} \quad \text{for} \quad t = 0 \ldots T - 1 \\ A_T &= K \end{aligned}$$

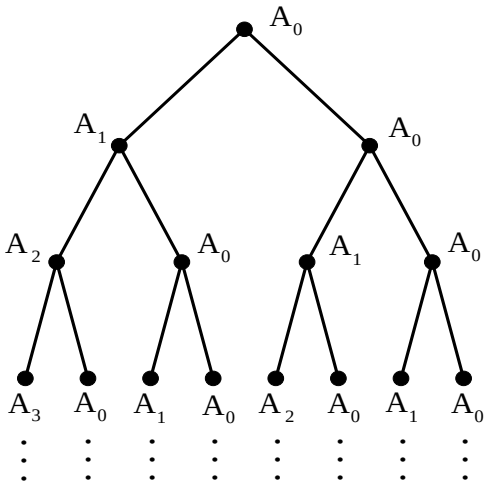


Figure 2.4: Tree of possible outcomes. A left move means no sale is made on the last time step, a right move means that a sale was made on the last time step. The labels represent the value of the product line at that state.

where π_t is the sale attenuation function, that is the probability that a sale is made at time t given that the last sale was made at time zero (a discrete version of the function $p(t, \Delta t)$ defined in section 2.2).

This is a linear system of $T + 1$ equations and $T + 1$ unknowns. Solve it to determine the value of the project under the assumption that the optimal time to abandon the project is if a sale has not been made in T time steps. Iterate this procedure with different choices for T and maximize A_0. This way we determine the current day value of the product line and the optimal time to abandon.

2.4 The Inventory Problem

The last section solved a problem in which we ignored the aspect of inventory. Here we shall set up the problem as it might be posed where the speed of production and the level of inventory are important factors.

Let us concentrate on a single product, in this case, a single line of planes. Looking at past sales history and estimating the future performance, management should establish which renewal density function should be used (include drift and mean reversion if desired). In light of the expected "unexpected surges in demand" it might be wise to consider a clustered sales model.

Again we will discretize the problem, maybe into one–day time steps. In what follows, please note the difference between a fixed one–time *cost*, and a *fee* which must be paid every time step.

We need to know the initial state of production, the possible states of production, the fee required to maintain each state, and the cost of switching between the states.

We suppose, for our hypothetical scenario, that a single plant can operate at the levels given in Table 2.1. Every time step we have the option to either open a new plant at a cost of D dollars, close a plant at a cost of A dollars, or pay $e_{i,j}$ dollars to change the production level at any existing plant from State i to State j.

We must also keep track of the number of planes in inventory at any time, which we will call I. At a given time step, if $I > 0$ then we must pay a fee equal to the cost of storing the planes, say cI. If $I < 0$ then we have over sold our stock and must pay a fee, say CI, which represents customer dissatisfaction. Note that we make no assumption about C and c. In fact, c might well be greater than C to a monopolist.

Another option, useful where competition is present, would be to have $I < 0$ affect the mean of the renewal function. That is, forcing the customer to wait for the delivery of the product may

	Airplane Production Rate	Fee to Maintain
State 1	100	15
State 2	200	5
State 3	400	4

Table 2.1: Production levels of a hypothetical assembly plant. The first column is the number of time steps required to produce an airplane. The second column is the fee required to maintain the plant at that level of production.

cause a loss in sales.

The optimal strategy for this problem is going to be a function of I, the state of production and the state of sales with respect to the density function. We do not have an analytic solution to this problem, but simulation can be used to price any fixed strategy (to some statistical accuracy) and then various methods can be used to try to optimize the strategy function.

2.5 References

[1] D.R. Cox, 1962. Renewal Theory. Science Paperbacks and Methuen & Co. Ltd.

Chapter 3

A Problem in Petroleum Reservoir Simulation

Alistair Fitt[1], Brian Seymour[2], John Stockie[3], Michèle Titcombe[4], Tony Ware[5]

Report compiled by Brian Seymour

3.1 Introduction

This report describes the results of the analysis of a problem in petroleum reservoir simulation presented by Peter Sammon of the Computer Modeling Group Ltd. (CMG). A summary of the problem, including contact information, is in the introductory pages of these proceedings.

The flow of two fluids (here, oil and water) in an underground petroleum reservoir is typically modeled by two partial differential equations: a parabolic equation for the fluid pressure in the reservoir, and a non-linear hyperbolic conservation law describing the fluid movement. In general, these equations are coupled through their coefficients, and are invariably solved numerically using finite difference schemes. The major difficulty in using these schemes is that they are very sensitive to the orientation of the finite difference grid.

The problem posed by Peter Sammon was to find an exact solution to the equations governing two-phase petroleum reservoirs that could be used to validate the numerical code used by the Computer Modeling Group in its simulations. Ideally, the analysis would result in an exact solution to the coupled system of equations for a specific rectangular geometry. It was soon realized that this was a hopeless task, so the group set out to find compromises that would still provide useful validation examples.

3.2 The Physical Problem

A simple model of the physical problem consists of a horizontal section of the petroleum reservoir in which water is pumped into an injector (source) at a specified velocity and oil extracted from a producer (sink) at a remote site. The geometry of the region shown in Figure 3.1 is typically part of a much larger region with many sources and sinks.

The saturation of water S is

$$S = \begin{cases} 1 & \text{if all water,} \\ 0 & \text{if all oil.} \end{cases} \tag{2.1}$$

[1]Faculty of Mathematical Studies, University of Southampton
[2]Department of Mathematics, University of British Columbia
[3]Department of Mathematics, Simon Fraser University
[4]Department of Mathematics, University of British Columbia
[5]Clinical Neurosciences, University of Calgary

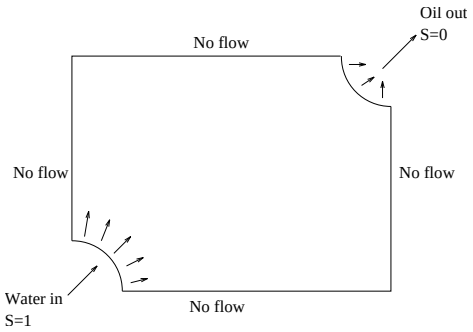


Figure 3.1: Typical reservoir simulation geometry.

3.3 Potential Problems

While there are conflicting opinions regarding the details of the model equations (e.g. a three-phase model), whichever model is used, the flow equations must be solved numerically, typically using upwind differencing. The major difficulty with this approach is that the numerical solutions exhibit grid orientation dependency. On grids such as those shown in Figure 3.2, the numerical scheme will give different (sometimes *very* different) results for the same problem. McCracken et al. [3] examine

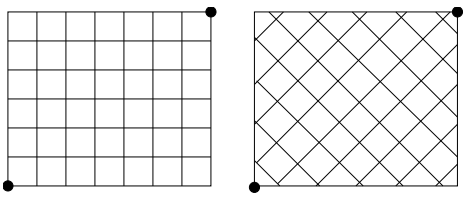


Figure 3.2: Rectangular grid orientations.

a numerical method for eliminating the dependence of the solution on grid orientation.

A question one might ask is whether or not the numerical scheme becomes more accurate as the grid size decreases? (See Section 3.8 for further comments regarding the well-posedness of the problem.)

3.4 Model Equations

The simplest 2-D model of petroleum reservoir simulation is described by the equations

$$\nabla \cdot (\lambda(S)\nabla p) = 0, \text{and } \phi S_t = \nabla \cdot (f(S)\lambda(S)\nabla p), \tag{4.2}$$

where p is the pressure, S the saturation and ϕ the porosity of the rock. The **total mobility**, $\lambda(S)$, and **fractional flow**, $f(S)$, are the key functions that determine the different forms of solutions.

The total mobility function is typically of the form:

$$\lambda(S) = \frac{k}{\mu_w}\left(\frac{(1-S)^\beta}{M} + S^\alpha\right), \tag{4.3}$$

where μ_w is the viscosity of water and k is the permeability of the rock. Similarly the fractional flow function is

$$f(S) = \frac{S^\alpha}{S^\alpha + \dfrac{(1-S)^\beta}{M}}, \tag{4.4}$$

where α and β are constants such that $\alpha \geqslant 1$ and $\beta \geqslant 1$. In these functions the **mobility** M is defined as

$$M = \frac{\mu_o}{\mu_w}, \tag{4.5}$$

and typically $M > 1$. For simplicity we take ϕ and k as constants.

Figure 3.3 shows the dependence of key functions on the constants M, α and β.

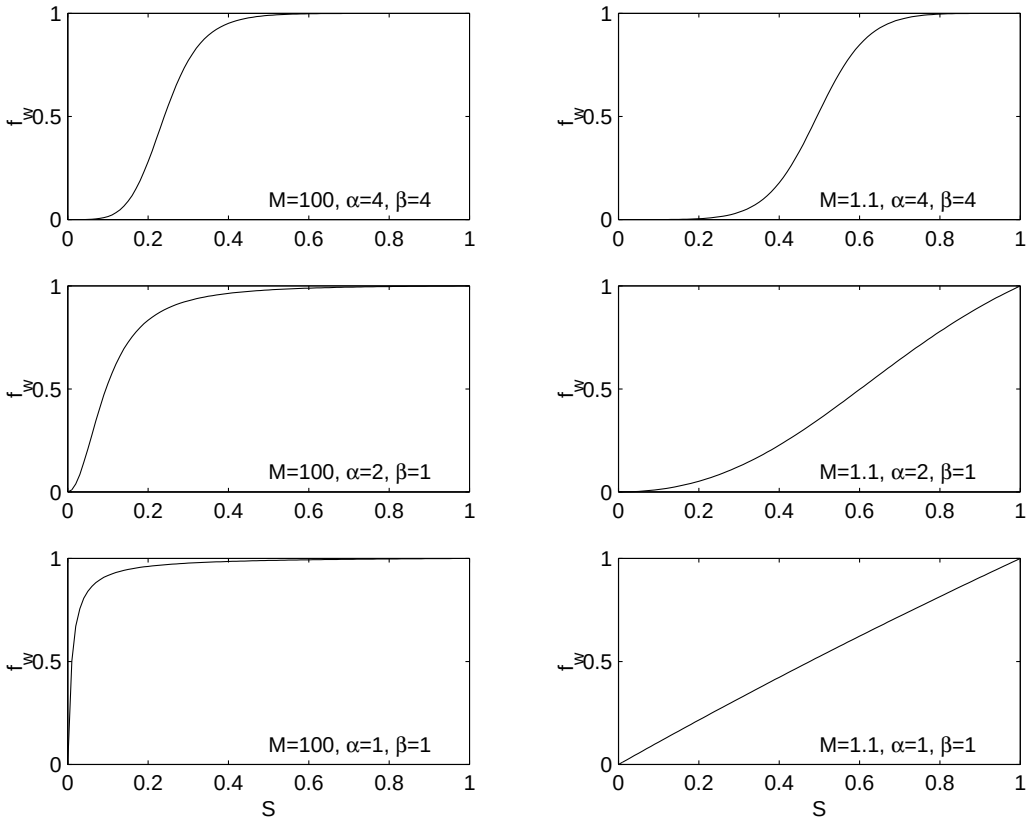


Figure 3.3: Dependence of key functions on M, α, and β.

3.5 Approaches for Finding Exact Solutions

Several approaches were initially considered:

[1] Two simple special solutions:

(a) Traveling waves of the form $S = S(ax + by - ct)$, and

(b) Similarity solutions: $S = t^q f(xt^\gamma, yt^\gamma)$. Although both will produce solutions, through coupled ordinary differential equations, there is no hope of satisfying realistic boundary conditions. Neither was pursued.

[2] Uncoupled solutions found by the following steps:

(a) Guess a velocity field (the least sensitive to changes at the front), that must be rotational with source/sink at corners.

(b) Given $(u, v) = -\lambda(S)\nabla p$, find S from $S_t = \nabla \cdot (f(S)\vec{v}) = 0$ and then find p. This approach turned out to be more difficult than originally estimated and was not pursued.

[3] Assume the flow is radially symmetric:

(a) This is the same as the *local* flow around a source. This was the approach that was considered in detail and in the next section we elaborate on this case. This approach has been considered before. For example, Bajor et al. [1] test the grid orientation problem on a radially symmetric problem, although their paper omits any details of the analytic solution. So far the exact radially symmetric solution has not been located in the literature.

An attempt was also made to check the solution with a simple numerical code. John Stockie had a code available from an undergraduate project. [It turned out that this code, from Waterloo, was originally developed in part by Peter Sammon.] This code is unlikely to produce anything new in a short time but could be useful as a check.

3.6 Radially Symmetric Flow

The equations governing the pressure p and water saturation S are:

$$\nabla \cdot (\lambda(S)\nabla p) = 0, \text{ and } \phi S_t = \nabla \cdot (f(S)\lambda(S)\nabla p). \tag{6.6}$$

Assuming p and S are only functions of the radial variable r, these equations become

$$(r\lambda p_r)_r = 0, \text{ and } \phi S_t = \frac{1}{r}(rf(S)\lambda(S)p_r)_r. \tag{6.7}$$

We consider a source at $r = a$ of finite size, then from equation (6.9)

$$a\lambda(S)p_r = -g(t), \tag{6.8}$$

where $g(t)$ is an arbitrary source function. The negative sign indicates that $p_r < 0$ for outflow. From this, we obtain

$$\frac{\phi}{g(t)}S_t = -\frac{1}{r}[f(S)]_r. \tag{6.9}$$

Defining the new time variable τ as

$$\tau = \int_0^t \frac{g(\bar{t})}{\phi} d\bar{t}, \tag{6.10}$$

in terms of τ, (6.9) becomes

$$S_\tau + \frac{1}{r}[f(S)]_r = 0. \tag{6.11}$$

Redefining the source as $g(t) \equiv G(\tau)$ the initial and boundary conditions are:

$$S = \begin{cases} G(\tau) & r = a, \tau \geq 0 \\ 0 & r \succ a, \tau = 0. \end{cases} \tag{6.12}$$

Many different solutions to the problem arise, depending on the behaviour of the fractional flow function $f(S)$ and the initial function $G(\tau)$. We will usually consider $G(\tau)$ as a step function, so will concentrate on various forms of $f(S)$. For convenience, from now on we set $\tau = t$.

Case 1. If $f(S)$ is convex ($f''(S) > 0$), then

$$S = \begin{cases} G(\alpha) & \left(t \geq \dfrac{r^2 - a^2}{2f'(G(0))}\right) \\ 0 & \left(t < \dfrac{r^2 - a^2}{2f'(G(0))}\right). \end{cases} \tag{6.13}$$

The characteristic variable, α, is defined by

$$t - \alpha = \frac{r^2 - a^2}{2f'(G(\alpha))}, \tag{6.14}$$

so that $t = \alpha$ at $r = a$.

In this case, we have a simple solution along characteristics. A shock occurs if

$$f''(G)G'(\alpha) = f'(G), \tag{6.15}$$

and therefore depends on f and G. For example, an $f = G^2$ shock occurs if $G'(\alpha) = G(\alpha)$ for some α.

If $r = U(t)$ is the shock path, it is determined by the shock condition:

$$\frac{dU}{dt} = \frac{f^+ - f^-}{(S^+ - S^-)U} \tag{6.16}$$

When $G(\alpha)$ is a step function, a single shock appears in plots of the saturation S as a function of radial distance r from the source. The shock propagates in the positive r direction as time increases. This case would provide a good numerical test for non-smearing and grid dependencies. The shock path is given by

$$U^2 - a^2 = 2t(f(1) - f(0)). \tag{6.17}$$

Case 2. When $f(S)$ is concave there is no shock for the same step function input. The solution is smooth, corresponding to a rarefaction wave connecting the characteristics carrying $S = 1$ and $S = 0$.

Case 3. The best test of all is for the real f. For example, f is convex for $0 < S < L$ and concave for $L < S < 1$. Here, $f'(L) = f(L)/L$ and $f'(1) \neq 0$. The general form of the solution for the water saturation is

$$\text{Constant } (S = 1) \mid \text{Rarefraction } (S = Q) \mid \text{Shock} \mid \text{Constant } (S = 0). \tag{6.18}$$

The characteristic net is divided into three regions with the corresponding solutions:

$$S = \begin{cases} 1 & \left(t \geq \dfrac{r^2 - a^2}{2f'(1)}\right) \\ N & \left(\dfrac{r^2 - a^2}{2f'(1)} \leq t \leq \dfrac{r^2 - a^2}{2f'(L)}\right) \\ 0 & \left(t < \dfrac{r^2 - a^2}{2f'(L)}\right), \end{cases} \tag{6.19}$$

where N is given by

$$L \leq N \leq 1;\ f'(N) = Q$$
$$Q = (r^2 - a^2)/2t,\ f'(1) \leq Q \leq f'(L).$$

This case tests both the smooth parts of the flow *and* shock-smearing. Figure 3.4 displays the characteristic curves for the mixed convex/concave f case.

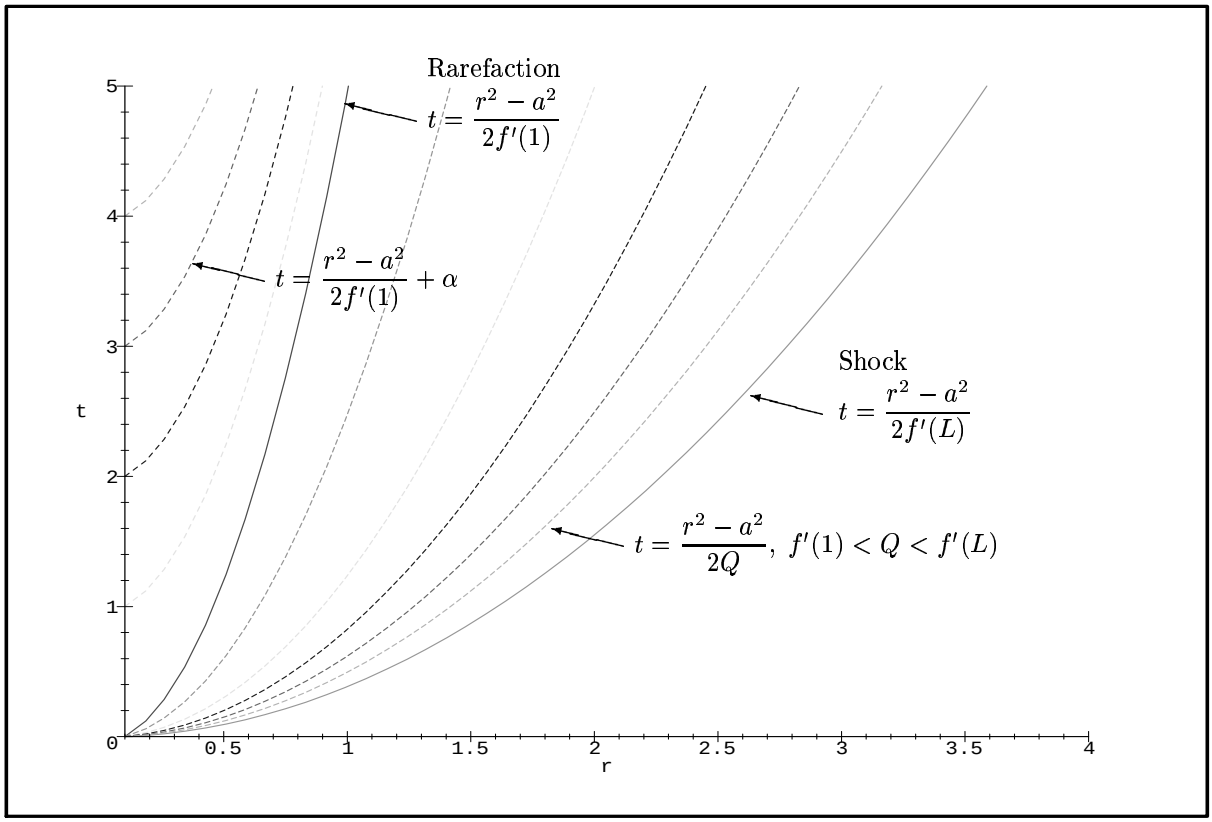


Figure 3.4: Characteristic curves for the mixed convex/concave f case.

3.7 Results and Discussion

We present the results of our study through various examples.

For $M = 10$, $\alpha = 4$ and $\beta = 1$, we have plotted the saturation profile in Figure 3.5 and the pressure profile in Figure 3.6 for two fixed times.

We computed the pressure and saturation solutions using a simple numerical code (implicit in pressure and explicit in saturation). Figures 3.7 and 3.8 display the resulting pressure and saturation fields along a ray emanating from the injector (source) to the producer (sink). Figures 3.9 and 3.10 show the corresponding pressure and saturation fields over the entire solution domain.

3.8 Stability and Well-posedness

Although we are confident that the radially symmetric problem described above should provide a good test for both shock-capturing and grid orientation dependence, it is worth expending some thought on the question of whether or not the generic flow problem is well-posed or not. To begin to understand why we might anticipate difficulties, we consider first the standard porous medium problem. This exhibits similarities to the full oil recovery problem considered above, but is technically much simpler.

Assume that the region to the left of the boundary $f(x, y, t) = 0$ (see Figure 3.11 for schematic details) is a porous medium filled with fluid, whilst to the right there is no fluid. By Darcy's law we

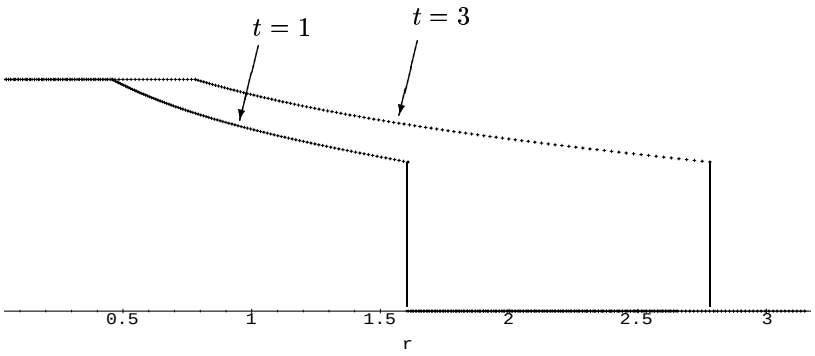


Figure 3.5: Saturation S versus radial variable r at times $t = 1$ and $t = 3$ for $M = 10$, $\alpha = 4$ and $\beta = 1$.

have

$$\boldsymbol{q} = -k\nabla p$$

where, as usual, $\boldsymbol{q}$ and p denote the velocity and pressure respectively and k (dimensions msec/kg) is the relative permeability. For incompressible fluid, conservation of mass now gives the equation of motion in the fluid-filled porous medium as

$$\nabla^2 p = 0.$$

As far as boundary conditions are concerned, we assume that the pressure is given on $f = 0$ by $p = p_a$, say. Another condition is also needed to determine the free boundary; this condition simply asserts that $Df/Dt = 0$ on $f = 0$ where D/Dt denotes the Lagrangean derivative. We find that, on $f = 0$, we must have

$$f_t = k\nabla p \cdot \nabla f.$$

This problem has a simple traveling wave solution with speed v; it is given by

$$p = p_a + \frac{v}{k}(vt - x), \qquad f = x - vt.$$

We wish to determine under what circumstances this traveling wave is stable to y-perturbations. Proceeding in the obvious fashion, we consider linear stability by assuming that

$$\begin{aligned} p &= p_a + \frac{v}{k}(vt - x) + \epsilon g(x,t)\sin ny \\ f &= x - vt + \epsilon e^{qt}\sin ny \end{aligned}$$

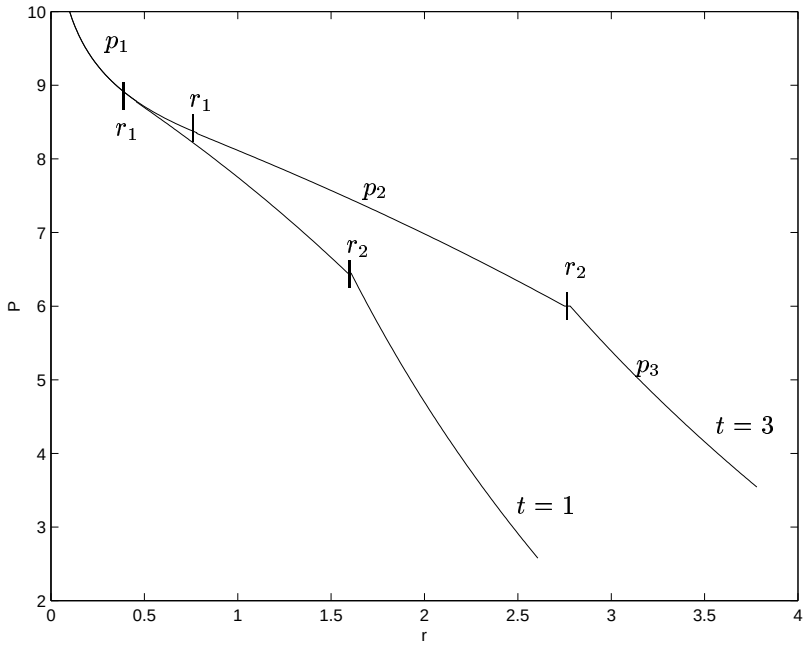


Figure 3.6: Pressure p versus radial variable r at times $t = 1$ and $t = 3$ for $M = 10$, $\alpha = 4$ and $\beta = 1$. For $a < r \leq r_1$, $p = p_1 = p_a - (A/\lambda(1))\ln(r/a)$. For $r_1 \leq r \leq r_2$, $p = p_2 = p_1(r_1) - A\int_{r_1}^{r}(1/\lambda(S(a))q)dq$. For $r > r_2$, $p = p_3 = p_3(r_2) - (A/\lambda(0))\ln(r/r_2)$.

where $\epsilon \ll 1$ and $g(x,t)$ and the amplification factor q are to be determined. The flow equation in the fluid-filled porous medium gives

$$g = A(t)e^{nx} + B(t)e^{-nx}$$

and assuming without loss of generality that $n > 0$, we must take $B = 0$ so that the perturbations become negligible as $x \to -\infty$. Imposing $p = p_a$ and $f_t = k\nabla p \cdot \nabla f$ on $x = vt - \epsilon e^{qt}\sin ny$ now gives the equations

$$\begin{aligned} 0 &= \frac{v}{k}e^{qt} + A(t)e^{nvt}, \\ \frac{q}{k}e^{qt} &= nA(t)e^{nvt} \end{aligned}$$

so that $q = -vn$ and $A(t) = (-v/k)\exp(-2nvt)$. Since n was assumed to be greater than zero, we recover the well-known result that *retreating* flow ($v < 0$) in a porous medium is unstable. (A very similar argument may be applied to the Hele-Shaw flow shown in Figure 3.12 to show that retreating fronts in such a cell are unstable.)

This result is potentially worrying. If we consider a retreating porous medium flow as discussed above as one where a 'more viscous' fluid (water) is being expelled by a 'less viscous' fluid (air), which is the case that pertains for nearly all of CMG's calculations, then the instability of the process could lead to gross errors in the numerical calculations. To be sure that this problem is present, however, we need to ensure that instability is present for the full equations.

3.8.1 Stability of a Paradigm Shock Solution

In this subsection we analyze a simple paradigm shock problem to finish developing the methodology necessary to attack the full problem.

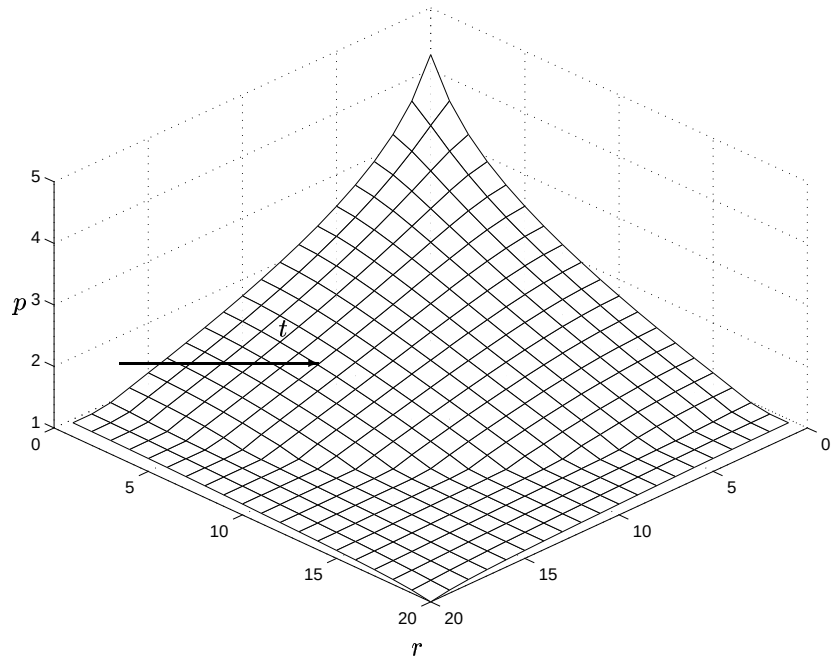


Figure 3.7: Pressure field p along a ray emanating from the injector to the producer (computed using the simple numerical code).

To illustrate the methods that we wish to use, consider the conservation law

$$S_t + (S^2/2)_x + (S^2/2)_y = 0.$$

It is easily confirmed that this possesses the one-dimensional shock solution

$$S = \begin{cases} S_L & (x < vt) \\ 0 & (x \geq vt) \end{cases}$$

where, according to the standard jump condition with $f(S) = S^2/2$

$$v = \frac{[f(S)]}{[S]} = \frac{S_L}{2}.$$

We note also that since $f'(S_L) = S_L$ and $f'(S_R) = 0$ then the entropy condition $f'(S_L) > v > f'(S_R)$ is satisfied and so according to the standard theorems (see, for example Smoller [5]) the solution that has been determined is unique.

We now wish to examine the stability of this solution. Using the *ansatz* that, for a perturbed shock,

$$\begin{aligned} \psi &= x - vt - \epsilon f(y,t) \\ S &= S_L + \epsilon g(y,t) \qquad \text{(left of the shock)} \\ S &= 0 + \epsilon h(y,t) \qquad \text{(right of the shock)} \end{aligned}$$

we find that h and g satisfy the partial differential equations

$$g_t + S_L(g_x + g_y) = 0, \qquad h_t = 0$$

whilst the jump condition (evaluated at the perturbed shock front) gives at $\mathrm{O}(\epsilon)$

$$-S_L f_t - v\tilde{g} + v\tilde{h} + S_L\tilde{g} - \frac{1}{2}S_L^2 f_y$$

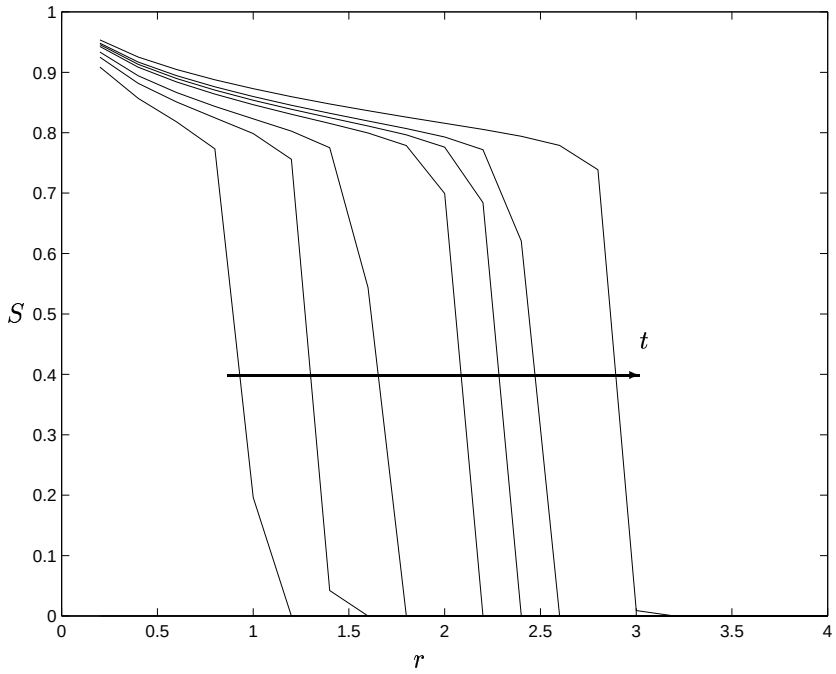


Figure 3.8: Saturation field S along a ray emanating from the injector to the producer (computed using the simple numerical code).

where $\tilde{g} = g(vt, y, t)$, $\tilde{h} = h(vt, y, t)$. Now suppose that the perturbation is of the obvious form

$$f(y,t) = Re(e^{iny}e^{\sigma t}).$$

Then the equations rapidly give $h = 0$ and $g = g_1(x)e^{iny}e^{\sigma t}$, followed by $g_1(x) = 2\sigma + inS_L$ and $\sigma = -iS_L n$. Thus the shock is neutrally stable (applied perturbations neither grow nor decay) and, on taking real parts,

$$\begin{aligned}
\psi &= x - \frac{S_L}{2}t - \epsilon\cos(n(y - S_L t)) \\
S &= S_L + \epsilon n S_L \sin(n(y - S_L t)) \qquad \text{(left of the shock)} \\
S &= 0 \qquad \text{(right of the shock).}
\end{aligned}$$

3.8.2 Stability for the Oil recovery problem

Guided by the methodology of the previous two subsections, we now consider the stability of the problem

$$(\lambda(S)p_x)_x + (\lambda(S)p_y)_y = 0 \tag{8.1}$$

$$(f(S)\lambda(S)p_x)_x + (f(S)\lambda(S)p_y)_y = S_t, \tag{8.2}$$

Here we have assumed without loss of generality that $\phi = 1$. For simplicity we consider the case $\alpha = \beta = 1$ (see later remarks) so that

$$f(S) = \frac{S}{S + \frac{1-S}{M}}, \qquad \lambda(S) = \frac{k}{\mu_w}\left(\frac{1-S}{M} + S\right).$$

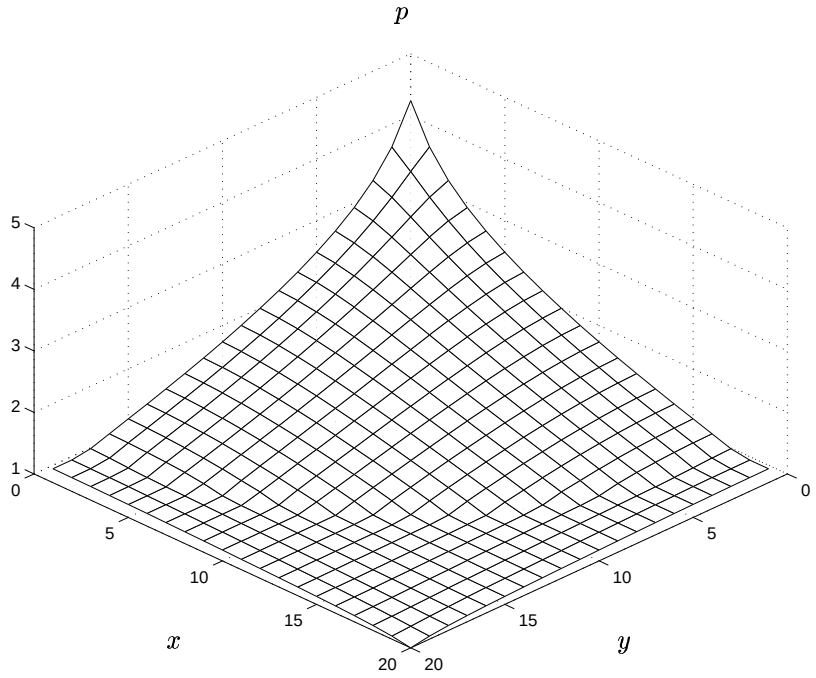


Figure 3.9: Pressure field p in the entire solution domain (computed using the simple numerical code).

Perturbing the problem in a similar fashion to the previous subsection, we assume that

$$\psi = x - t - \epsilon e^{\sigma t} e^{iny} + O(\epsilon^2) \tag{8.3}$$

$$S_L = 1 - \epsilon a(x,y,t) + O(\epsilon^2) \tag{8.4}$$

$$S_R = \epsilon b(x,y,t) + O(\epsilon^2) \tag{8.5}$$

$$p_L = -\frac{\mu_w x}{k} + \frac{t}{k}(\mu_w - \mu_o) + \epsilon c(x,y,t) + O(\epsilon^2) \tag{8.6}$$

$$p_R = -\frac{\mu_o x}{k} + \epsilon d(x,y,t) + O(\epsilon^2) \tag{8.7}$$

where by convention real parts are taken throughout. Note that the form of this *ansatz* ensures (a) that the pressure is continuous to leading order at the shock front, (b) that the one-dimensional solution is a shock of amplitude 1, and so by the jump conditions the shock speed is one, and (c) in the one-dimensional solution $\lambda p_x = -1$, so that the flux of material is constant in time and the front moves from left to right. (Fluxes that are functions of time may easily be dealt with by rescaling time as described above.)

We now need to pick a, b, c, d and σ so that the following conditions are satisfied:

- s_L and p_L satisfy the partial differential equations (8.1) and (8.2) to the left of the shock.
- s_R and p_R satisfy the partial differential equations (8.1) and (8.2) to the right of the shock.
- The correct jump conditions for (8.1) and (8.2) are satisfied.
- The pressure is continuous to O(ϵ) at the shock front.

Courant & Hilbert [2] give the jump conditions for a system of k conservation laws in n independent variables. If the system is written in conservation form as

$$\sum_{i=1}^{n} \frac{\partial P_{ij}}{\partial x_i} = 0 \qquad (j = 1, 2, ..., k)$$

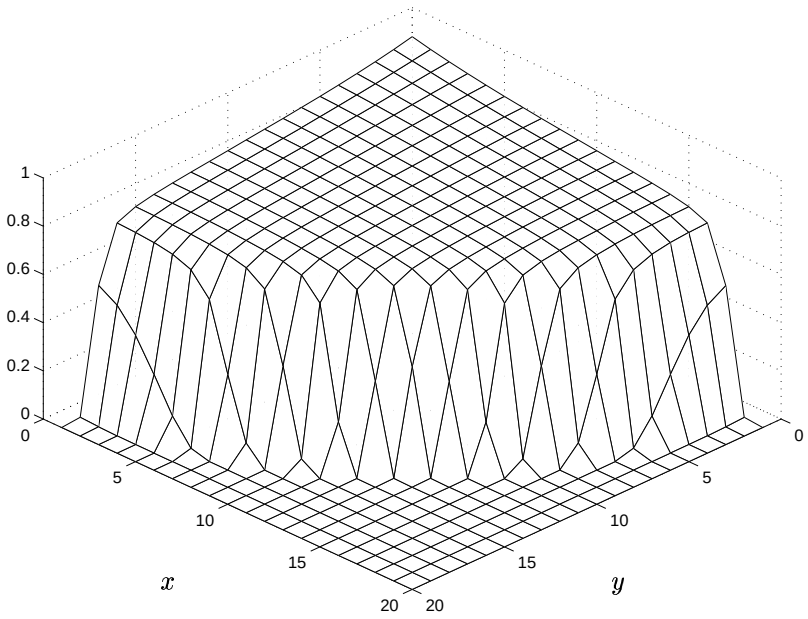


Figure 3.10: Saturation field S in the entire solution domain (computed using the simple numerical code).

some simple analysis shows that if the shock front is given by

$$\psi(x_1, x_2, ..., x_n) = 0$$

then the jump conditions are

$$\sum_{i=1}^{n} [P_{ij}] \frac{\partial \psi}{\partial x_i} \qquad (j = 1, 2, ..., k)$$

where, as usual, [.] denotes the jump in a quantity as the shock $\psi = 0$ is crossed. For the equations above the jump conditions are therefore

$$[\lambda p_x]\psi_x + [\lambda p_y]\psi_y = 0 \qquad [S]\psi_t - [f\lambda p_x]\psi_x - [f\lambda p_y]\psi_y = 0.$$

Inserting the *ansatz* into the relevant equations in the order indicated above, we find that at $O(\epsilon)$ the following seven equations must be satisfied:

$$c_{xx} + c_{yy} + \frac{\mu_w}{Mk}(M-1)a_x = 0 \tag{8.8}$$

$$c_{xx} + c_{yy} + \frac{\mu_w}{k}(a_t + a_x) = 0 \tag{8.9}$$

$$d_{xx} + d_{yy} + \frac{\mu_o}{k}(1-M)b_x = 0 \tag{8.10}$$

$$b_t + Mb_x = 0 \tag{8.11}$$

$$c_x - \frac{1}{M}d_x + \frac{(1-M)}{Mk}[-a\mu_w - b\mu_o] = 0 \tag{8.12}$$

$$-\sigma e^{\sigma t} e^{iny} + b(1-M) - \frac{k}{\mu_w}c_x = 0 \tag{8.13}$$

$$\frac{\mu_o - \mu_w}{k} e^{\sigma t} e^{iny} + c - d = 0 \tag{8.14}$$

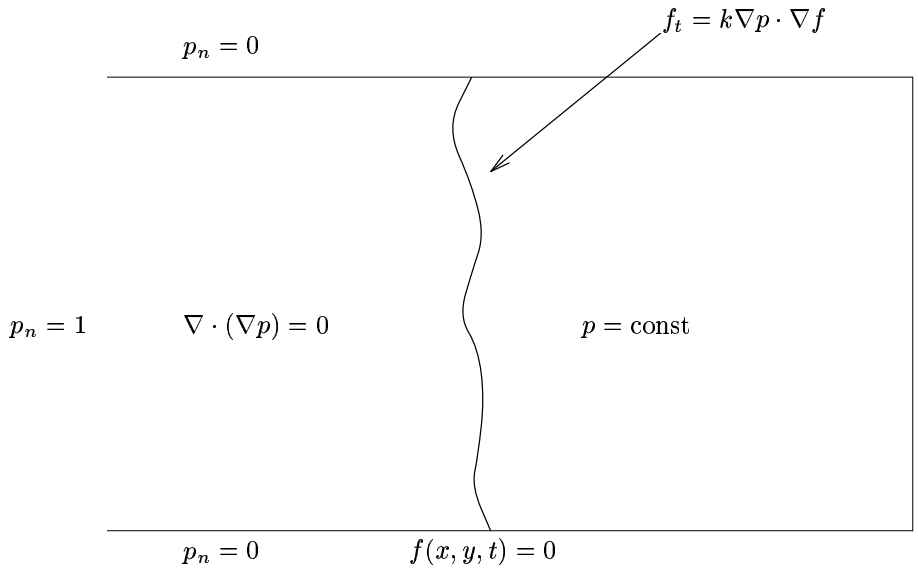


Figure 3.11: Darcy porous medium.

The first four of these equations are partial differential equations for functions of x, y and t, whilst the last three are to be evaluated at $x = t$ and are therefore equations in the independent variables y and t only.

To solve (8.8)-(8.14), we make the further *ansatz* (again assuming that real parts are taken throughout) that

$$a = a_1 e^{a_2(Mx-t)} e^{iny} \tag{8.15}$$
$$b = b_1 e^{b_2(x-Mt)} e^{iny} \tag{8.16}$$
$$c = c_1 e^{a_2(Mx-t)} e^{iny} \tag{8.17}$$
$$d = d_1 e^{b_2(x-Mt)} e^{iny}, \tag{8.18}$$

these choices having been made so as to automatically satisfy (8.11) and the simple equation that results from subtracting (8.9) from (8.8). Now (8.13) gives the three equations

$$b_2(1-M) = \sigma \tag{8.19}$$
$$a_2(M-1) = \sigma \tag{8.20}$$
$$(1-M)b_1 - \frac{k}{\mu_w} M a_2 c_1 = \sigma \tag{8.21}$$

whilst (8.8), (8.10), (8.12) and (8.14) in turn give

$$M^2 c_1 a_2^2 - n^2 c_1 + \frac{\mu_w}{k}(M-1)a_1 a_2 = 0 \tag{8.22}$$
$$d_1 b_2^2 - n^2 d_1 + \frac{\mu_o}{k}(1-M)b_1 b_2 = 0 \tag{8.23}$$
$$M c_1 a_2 - \frac{1}{M} b_2 d_1 + \frac{(1-M)}{Mk}[-\mu_w a_1 - \mu_o b_1] = 0 \tag{8.24}$$
$$c_1 - d_1 + \frac{\mu_o - \mu_w}{k} = 0. \tag{8.25}$$

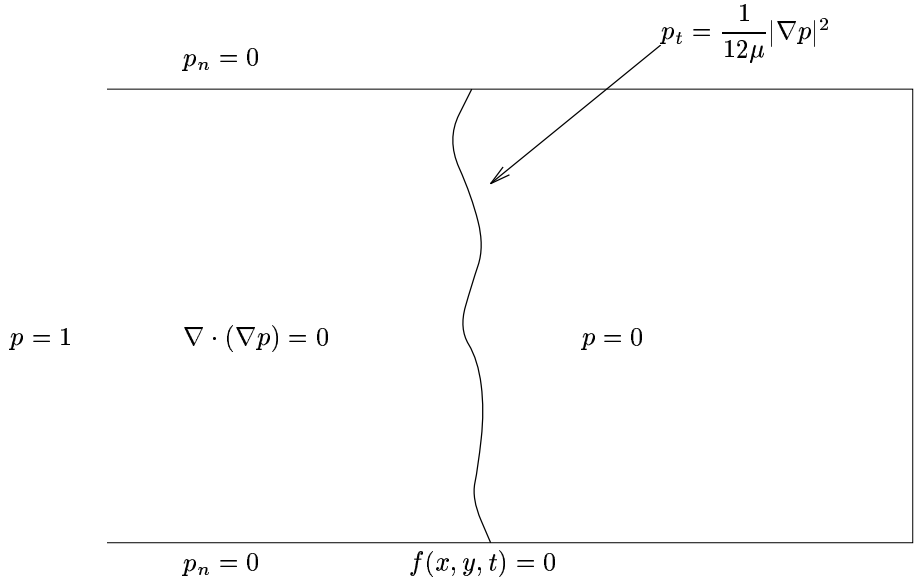


Figure 3.12: Hele-Shaw.

The equations (8.19)-(8.25) constitute seven equations in the seven unknowns a_1, b_1, c_1, d_1, a_2, b_2 and σ so nominally the problem formulation is complete.

Before proceeding further, let us consider how the *ansatz* that we have made is related to these equations. To the left of the shock, the perturbations must decay as $x \to -\infty$, so that necessarily $a_2 > 0$. But now if $\mu_o > \mu_w$ so that M exceeds 1 (as is normally the case), then (8.20) suggests that σ is positive. A similar argument applied to the right of the shock shows that $b_2 < 0$ and gives the same conclusion regarding the sign of σ. Thus according to (8.3) the shock position is unstable and perturbations grow exponentially. This instability is likely to manifest itself as 'fingering', where the water infiltrates the oil in an unstable manner, similar to the porous medium instability discussed above. As far as the size of the shock is concerned, we note from (8.15) that since $a_2 > 0$, perturbations to the left of the shock decay exponentially in time. To the right of the shock, however, since $b_2 < 0$ th perturbations grow exponentially. Thus both the position of the shock front and S are linearly unstable. (Since we have examined the case $\alpha = \beta = 1$ there is no technical reason why S cannot take values less than zero, which is unphysical. When α and β are not one, however, this problem may easily be dealt with.)

Although the conclusion that the simple shock front solution is unstable now seems unavoidable, it is necessary to check (for consistency) that amplification factors with positive real parts actually exist. Eliminating a_2, b_1, b_2 and d_1 we find from (8.22), (8.23) and (8.24) that

$$c_1 \left(\frac{\sigma^2 M^2}{(M-1)^2} - n^2 \right) + \frac{\mu_w a_1 \sigma}{k} = 0$$

$$\left(c_1 + \frac{\mu_o - \mu_w}{k} \right) \left(\frac{\sigma^2}{(1-M)^2} - n^2 \right) + \frac{\mu_o \sigma^2}{k(1-M)} \left(1 + \frac{kMc_1}{\mu_w(M-1)} \right) = 0$$

$$\frac{c_1 M \sigma}{M-1} + \frac{\sigma}{M(M-1)} \left(c_1 + \frac{\mu_o - \mu_w}{k} \right) +$$

$$\frac{(1-M)}{kM}\left(-a_1\mu_w - \frac{\mu_o}{1-M}\left(\sigma + \frac{k\sigma M c_1}{\mu_w(1-M)}\right)\right) = 0$$

and a_1 and c_1 may now be eliminated to give an equation involving σ alone. After some (rather tedious and lengthy) calculations, we find that the amplification rate σ satisfies the quintic equation $\theta(\sigma) = 0$ where

$$\theta(\sigma) = K\sigma^5 + M(1-M)^2\sigma^4 - Kn^2(1-M)\sigma^3 - n^2M^2(1-M)^3\sigma^2 + n^4(1-M)^5 \qquad (8.26)$$

and $K = k/\mu_w$. This equation is tedious to analyze (though various special cases where the parameters K, M and n are large or small may be examined), but we note immediately that $\theta(0) < 0$ for $M > 1$ whilst for $\theta(\sigma) \to +\infty$ as $\sigma \to +\infty$. It is therefore certain that a real root greater than zero exists, confirming that the problem as posed is, indeed, ill-posed.

What may we conclude from this analysis? Unfortunately it seems clear that, as might be expected, the simple porous medium result that forcing a more viscous fluid with a less viscous one is unstable is not altered for the more complicated oil recovery problem. This is likely to lead to difficulties in computing numerical results, and may be one of the factors that influences the grid orientation dependence. As far as the computation of numerical results is concerned, of course, the code is unlikely to 'see' the equations (8.1) and (8.2). Instead, the discretization used will introduce 'artificial' viscosity and other dissipative mechanisms which may act to at least partially damp out numerical instabilities.

Finally, a few general comments are in order. Although for simplicity we only considered the case $\alpha = \beta = 1$ in detail, other values of α and β may be dealt by making various changes to the solution *ansatz*. We have not analyzed these fully yet, but it seems likely that the conclusions will be similar. We also note that, whilst the first two problems analyzed in this section are standard book work, we have not been able to find an analysis of shock front stability for the full oil recovery problem in the literature, though it seems likely that one must exist somewhere.

Acknowledgment

We would like to thank Peter Sammon for explaining the petroleum reservoir simulation problem to us and for answering our constant deluge of questions.

3.9 References

[1] Bajor and Cormack, The Quantitative Characterisation of the Grid Orientation Problem, AOSTRA of Research, Vo. 235, 1992.

[2] R. Courant and D. Hilbert, Methods of Mathematical Physics, Wiley-Interscience, Vol. 2, New York, 1962.

[3] T.A. McCracken and J L. Yanosik, A nine-point, finite difference reservoir simulator for realistic prediction of adverse mobility ratio displacements, Transactions AIME, Vol. 267, 1979, pp. 253–262.

[4] M.K. Abdou et. al., Impact of grid selection on reservoir simulation, Journal of Petroleum Technology, Vol. 45, 1993, pp. 664–669.

[5] J. Smoller, Shock Waves and Reaction-Diffusion Equations, Springer-Verlag, 1983.

Chapter 4

On Seismic Imaging: Geodesics, Isochrons, and Fermat's Principle

V.E. Agapov[1], R. Ait-Haddou[2], C.S. Bohun[3], A.T. Dawes[4], A.M. Duffy[5], J.P. Grossman[2], A.J. Irwin[6], S.T. Jensen[7], J.V. Modayil[1], E. Nyland[1], V.E. Shapiro[4], M.A. Slawinski[2], P.S. Webster[2]

Report compiled by P.S. Webster

4.1 Introduction

The problem described in this report deals with the propagation of waves in elastic solids. The particular application was suggested by Baker Atlas, PanCanadian Petroleum, Petro-Canada Oil and Gas, and Talisman Energy, i.e., the companies involved in seismic data acquisition, processing and interpretation and sponsoring The Geomechanics Project at The University of Calgary.

Seismic data is used to obtain the information about the subsurface, which includes spatial positions of reflectors. In the case of exploration seismology, the known parameters are: the location of the source, the location of the receiver, and traveltime of a signal between the source and receiver. (In earthquake seismology the first and the third parameter might have to be inferred.)

To proceed with the geophysical investigations one often needs to assume knowledge of the material field which describes the velocity of a signal as a function of position (inhomogeneity) and direction (anisotropy). Consequently, fitting the observed data with the assumed characteristics of the material field leads to an image of the subsurface. We investigate the scattering from a subsurface point of a signal generated by the source and recorded at the receiver. If the signal velocity in the medium is constant, i.e., the medium is homogeneous and isotropic, the set of all possible reflection points for the measured traveltime forms an ellipse whose foci correspond to the source and receiver positions and whose "string length" is equal to the product of the assumed velocity and the measured traveltime. This ellipse is an isochron, the locus of the same traveltime, corresponding to all possible reflection points of the traveltime geodesics. (The geodesics are the least-time raypaths between a given source and receiver.) A continuous set of reflection points, such as a planar reflector, corresponds to the reflective surface whose location coincides with intersection of numerous ellipses generated by numerous source-receiver configurations. The location of the intersection points also coincides with the outer envelope (convex hull) containing all these ellipses.

[1]University of Alberta
[2]University of Calgary
[3]University of Victoria
[4]University of British Columbia
[5]University of Saskatchewan
[6]Queen's University
[7]McGill University

The intersection points have served as a standard way for creating a subsurface image based on seismic data derived from numerous sources and receivers.

A more accurate model of the material field should take into account both inhomogeneity and anisotropy of the subsurface. Consequently, the geodesics might not form straight lines and isochrons might not form ellipses. Furthermore, the intersection points might not coincide with the convex hull. For instance, complex material fields required for accurate imaging of seismic data acquired in the foothills might not allow the benefit of certain simplifying assumptions. Notably, the convenient coincidence of intersection points and the envelope should be revisited. In order to provide a more general approach we allow sources and receivers to be located at different levels, for instance in data acquisition over a complicated surface, or positioning of receivers in the wellbore, as in Vertical Seismic Profiling.

This report constitutes a significant contribution of all the authors. In particular, the selected sections are derived from written documents provided by A.T. Dawes, A.M. Duffy, A.J. Irwin, S.T. Jensen, and E. Nyland. The aforementioned documents are available upon request. Also, C.S. Bohun, E. Nyland and P.S. Webster wrote computer code in Mathematica and Maple, which is available on request.

4.2 Background Of The Problem

4.2.1 Mathematical foundations

The differential equations that describe the propagation of seismic pulses in a heterogeneous anisotropic material are well known [3]. If a continuum with elasticity tensor C_{ijpq} and mass density ρ undergoes a space and time dependent deformation u_i and the spatial derivatives of the displacement are small, then

$$(C_{ijpq}up,q)_{,j} = \rho\ddot{u}_i.$$

Suppose this equation has a solution of the form

$$u_i(\boldsymbol{r},t) = U_i(t - T(\boldsymbol{r}))f(\boldsymbol{r})$$

where U_i is essentially zero unless the time t is near the traveltime, $T(\boldsymbol{r})$, required for a wavefront to travel from the origin of coordinates to position $\boldsymbol{r}$. The scalar function f is smooth. Clearly $U_i \ll \dot{U}_i \ll \ddot{U}_i$ and all derivatives of f are also small. By ignoring all terms involving first or no derivatives of U and f, this elastodynamic equation can be written as

$$(\rho\delta_{ik} - C_{ijkl}T_{,j}T_{,i})\ddot{U}_k f \approx 0 \Longrightarrow \mid \rho\delta ik - C_{ijkl}T_{,j}T_{,i} \mid = 0.$$

In the case of isotropic media the last equation becomes

$$(\nabla T \bullet \nabla T - 1/\alpha^2)(\nabla T \bullet \nabla T - 1/\beta^2)^2 = 0$$

where α and β are the velocities of propagation of compressional and torsional pulse fronts respectively. The surface T = constant defines the pulse front and these eikonal equations imply that the perpendiculars to this pulse front, the ray trajectories, must satisfy the parametric equations

$$\partial_s \frac{\partial_s \boldsymbol{r}(s)}{c(\boldsymbol{r})} = \nabla \frac{1}{c(\boldsymbol{r})}$$

where c is either α or β and the solution is sought in a domain where the velocity is essentially constant over the width of the pulse. These equations for the ray trajectory are equivalent to Fermat's principle of stationary time along the trajectory

$$\delta \oint_{\boldsymbol{r}_1}^{\boldsymbol{r}_2} \frac{1}{c(\boldsymbol{r})} ds = 0 \Longleftrightarrow \partial_s \frac{\partial_s \boldsymbol{r}(s)}{c(\boldsymbol{r})} = \nabla \frac{1}{c(\boldsymbol{r})}$$

4.2.2 A physical model of propagation of a seismic pulse

This mathematical development is a strict definition of a model for wave propagation first proposed by Christian Huygens in the 17th century. It is possible to take the view that a propagating disturbance has a well-defined wave front at any time t and that it interacts with the medium only on the wave front. At every such point in the medium, energy is scattered in all directions. After an instant the new wave front is the envelope of all such scattering "wavelets", Huygens wavelets.

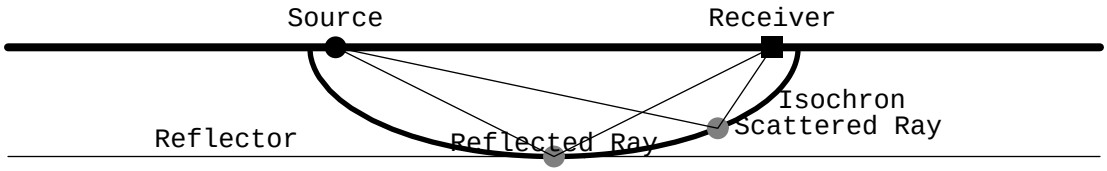
In seismology, a disturbance that travels in a fixed time from a source to a receiver could have been scattered as a Huygens wave from a scatterer that can have a multitude of positions. The collection of scatterers that lead to constant travel time are called an isochron. For travel in three dimensions, the isochron is a surface while in two dimensions the isochron is a curve.

Once the speed is determined as a function of position and direction, this scattering isochron can be calculated. In a medium of constant velocity the scattering isochron is obviously an ellipse with the source of the energy at one focus and the receiver at another. These isochrons are more complicated if the velocity varies with position or the elasticity tensor is anisotropic. If isochrons corresponding to different source receiver pairs-intersect, the intersection points define the times at which energy on the two records is scattered from the same point. If the scattering is from a reflector a little thought will show that the reflector must be the envelope of all the scattering isochrons that include a point on the reflector.

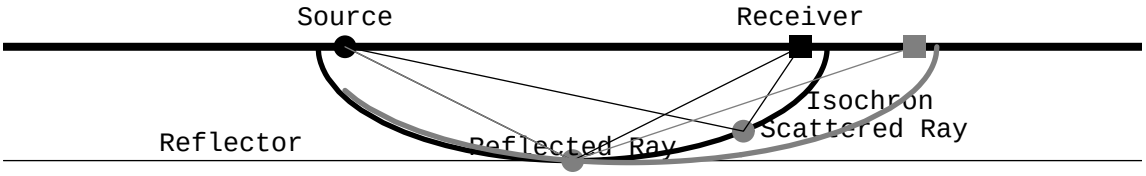
4.2.3 Single horizontal reflector

In order to clarify the concepts, consider the case when a single, horizontal, reflector exists within a medium in which velocity is generally a slowly varying function of position. Although the isochrons are not exactly ellipses it will be convenient to draw them as such.

The seismologist does not know of the existence of the reflector before doing measurements. If scattered energy is observed in a detector at time T, after the source was initiated it is possible that this energy was scattered from any point in the medium that lies on the isochron. If the isochron includes the actual reflection point the geometrical relations are easy to illustrate



but the seismic record at the receiver usually has signals that arrived at other times as well. Such signals were scattered from points defined by other isochrons. Consider now the effect of moving the detector of energy along the surface.



The isochrons for the second detector position generally intersect those generated for the first. In particular, one of the new isochrons intersects the reflection point for the first position and the energy seen at the time corresponding to the new receiver position could therefore have been scattered in part from the same point on the reflecting horizon. Since the signals from the same scattering point might reasonably be expected to be in phase, the scattered signal from this reflection point could be enhanced by "stacking" the two portions of the response together. It is elementary to show that for

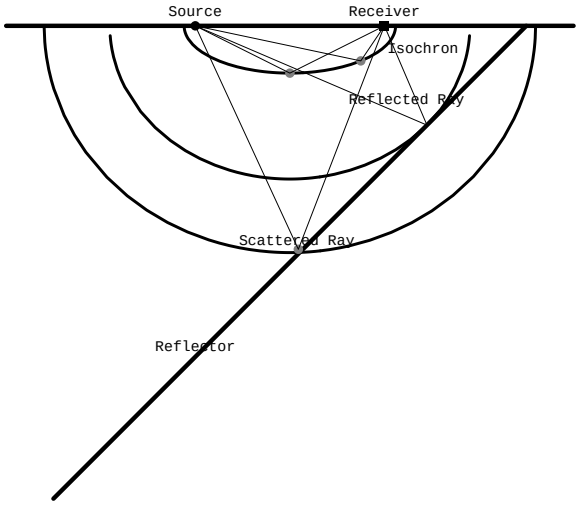
a layer of constant velocity c, the arrival time for energy scattered from a reflecting point at offset X from the source at the origin of coordinates is determined by

$$T^2 - \frac{x^2}{c^2} + 2\frac{xX}{c^2} = 2\frac{d^2 + X^2}{c^2}$$

Notice that the minimum travel time is at $x = X$, as it ought to be. When considered on a plot of signal as a function of detector position x and time along the recording T, this stacking curve is a hyperbola with its apex directly below the position of the reflecting point.

4.2.4 A single dipping reflector

The geometry changes if the reflector dips, or if the detector positions are on the side of a hill and the reflector is flat. Find the points at which the isochron intersects the reflector as a function of the path length. If these points coalesce you have the reflection point.



For a constant velocity layer the isochrons are still ellipses with foci at the source and receiver positions, but different isochrons now contain the reflection point. The nature of the stacking curve remains the same though because scattering from a point is assumed to depend only on the position of the scattering point, not the slope of the reflector. The equation of the scattering hyperbola is now

$$T^2 - \frac{x^2}{c^2} + 2\frac{xX}{c^2} = 2\frac{((d_0 - x\tan(\theta))^2 + X^2)}{c^2}$$

and the horizontal position at which the scattering path is a minimum is now

$$x = \frac{X + 2d_0\tan(\theta)}{1 + 2\tan(\theta)}$$

4.2.5 Vertical seismic profiling

It is much easier to place detectors on the surface of the earth, but if a well exists in the neighborhood of the target for the seismic investigation, it is perfectly reasonable to use detectors inside the well to measure travel times closer to the reflector. The isochrons are still ellipses but now one of the foci is on the surface while the other is beneath it. The stacking curves are now far from simple, particularly if the offset of the source from the well is significant. It is this problem that was explored at PIMSIPS 98.

4.3 Theoretical Approach

4.3.1 Constant velocity field

For a constant velocity field $v(z) = A$, we can formulate the Lagrangian of the total time taken by the acoustic wave to travel from source to receiver as

$$\begin{aligned}\mathcal{L} &= \int_a^0 \frac{\sqrt{1+\dot{z}^2}}{v(z)}\ dx \\ &= \int_a^0 \frac{\sqrt{1+\dot{z}^2}}{A}\ dx.\end{aligned}$$

The Euler-Lagrange equation is employed to minimize this Lagrangian function, yielding

$$\frac{\partial}{\partial x}(\dot{z}\frac{\partial \mathcal{L}}{\partial \dot{z}} - \mathcal{L}) = 0,$$

$$\frac{\partial}{\partial x}(\frac{-1}{v\sqrt{1+\dot{z}^2}}) = 0.$$

The vanishing derivative of this expression implies that this expression is constant, thus

$$E = \frac{-1}{v\sqrt{1+\dot{z}^2}} \Rightarrow \dot{z} = \sqrt{\frac{1-v^2E^2}{v^2E^2}} \Rightarrow \int_a^x \sqrt{\frac{1-v^2E^2}{v^2E^2}}\ dx = \int_{z(x)}^{z(a)} dz.$$

Evaluating at the point $(0, d)$ we have

$$E^2v^2 = \frac{a^2}{a^2+d^2}.$$

Our integral, with this substitution, is then evaluated as

$$\int_a^x \frac{d}{a}dx = \int_{z(x)}^{z(a)} dz \Rightarrow \frac{z(x)-z(a)}{x-a} = \frac{d}{a}.$$

This result confirms that the geodesic curve for a constant velocity field is a straight line. We can now assume that an acoustic wave that is reflected once within the medium will follow a path consisting of two straight lines, first from source to point of reflection, and then from point of reflection to receiver:

$$l_1 = \sqrt{(a-x)^2+z^2}, \quad l_2 = \sqrt{x^2+(d-z)^2}$$

The minimum time taken is therefore the sum of the time taken along both of these lines.

$$vT_{total} = l_1 + l_2,$$

$$vT_{total} = \sqrt{(a-x)^2+z^2} + \sqrt{x^2+(d-z)^2}.$$

For given a, d, v, T_{total}, this relationship forms an ellipse on the (x, z)-coordinate axis. By performing many experiments with different values of a and d, corresponding to different positions of the emitter and detector, we will get a range of values for T_{total}. We can then plot the family of ellipses arising from these data points and the profile of the reflective surface will be indicated by where short arcs of the ellipses are concentrated.

4.3.2 Linear velocity field

Determination of direct signal path

The direct signal, which travels from the source to detector in the least amount of time, will follow a path which can be determined by minimizing the functional which gives travel time as a function of path. In the case where the velocity field varies only with z, the problem is effectively two-dimensional. That is to say, if the signal starts in the (x,z)-plane, the path of the direct signal will not deviate from this plane. In this case the travel time functional as a function of path is:

$$T[x(z)] = \int_0^Z F(z,x,x')\,dz = \int_0^Z \frac{\sqrt{1+x'(z)}}{a+bz}\,dz, \tag{3.1}$$

where the path is described by $x(z)$ and the integration is from the surface at $z=0$ to the depth of the detector at $z=Z$. The Euler-Lagrange equation for a functional of this form is:

$$\frac{\partial}{\partial x}F(z,x,x') - \frac{d}{dz}\left(\frac{\partial}{\partial x'}F(z,x,x')\right) = 0. \tag{3.2}$$

In the case that $F(z,x,x')$ does not depend on x this simplifies to:

$$\frac{d}{dz}\left(\frac{\partial}{\partial x'}\frac{\sqrt{1+x'\,(z)^2}}{a+bz}\right) = 0, \tag{3.3}$$

or:

$$\frac{\partial}{\partial x'}\frac{\sqrt{1+x'\,(z)^2}}{a+bz} = \frac{x'\,(z)}{(a+bz)\sqrt{1+x'\,(z)^2}} = \text{constant}, \tag{3.4}$$

with boundary conditions determined by the source and detector configuration. The boundary conditions are simply that the signal starts from the source at $(X,0)$ and is received by the detector at $(Z,0)$. The first boundary condition is $x(0)=X$, since at the surface (where $z=0$) the signal starts at the source which is a distance X from the detector well. The second boundary is $x(Z)=0$, since the signal is received by the detector at depth $z=Z$ with no displacement from well, $x=0$. The above differential equation may be solved to yield:

$$x(z) = \frac{X}{2} - \frac{Z(2a+bZ)}{2bX} + \sqrt{\left(\frac{X}{2}+\frac{Z(2a+bZ)}{2bX}\right)^2 - \frac{z(2a+bz)}{b}} \tag{3.5}$$

which describes the path of the direct signal for the prescribed velocity field, and represents an arc of a circle.

4.3.3 Direct signal travel time

In order to calculate the travel time of the direct signal, the solution must be differentiated with respect to z and substituted into the travel time functional. The differentiation of equation 3.5 yields:

$$\frac{\partial}{\partial z}x(z) = -\left(\frac{a+bz}{b}\right)\left(\sqrt{\left(\frac{X}{2}+\frac{Z(2a+bZ)}{2bX}\right)^2 - \frac{z(2a+bz)}{b}}\right)^{-1}. \tag{3.6}$$

This is substituted into the integrand of equation 3.1 and the resulting expression is integrated along the ray path from source to detector to yield the time for the direct signal to arrive as a function of the X, and Z for given constants a and b. The integrand does not converge uniformly as X goes to zero, thus the case $X=0$ must be treated separately. The following is valid for $X=0$:

$$\begin{aligned} T(X,Z) &= \int_0^Z \frac{1}{a+bz}\,dz \\ T(X,Z) &= \frac{\log a+bZ}{b} - \frac{\log a}{b}. \end{aligned} \tag{3.7}$$

as for the case $X \neq 0$ (henceforth this will be the case, unless stated otherwise) we get:

$$T(X,Z) = \int_0^Z \frac{1}{a+bz}\sqrt{1+\left(\frac{a+bz}{b}\right)^2\left(\left(\frac{X}{2}+\frac{Z(2a+bZ)}{2bX}\right)^2-\frac{z(2a+bz)}{b}\right)^{-1}}\,dz$$

$$T(X,Z) = \frac{1}{b}\tanh^{-1}\left(\frac{\sqrt{(Z^2+X^2)(4a^2+4abZ+b^2Z^2+b^2X^2)}}{|2aZ+bZ^2+bX^2|}\right). \tag{3.8}$$

Thus for different sets of values for X and Z and measured direct signal times T, one could solve the resulting system of equations for a and b.

The previous equation can then be rewritten as:

$$T(X,Z) = \frac{1}{b}\tanh^{-1}\left(\frac{\sqrt{(Z^2+X^2)(Z^2+4Za/b+4a^2/b^2+X^2)}}{|2Za/b+X^2+Z^2|}\right),$$

which may be solved for b to yield:

$$b = \frac{1}{T(X,Z)}\tanh^{-1}\left(\frac{\sqrt{(Z^2+X^2)(X^2+Z^2+4Za/b+4a^2/b^2)}}{|2Za/b+X^2+Z^2|}\right)$$

Thus it is possible to determine both a and b.

4.4 Numerical Approach

In order to be able to generate pictures showing how the isochrons look for various velocity functions, the Maple symbolic manipulation package was used to perform the calculations.

Firstly, the positions of the sources and receivers was given and then a reflective surface was assumed, so the program could calculate the travel times to those detectors from the sources.

We imagine that the *only* information we have are the travel times, and positions of the sources and receivers as would be the case for a geophysicist. What can we say about the reflective surface that gave these travel times?

We know the reflection (supposing there to be only one reflection) must come from some point of the isochron for that traveltime for the given receiver and source. Maple calculates what the isochron looks like for each traveltime in a given velocity field and shows all the isochrons together.

Once we have all the isochrons shown, we can also display on the same diagram the original reflective surface and from this determine any relation between the two.

4.4.1 Method of solution

The equation for the reflector was taken to be a straight line given by the expression $z = 1500 - x$. The position of the source was at $z = 0$, $x = 635m$ and the receivers were taken to be at $x = 0$ and $z = 1200$, 1250, 1300, 1350 and $1400m$. The specific representative velocity profiles used were

$$v(z,\dot{z}) = \begin{cases} 2315 & \text{A: isotropic} \\ 2315\left(1+\dfrac{0.5}{1+\dot{z}^2}\right) & \text{B: anisotropic} \\ 2315+0.8z & \text{C: linear} \end{cases}$$

Isochronous curves for the three separate velocity profiles were generated using the Maple symbolic manipulation package. For a given profile the geodesics are determined by computing the class of curves C that minimize the total time of traversal

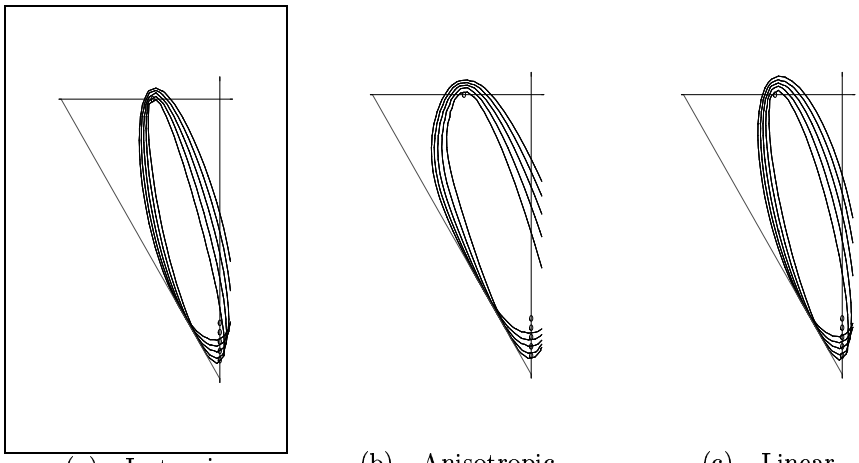
$$T = \int_C L(z,\dot{z})\,d\tau = \int_C \frac{(1+\dot{z}^2)}{v(z,\dot{z})}\,d\tau.$$

The curve C joins two points (x_0, z_0) and (x_1, z_1) and is describing using the parameter τ. Since $L(z, \dot{z})$ is independent of x, the Euler–Lagrange equation becomes

$$\frac{d}{dx}\left(\dot{z}\,\frac{\partial L}{\partial\,\dot{z}} - L\right) = 0.$$

In both the isotropic and anisotropic cases the geodesics are straight lines whereas in the case of the linear velocity profile, they are circular arcs. For each case, a parametric for of the geodesics is entered externally into the model. Once the geodesics are known, an expression is found for the total traveltime of traversal from the source to each of the five receivers where the condition that the signal hits the reflector only once is imposed. By parameterizing the equation for the reflector, local minima in the total travel time are found. The location on the reflector of these local minima correspond to the reflection point on the reflector.

To get the corresponding isochronous curves, one again considers the expression for the total traveltime where the point of reflection on the reflector is allowed to vary and the traveltime is the value determined above with the minimization procedure. A contour plot of this two parameter surfaces generates the isochronous curves. Isochronous curves for the three velocity profiles discussed, with a linear reflector, are plotted below.



(a) Isotropic (b) Anisotropic (c) Linear

Figure 4.1: Isochronous curves for the three profiles. These three plots show the isochronous curves and illustrate how the envelope of the curves reproduce the location of the reflector, for three different velocity profiles. For all cases the source is indicated at $635m$ and the five receivers are at depths of 1200, 1250, 1300, 1350 and $1400m$.

Clearly this technique can be extended to more complicated velocity profiles such as combining linear velocity dependence with angular dependence.

Lastly, it should be pointed out that for *isotropic* media (i.e. with no angular velocity dependence) the convex hull of the isochrons *must* be the reflective surface (for many isochrons). This is because for this type of media the angle of reflection is equal to the angle of incidence and can thus be considered to reflect off a line that is *perpendicular* to the normal defined by these angles. But this line is also the tangent to the isochron and thus the reflective surface must be along the collection of tangents of all the isochrons. This is the definition of a convex hull.

4.5 Conclusions

Both the greater accuracy of data acquisition/processing and increased scarcity of petroleum resources promote and necessitate rigorous investigations. Consequently, we have attempted to remove certain simplifying assumptions in order to increase the reliability of the seismic image. Firstly, we

Receiver	Depth of Receiver	Velocity profile	Contact position x	z	Travel time (ms)
1	1200	A	386.3	1113.7	663.9
		B	499.2	1000.8	594.6
		C	380	1120	539.0
2	1250	A	336.3	1163.7	669.0
		B	425.4	1074.6	603.7
		C	335	1165	541.9
3	1300	A	281.7	1218.3	674.6
		B	348.9	1151.1	613.0
		C	285	1215	545.4
4	1350	A	221.7	1278.3	681.0
		B	270.0	1230.0	622.7
		C	230	1270	549.5
5	1400	A	186.8	1344.6	687.9
		B	499.2	1313.2	632.7
		C	168	1332	554.4

Table 4.1: Summary of numerical results.

The three velocity profiles studied are A: isotropic, B: anisotropic and C: linear. The depth of the receivers and the contact position with the reflector are given in meters whereas the travel time of the signal is given in milliseconds.

have derived analytic expressions for the isochrons in an inhomogeneous medium characterized the constant vertical velocity gradient, and for the anisotropic medium characterized by the elliptical velocity dependence. The derived algorithms were coded and hence numerical as well as graphical results were generated. The results were generated for a single surface source and five wellbore receivers over a dipping interface. Three types of media were assumed: homogeneous/isotropic, homogeneous/anisotropic, and linearly inhomogeneous/isotropic. The traveltime and reflection-point results appear to be sensitive to a given assumption of the velocity field. In considering complex geological areas, e.g., foothills, one should examine several velocity fields used in data processing and quantify the difference in derived location of reflection points (e.g., see Table 4.1). Moreover, in such an area, the location of the planar reflector is tangent to the convex hull of isochrons and might not correspond to their intersection points since the two concepts need not coincide. Note that the latter method is commonly used in the Kirchhoff migration procedures.

The immediate further work appears to be the derivation of analytic expressions or numerical algorithms providing the range of positions of the subsurface features depending on assumed velocity fields. Notably, such a quantitative tool would play an important role in the drilling process. Also, the indication of likely scope of displacement in seismic mapping may yield an optimal drilling trajectory such as to successfully encounter the desired subsurface feature.

Several topics are investigated further for rigorous direct applications. For instance, R. Ait-Haddou and M.A. Slawinski are elaborating a formalism involving the notions of differential geometry, in particular, following the Finslerian approach via the local Minkowskian properties of this space. Notably, this method yields a geometrically fruitful context for inhomogeneous, anisotropic media by invoking the notion of continuously varying metrics. The initial results are prepared for the PIms PDF workshop in Vancouver (September 1998).

4.6 References

[1] M. B. Dobrin and C. H. Savit, *Introduction to geophysical prospecting*, McGraw Hill, 1988.

[2] E. Robinson, T. S. Durrani, and L. G. Peardon, *Geophysical signal processing*, Prentice

Hall, 1986.

[3] AKI, K. AND RICHARD, P.G. *Quantitative Seismology: Theory and Methods*, W.H. Freeman & Co., 1980.

Chapter 5

Trip Wire Detection for Land Mines

Chris Jessop[1], Shabnam Kavousian[2], Michael Lamoureux[1], Joshua Madden[3], Edo Nyland[4], Bruce Rout[2], Arunas Salkauskas[1], David Saunders[5], Bill Vetter[6], Satoshi Tomoda[1]

Written by Chris Budd[7] and John Stockie[2]

5.1 Introduction.

The safe detection of anti-personnel land mines is essential for the recovery of nations after an armed conflict. Land mines are often left as booby traps in populated areas, and so it is not appropriate to simply blow them up. Rather, the land mines need to be detected and then deactivated. A common mode of setting triggers for anti-personnel land mines is with trip wires which are hidden behind foliage or camouflaged in some other way. In order to minimise the risk to the people involved, a method is needed to identify suspect wires quickly and accurately. There are also many other practical situations where it is important to be able to detect wires of this sort; for example, helping light aircraft to avoid power cables.

The method of land mine detection being considered by ITRES Research Limited (which posed the problem for the PIMS Workshop) is to mount a detecting camera on a boom ahead of a slowly moving truck. The camera looks vertically downward at the ground beneath the boom and uses a charge-coupled device (CCD) detector to survey the resulting image[8]. In this detector, the image is acquired as a sequence of lines, which are constantly updated as the truck moves forward. Although the trip wires are often purposely camouflaged by being made of an optically transparent material, it is reasonable to assume that they are opaque in at least one of the scanned wavelengths. It is then hoped that an automatic algorithm (running in a well-protected computer on the truck) can be used to find the trip wires appearing in the images. If any wire (or suspect wire) is detected, then the algorithm will sound an alarm which will be a signal to the human operator to halt the truck and initiate a response by mine disposal experts.

[1]University of Calgary
[2]Simon Fraser University
[3]University of British Columbia
[4]University of Alberta
[5]University of Toronto
[6]Geospace Research
[7]University of Bath

[8]More information about this type of detector and its applications is available at ITRES' web site – `http://www.itres.com/`.

5.2 Detailed Problem Description.

The trip wire detection problem can be reduced to the following:

> *Can a computer algorithm be designed that will detect a trip wire in a single image, assuming that the wire can be recognised with the naked eye?*

A typical image with several wires hidden in foliage is given in Figure 5.1. In this picture there are

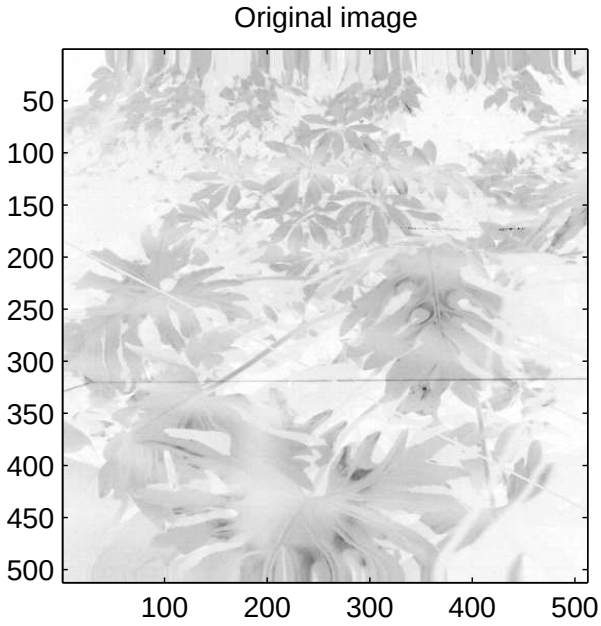


Figure 5.1: A test image provided by ITRES containing two trip wires and a third suspect linear feature.

two trip wires: one dark, horizontal wire located just below the centre of the image; and a second light-coloured one, obscured by the foliage, and running at an angle of approximately 30° to the first. We can see immediately some of the difficulties inherent in detecting a trip wire, namely

- wires can be partially covered by foliage;
- wires are not uniformly illuminated;
- many features, such as the stems of leaves, are also linear in nature and so might easily be mistaken for a wire;
- the image can be blurred or noisy, either making the wires difficult to detect, or introducing spurious linear features such as the dark, grainy feature at the upper right of the image.

The problem faced in this Workshop was to develop an algorithm for detecting a trip wire which satisfied the following conditions:

[1] it is *reliable*, and can detect a wire partially hidden and non-uniformly illuminated,

[2] it is *fast* and can process the image and detect the wire in real-time, before the advancing truck runs over the mine,

[3] it has a low rate of *false detections* so that it does not get confused between trip wires and leaf stems or other naturally-occurring linear features.

An algorithmic approach which has already been investigated by ITRES is one based on the *Hough transform*, which is a common technique used in image processing for detecting lines in images [3]. However, this method has apparently not been very satisfactory.

During the Workshop, we considered three algorithms for detecting trip wires:

- The first used a generalisation of the Hough transform called the *Radon transform*. To enhance the detection capability of this method, the image was first pre-processed in a way that exploits various properties of wires, which we discuss in Section 5.3.1. This method proved reasonably successful and the key aspect of implementing it turned out to be the development of an image-independent thresholding condition to distinguish between linear and nonlinear features.

- The second algorithm was based upon taking a series of one-dimensional slices through the image and looking for the movement of a sharp feature across this slice as the position of the slice was changed. This method was implemented, but was found to be very much affected by external clutter (such as foliage) and much less effective than the Radon transform method.

- The final approach was similar to the previous one, but attempted to take advantage of the fact that the CCD detector acquires the image in thin, horizontal strips. By looking at a sequence of contiguous strips, we attempted to locate the trip wire by looking for coherence from one slice to the next. This method suffered from the same problems as the previous one.

In this report, we will concentrate primarily on the Radon transform method, since it was the most successful of the three. When tested on several images, some supplied by ITRES and other generated artificially during the Workshop, it proved fairly effective in detecting hidden wires. The algorithm was implemented in MATLAB and proved rather slow. However, we believe that our implementation is far from optimal (indeed, there is a difference between what MATLAB claims it can do from the manual and what it does in practice!), and a dedicated implementation will certainly run *much* faster. We also believe that there is potential for a much faster implementation of the Radon transform that makes use of the fact that the image is updated one line at a time.

5.3 Identifying Linear Features using the Radon Transform.

Suppose that we look at a two-dimensional image which is made up of pixels. The image intensity can be regarded as a function $u(x, y)$ of the position (x, y) in the image. Before describing the Radon transform and the algorithm based on it, we first need to define what is meant by a trip wire, or a *linear image feature*, in a more mathematical way.

5.3.1 What is a wire?

We identify three characteristics of the intensity function $u(x, y)$ which make a wire distinct from the surrounding background:

[1] A wire is *narrow* and represents a portion of the image in which the function $u(x, y)$ has a *high curvature*; that is, where $u(x, y)$ changes very rapidly.

[2] A wire is locally straight and presents a *sharp edge*.

[3] A wire can be partially hidden by shade, foliage or other natural obstacles, and as such it consists of *a sequence of co-linear line segments*, rather than an uninterrupted, straight line.

Although one or two of these properties may be satisfied by other natural features such as the stem of a leaf, the combination of the three appears to define a wire uniquely.

5.3.2 The Radon Transform.

The Radon transform of a 2D image $u(x, y)$ is commonly written as [7]

$$U(\rho, \theta) = \int_{\mathbb{R}^2} u(x, y) \cdot \delta(x \cos\theta + y \sin\theta - \rho)\, dx\, dy \tag{3.1}$$

with ρ and θ defined as shown in Figure 5.2 We assume that the origin for the measurement of ρ is always chosen to be the centre of the image. The advantage of using the Radon transform for the

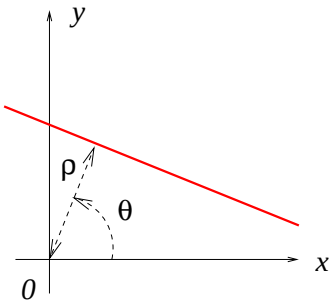


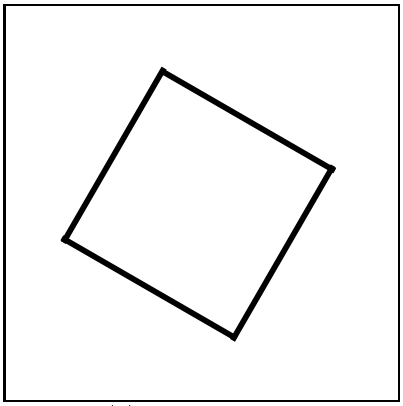
Figure 5.2: Definition of ρ and θ in the Radon transform.

detection of a trip wire is that a linear feature in the image, approximated by the straight line with equation

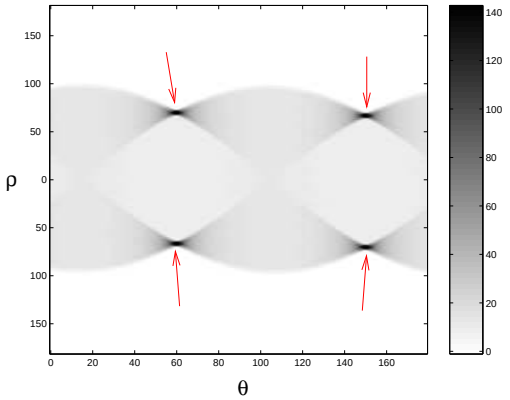
$$x \cos\theta + y \sin\theta = \rho,$$

corresponds to a point where $U(\rho, \theta)$ is large. Furthermore, the integral operation in (3.1) is insensitive to whether the ray along which the integration is performed has a linear feature that is contiguous or broken. The maximum value of $U(\rho, \theta)$ should thus correspond to the "most significant" linear feature in the image.

As an example, consider the simple test image in Figure 5.3(a), consisting of four line segments arranged in the shape of a square. The Radon transform clearly indicates four peaks corresponding



(a) Test image.



(b) Radon transform.

Figure 5.3: A simple test image and its Radon transform.

the four sides of the square: two at an angle of $\theta = 60°$ (with approximate offsets $\rho = 65$ and -65 from the centre of the image) and two more at an angle of $\theta = 150°$ (with the same offsets).

5.3.3 The algorithm.

The Radon transform is well-known for its ability to pinpoint linear features even in very noisy images [6],[2]. However, for the images supplied by ITRES, we found that the transformed intensity $U(\rho, \theta)$ often failed to detect real wires, and also exhibited spurious maxima totally unrelated to trip wires that were clearly visible to the naked eye. Therefore, the Radon transform by itself is not appropriate for trip wire detection.

The question now is this: *How can we exploit the three characteristics of a "linear feature" listed in Section 5.3.1?* The Radon transform makes use of only the third item in our list, by identifying multiple segments that lie on a single straight line. We will now describe an algorithm that uses two pre-processing steps which exploit the fact that trip wires are associated with edges, and high image curvature. Starting with a two-dimensional image $u(x, y)$, with values in the range $[0, 1]$, we proceed as follows:

Trip Wire Detection Algorithm.

---

1. Apply a *Laplacian filter*

$$L = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

to each point in the raw image, which is analogous to taking the second derivative of the image, and is designed to accentuate regions of high curvature.

2. Perform *edge detection* on the filtered image (we use the *Sobel* method in most cases since it is the default in MATLAB).

3. Apply the *Radon transform.*

4. *Threshold* the transformed image, by looking for all points (ρ, θ) with $U(\rho, \theta) > T$ for some threshold T, in order to differentiate candidate lines from surrounding clutter.

5.* (for testing only) Apply the *inverse Radon transform* to locate candidate trip wires, and then compare them manually with the original image.

Step 5 is not required, but is very useful in terms of testing the success of the algorithm to see which linear features are recognized. In an actual mine detection scenario, a transformed image which has a value that exceeds the threshold T in Step 4 will set off an alarm that causes the human operator to stop and take action. The main conceptual difficulty in implementing this algorithm is the selection of an appropriate threshold, which will be described in detail in the next section.

This algorithm has the advantage of eliminating background intensity variations and bright spots and identifying linear features which may be both noisy and disconnected in the original image. We experimented with leaving out individual steps, but our overwhelming feeling is that steps 1–4 are all required for an effective algorithm.

The algorithm was very straightforward to implement in MATLAB, since all the components (except the "optional" inverse Radon transform) are built-in procedures in MATLAB's *Image Processing Toolbox* (see the Appendix for a listing of the code). We observed that steps 1, 2, 4 and 5 are all reasonably fast and that most of the computational time is taken up by the Radon transform in Step 3. This is partly because the MATLAB implementation is very slow and does not use the very important result that the Radon transform can be computed rapidly by using a Fast Fourier Transform

(FFT) [1]. Even though the FFT algorithm is mentioned in the MATLAB documentation [5], it is not implemented in the current version of the *Toolbox*. Communications with a technical representative of The MathWorks Inc. have confirmed this fact.

We also implemented the trip wire detection algorithm using another set of routines in the Kansas University Image Processing Library (KUIM) [4], but found them to be equally slow.

5.3.4 Approaches to thresholding.

As mentioned before, the choice of an appropriate threshold in Step 4 of the algorithm on page 53 is of paramount importance. When applying the threshold T, the only points (ρ, θ) retained in the transformed image are those satisfying $U(\rho, \theta) > T$. This produces a relatively small set of points (if T is chosen appropriately), and so inverting the thresholded image to give a set of candidate trip wires is then a fast and easy procedure (indeed it is very much faster than doing a direct inverse Radon transform).

There are several aspects of the thresholding problem that make the choice of T particularly problematic:

- If the threshold level is chosen too high, then no wires are found; alternately, if T is too low then too many false wires will be detected. Thus, there should be some "best" intermediate value in which only true trip wires are detected.

- The Radon transform integral (3.1) can be interpreted in the sense of Riemann sums as summing image intensities along linear slices. Therefore, an image that has bright patches throughout will have much larger transform values than an identical image that has no such patches. Some sort of scaling of the transformed image U based on average or maximum intensity is hence necessary if the linear features in both of these images are to be detected at the same threshold level.

- Equation (3.1) should also be insensitive to the dimensions of the image, and so U may also have to be scaled by some size factor. For example, if a trip wire is detected in an $M \times N$ image, then the same threshold should be sufficient to detect the wire in a smaller sub-image that fully encompasses the wire. In fact, the MATLAB implementation of the Radon transform assumes the pixels have dimension 1 square unit (that is, $\Delta x = \Delta y = 1$), and so a rescaling based on image size *is necessary*.

We now propose several alternatives for scaling the transformed image. Assume for the remainder that the image intensity values $u(x, y)$ lie between 0 and 1[9], and that the image has dimensions $M \times N$ pixels. Our approach to thresholding will be to look for points in a *rescaled* Radon transform $\tilde{U}(\rho, \theta) = U(\rho, \theta)/\alpha_u$ that satisfy $\tilde{U} > T$, and where α_u is a scaling factor that depends on the image. Based on the above observations, we will consider the following choices for the scaling factor:

S-1	no scaling	$\alpha_u = 1$
S-2	scale by the maximum of U	$\alpha_u = \max(U)$
S-3	scale by the average of U	$\alpha_u = \text{mean}(U)$
S-4	scale by the number of pixels	$\alpha_u = M \cdot N$
S-5	scale by the largest image dimension	$\alpha_u = \max(M, N)$

The reason for including **S-5** is that the **S-4** scaling actually serves to distort pixels in non-square images, while **S-5** weights both directions the same regardless of image size. The five different scalings will be applied in the next section to a variety of images. Our hope is to find an "ideal," universal threshold that is independent of image properties, and does not have to be tuned for a particular image.

[9] The image at this stage has already been through the edge detection process, and so is already a binary (0/1) image.

5.3.5 Results.

At the start of the Workshop, we were provided with a single digital trip wire picture (Image 5), with a well-defined wire passing along its entire length. To help in the testing process, we created several additional images based on the given one, with "artificial wires" having different intensities, orientations and noise levels. One such image, which we label Image 5c, was the one pictured earlier in Figure 5.1. The output of the trip wire detection algorithm for this image is pictured in Figure 5.4.

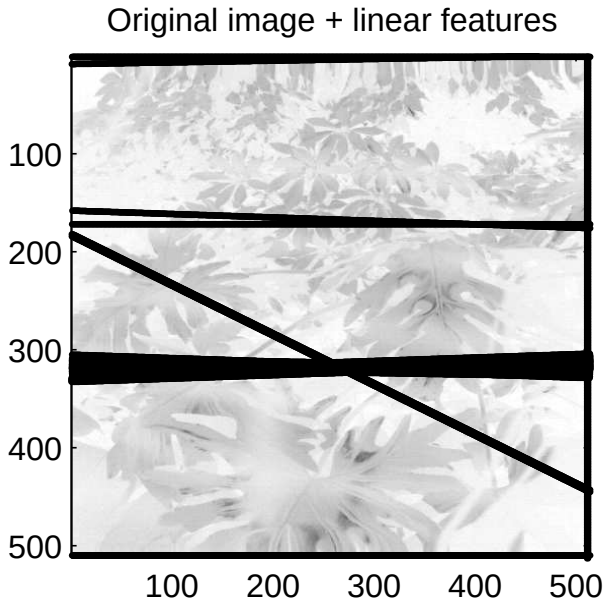


Figure 5.4: Sample inversion for Image 5c, with $T = 2.4 \times 10^{-4}$ and scaling **S-4**. The lines along the edges of the image are a result of the boundary conditions on the Laplace filter, and can be ignored.

Scaling S-1, S-2 and S-3:

The original image has three linear features: one an obvious bright trip wire running horizontally across the bottom of the image, one artificially created dark trip wire running obliquely from left to right, and a third noisy horizontal bright wire at the upper right. Using any of these three scaling methods, all of the linear features could be identified. A series of other images were constructed manually from this one, that contained other trip wires and/or more noise. Scaling **S-3** produced the best results, but the threshold had to be tuned for each individual image.

Scaling S-4:

Table 5.1 summarises the number of linear features found by the algorithm with scaling **S-4** compared to what we'd like to find (in the "Actual" column). It is clear that the rescaling based on image

[d] With the addition of noise in these images, the feature in the upper right of the image was much less coherent and thus hard to detect.

[e] This image has a lot of noise compared to the original one, which is what affects the $\max(\tilde{U})$ value. The artificial dark line is not picked out until the threshold is reduced to 2.1E-4. And there are also some other bogus vertical linear features picked out on the side of the image at this threshold.

[f] For each of these three subimages, the threshold detects far too many linear features to be useful.

[g] This is another noisy image like Image 5a, but the trip wire detection works better in this case.

Image	Actual	# of linear features found with threshold $T =$			$\max(\tilde{U})$
		2.4E-4	3.0E-4	5.0E-4	
Image 5 (512×512)	2	2	1	1	8.4E-4
Image 5a (512×512)	2-3[d]	1[e]	1	0	3.1E-4
Image 5b (512×512)	3	3	1	1	8.5E-4
Image 5c (512×512)	3	3	2	1	8.5E-4
200×512 sub-image	2	–[f]	–	–	2.4E-3
150×512 sub-image	2	–[f]	–	–	2.8E-3
100×512 sub-image	2	–[f]	–	–	4.2E-3
Image 5d (512×512)	2-3[d]	2[g]	2	0	3.2E-4
Image 5e (512×512)	3	2	2	1	8.5E-4

Table 5.1: How many linear features are recognized in each image, using the **S-4** scaling? The sub-images of Image 5c are horizontal strips of width 512 pixels, centered vertically on the main trip wire.

size seems to be fine for square images of the same size. There is a fairly consistent value of the threshold $T \approx 2.4 \times 10^{-4}$ which identifies the linear features we're looking for.

However, the **S-4** scaling is not adequate when considering non-square images (which are probably more typical of what would be obtained from a CCD camera). This can be seen by looking at the last column, where the maximum value of the scaled Radon transform has significant variation with image strip width N. Therefore, we cannot expect the thresholding in this case to be independent of the image size.

Scaling S-5:

Table 5.2 summarises the results for Image 5c using the scaling **S-5** with $\alpha_u = \max(M, N)$. The key thing to notice is that the $\max(\tilde{U})$ column is almost constant for the various dimensions of sub-image.

Image	Threshold $T =$			$\max(\tilde{U})$
	0.1	0.12	0.2	
Image 5c				
full 512×512 image	–	3	1	0.433
200×512 sub-image	2	1	1	0.492
150×512 sub-image	1	1	1	0.430
100×512 sub-image	1	1	1	0.427

Table 5.2: How many linear features are recognized in each image, using the **S-5** scaling? The sub-images of Image 5c are horizontal strips of width 512 pixels, centered vertically on the main trip wire, which should optimally show two linear features.

Here a threshold value T lying in the range 0.12 to 0.2 seems to reasonable (although it may be taken as large as 0.4 and still catch the main horizontal wire). While the second oblique feature in the three subimages does not get detected in this range of T, it is important to ask ourselves whether this is more important to detect *all* suspect linear features, or just to signal an alarm that there is *at least one* suspect wire.

One final observation: as the strip is reduced in height, the oblique feature no longer runs the length of the image and so is much harder to detect. This spells trouble for detection of trip wires where we have long, narrow images in which the wire does not run horizontally.

Idealised images.

To verify some of our previous conclusions we also generated a set of artificial images of wires of various types to check how the detection algorithm works on non-square images in an idealised situation. Four 300×300 images are pictured Figure 5.5.

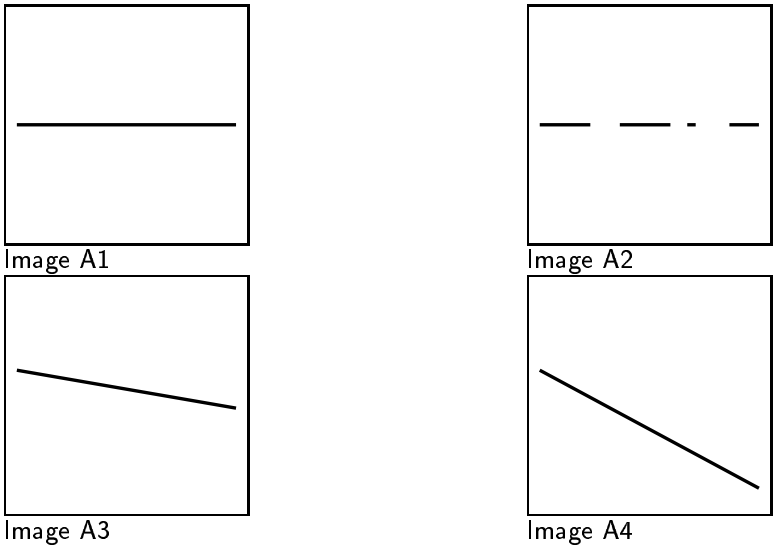


Figure 5.5: The four "idealised wire" images.

The wire detection algorithm is run on all four images with a threshold of $T = 0.2$ and scaling **S-5**, and the results are summarised in Table 5.3. Again, the scaling **S-5** with a threshold value of

Image	$\max(\tilde{U})$ for sub-image of size ...			
	300×300	150×300	75×300	38×300
A1	0.68	0.68	0.68	0.67
A2	0.47	0.47	0.47	0.45
A3	0.54	0.53	0.53	0.42
A4	0.62	0.42	0.26	0.16 (*)

Table 5.3: Maximum scaled Radon transform for the four artificial wire images, with scaling **S-5**. The "(*)" indicates that threshold with $T = 0.2$ would fail.

0.2 works well, except in the case of a wire at a large angle to the horizontal (Image A4), where it may fail if the image window is too narrow.

Difficult images.

Finally, we report on a second set of particularly "difficult" images obtained from ITRES near the conclusion of the Calgary Workshop. They comprised a sequence of 6 images, each taken in 5 filter bands (and named `file`m`_b`n`.tif`, with m corresponding to the image number, and n the band).

The results for this set of images were somewhat disappointing. The images taken in bands $n = 4$ and 5 were the clearest, and yet we were still unable to recognize the wire in any of these when the entire 900×461 image was processed. However, if we took the original image and grabbed a sub-image as small as 200×461 (pictured in Figure 5.6), then the wire can be detected, provided the edge detection method was changed. The default `sobel` edge detection option used in MATLAB does not work here, and we needed to use the `log` option instead (see Figure 5.7). After running

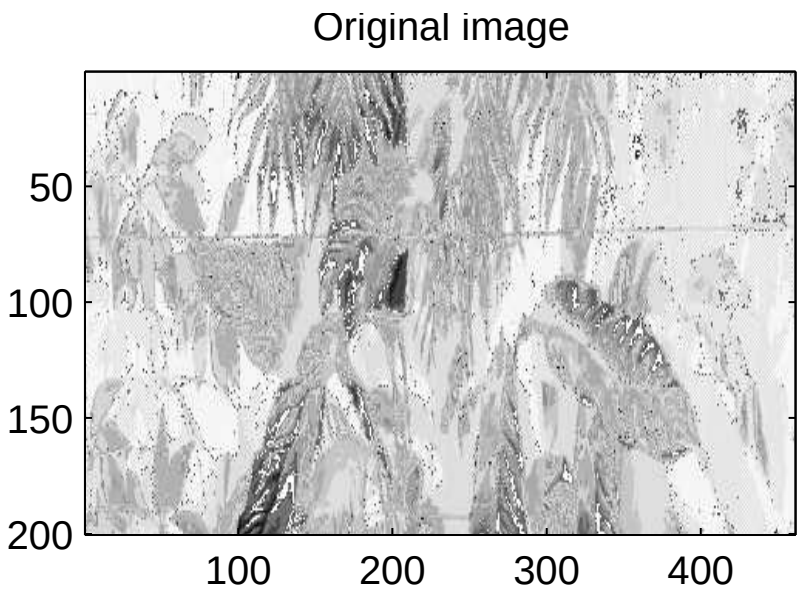


Figure 5.6: The top 200×461 portion of the image `file5_b4.tif` (original size: 900×461).

the inverse Radon transform on the `log` edge detected image, we obtain the plot in Figure 5.8. In constrast with the original set of images, the inversion here is very sensitive to the choice of threshold: $T = 0.25$ finds a large number of spurious linear features, while $T = 0.30$ fails to locate any wires. While this level of threshold is too high to detect *all* of the linear features from the set

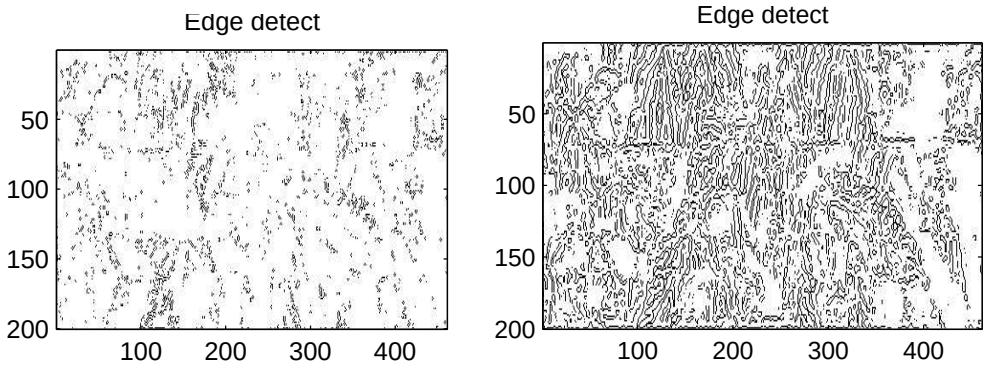


Figure 5.7: Edge detected image using options `sobel` (left) and `log` (right).

of "difficult" images, it is still sufficient to detect the major ones in Image 5c, for example.

We now summarise by considering the possible reasons for the failings of our algorithm with this last set of images:

- First of all, we may need to restrict the size of the image to a certain minimum, say $\max(M, N) < 500$. When using smaller subimages, the algorithm seems better able to detect individual linear features. It may be due to the fact that chance alignment of points in the edge-detected image is much more likely when the image is larger. Indeed, if we are to process CCD camera-generated images in real time, then we will likely be running on thin-strip type images anyways!

- second, these "difficult" images contain more noise than the others, since the wire appears to have a much more grainy texture. This seems to be causing a problem in the edge detection

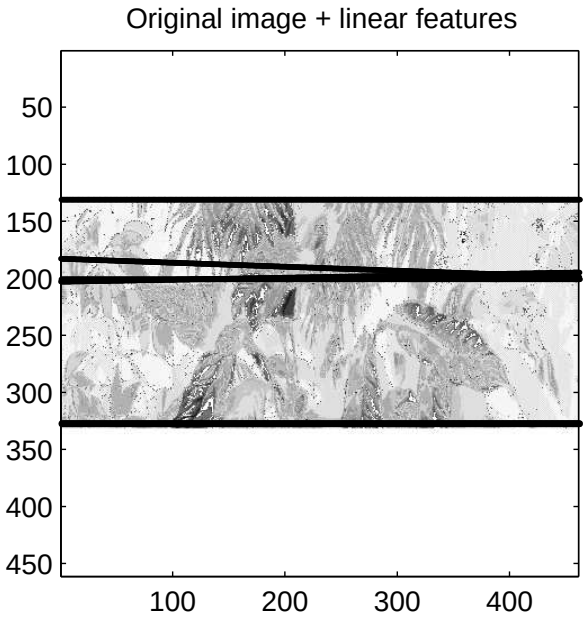


Figure 5.8: Inverse transform of the `log` edge detected image with $T = 0.27$ and scaling **S-5**.

step (even with the `log` results from Figure 5.7, where the wire is not as well resolved by the edge detection). Perhaps there is another edge detection algorithm that will deal better with grainy lines by "smoothing" the edge detected image.

Preliminary tests suggest that the choice of `log` edge detection does not significantly affect the results from the other sets of images considered provided that the threshold is increased to $T = 0.30$. Also, the results seem more sensitive to the choice of threshold than with `sobel` edge detection. Clearly, some more work is required here, to find an approrpriate edge detector and determine an image-independent threshold.

5.3.6 Speeding up the algorithm.

To get an idea of how fast the algorithm needs to be in order to process the images in real time, we can use the following information provided by ITRES:

- the CCD camera acquires one image line in 20 ms, and so the image is acquired at a rate of 50 rows per second;
- the truck on which the camera is mounted moves relative to the ground at 0.25 km/h or 7 cm/s;
- the boom length is on the order of 1 m;
- approx. 10 seconds to recognize a trip wire from image acquisition to danger zone

The algorithm as described was implemented directly in MATLAB using the default Laplacian filter, Sobel edge detector and Radon transform. For an image with N^2 pixels the MATLAB Radon transforms used order $\mathcal{O}(N^4)$ operations which is very inefficient and will make it very hard to use the method in real time. In contrast, methods based upon the FFT (as advertied in the MATLAB manual, but not implemented in the code) should take $\mathcal{O}(N^2 \log(N))$ operations. In an article by Brady [1] a direct algorithm (i.e. not one using the FFT) is proposed for the Radon transform which also takes $\mathcal{O}(N^2 \log(N))$ operations. Both of these methods are worth implementing.

It is possible that by exploiting special features of the image we can produce an algorithm which is $\mathcal{O}(N^2)$. To do this we observe that in detecting a trip wire we do not look at one image, but rather

a *sequence* of images which are produced as the truck moves forward. Given an original reference image, the next image will be formed from this by adding one new row of N pixels and deleting a row of N pixels. Furthermore a straight line in the original image can be clearly extended to a straight line in the new image. Thus to find the projection of the new image onto such a line we can take the projection of the reference image onto this same line, delete the points from the pixels to be discarded and add on the points from the new line of pixels. For any given line we would only have to add on a bounded number of pixels and there are $\mathcal{O}(N^2)$ such lines. Thus the Radon transform of the new image can be obtained from that of the reference image with $\mathcal{O}(N^2)$ operations, which is *much* faster than taking the Radon transform of the new image without reference to the earlier image. This method has some messy book-keeping associated with it (for example the offset ρ will change when considering the same line in the reference and new images, although its angle θ will not) but should be not too difficult to implement. We did not attempt to do this during the course of the Workshop.

5.4 Other Methods.

We now return briefly to the two other methods for trip-wire detection mentioned earlier that relied on taking a series of one-dimensional slices of the image.

Suppose that we have an $M \times N$ image of M horizontal and N vertical pixels of an image $u(x, y)$. Each of the horizontal lines of M pixels at $N = n$ gives a one-dimensional sequence $v_n(x)$ which can be processed. A trip wire at an oblique angle should in principle give a localised maximum in the function $v_n(x)$ at a point $x(n)$ that depends upon n. A feature of the linearity of the wire is that $x(n)$ should be a *linear* function of n.

This hypothesis was tested by taking a series of horizontal slices of the test images described in the earlier sections. When applied to the original image the results were very inconclusive. Indeed, the wire did not show up very well as a localised maximum and was very easy to miss. To improve the resolution, the images were pre-processed as before by applying a Laplacian filter to emphasise the curvature. When horizontal slices were taken from the filtered images, the wire showed up much more clearly, and for certain examples under certain conditions the motion of the maximum from one horizontal slice to the next showed up quite clearly, but were still in general rather hard to detect. Attempts to automate this process (for example, by stringing together a series of one-dimensional slices and using a spectral analysis of the resulting string) were not very effective.

These procedures seem less certain and much harder to automate than the Radon transform methods.

5.5 Summary and Future Directions.

In summary, the pre-processed, Radon transform-based detection method seems the best one to use. There are some issues that arose while testing the method:

- On the set of "difficult" images, the wire detection was most problematic. The `log` method of edge detection seemed better than the default `sobel`, but it is still worth considering other edge detection algorithms that might deal more effectively with noisy or grainy images.

- The algorithm works best when the wire is horizontal and spans the length of the image. If we are constrained to thin-strip type images, then it may not be possible to detect oblique wires reliably. A simple solution is to run the algorithm twice, once with the image and a second time with the same image rotated through 45°.

More investigation is clearly needed on different approaches to edge detection and thresholding. If the image-dependency in these aspects of the algorithm can be eliminated, then a fast, FFT-based implementation will make real-time, automated detection of trip wires a possibility.

Acknowledgments

We would like to express our appreciation to ITRES Research Limited for posing this problem, and for their help throughout the Workshop.

Appendix A: MATLAB Code

The MATLAB code for reading an image and applying steps 1–3 of the trip wire detection algorithm is included below. The thresholding and inverse Radon transform are performed within the function `plotinvr.m`, which can be downloaded from the Web site

`http://www.jms.sfu.ca/radon.html`.

All of the other code and test images used in generating the results in this report are also available at this site, along with a complete description.

5.5.1 Main program.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% Main Program:
%%%   Perform the initial stages of the trip wire detection
%%%   algorithm using a combination of Laplacian filter,
%%%   edge detection, and Radon transform.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%   Read the image and normalise the intensities
%%%   to the range [0,1]:
I    = double( imread( 'images/VEG5C.TIF' ));
Iraw = I;         % save original image
I    = (I - min(I(:))) / (max(I(:)) - min(I(:)));

%%%   Transpose the image so the number of columns is largest:
[nx, ny] = size(I);
if ny < nx
  I = I', [nx, ny] = size(I);
end;

%%%   Filter image using Laplacian:
h = [-1 -1 -1; -1 8 -1; -1 -1 -1];
Ilap = filter2( h, I );

%%%   Edge detection (options are log, prewitt,
%%%   roberts, zerocross and sobel (the default)):
edgealg = 'sobel';
Iedge = edge(Ilap, edgealg );

%%%   Compute the Radon transform:
th = 0:179;
[R, xp] = radon(Iedge, th);

%%%   Threshold and invert the Radon transform:
plotinvr( I, R, th, xp, 0.2, 5, 50 );
%%%   END.
```

5.6 References

[1] M. L. Brady, "A fast discrete approximation algorithm for the Radon transform," *SIAM Journal on Computing*, 27(1):107-119 (1998).

[2] A. C. Copeland, G. Ravichandran and M. M. Trivedi, "Localized Radon transform-based detection of ship wakes in SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, 33(1):35-45 (1995).

[3] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Communications of the ACM*, 15(1):11-15 (1972).

[4] *The KUIM Image Processing Library*, Department of Electrical Engineering and Computer Science, Kansas University, 1998. `http://www.tisl.ukans.edu/ jgauch/kuim/kuim.html`.

[5] MATLAB *Image Processing Toolbox Manual, Version 2.0*, The MathWorks Inc., 1997.

[6] L. M. Murphy, "Linear feature detection and enhancement in noisy images via the Radon transform," *Pattern Recognition Letters*, 4:279-284 (1986).

[7] P. Toft, *The Radon Transform*, `http://eivind.imm.dtu.dk/staff/ptoft/Radon/Radon.html`.

Chapter 6

Torsion in Multistrand Cables

Ainul Akhtar[1], Ben Aggarwala[2], Sean Bohun[3], Alistair Fitt[4], Huaxiong Huang[5], Christina Stoica[3], Rex Westbrook[2]

Report written by Rex Westbrook

6.1 The Problem

Multistrand cables are used widely in industry. Overhead electrical conductors, wire ropes and suspension cables in mine hoists are typical examples. The simplest construction involves a straight core wire surrounded by concentric layers of strands, while more complex cables involve a stranded core. (See Figure 6.1.) All strands in a given layer remain twisted about the cable axis at a fixed angle known as the layer angle, as in Figure 6.2. Adjacent layers remain twisted in opposite directions to one another, clockwise or anticlockwise in order to minimize torsion when the cable is loaded axially. Such a construction involves interaction between strands in contact.

A simple model, developed recently by Lanteigne and Akhtar [1], predicts the maximum failing load and torsion of the cable using data on the wire strands as the input parameters. That model has made the simplifying assumption that no frictional interaction occurs between adjacent strands. Experiments carried out by Akhtar and Lanteigne [2], are in excellent agreement with the predictions of the model for multistrand conductors made with aluminum alloy strands and for the measured values of true tensile stress for all cables. However, the model predicts torsion values for cables containing galvanized steel strands that deviate substantially from those measured experimentally. It has been concluded [2] that interstrand and frictional interaction does not occur between strands

[1]Powertech
[2]University of Calgary
[3]University of Victoria
[4]University of Southampton
[5]U.B.C. and PIMS

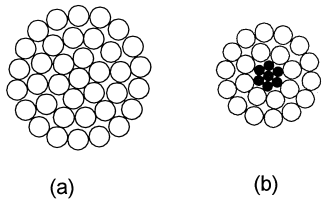


Figure 6.1: Cross-section of a multistrand cable

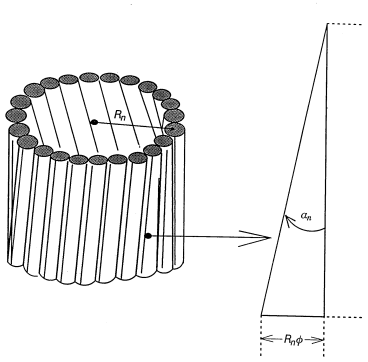


Figure 6.2: Single layer and angle of strands

made of aluminum and aluminum alloys and that frictional interaction occurs when galvanized steel strands are used for the construction of the multistrand cable.

Proposed objective for the PIMS Workshop 98: *Modify the existing model [1] or develop an alternative model to predict torsion in multistrand cable involving interstrand friction.*

Dr. Akhtar made the programme "conse" used in the calculation of [1] available to us as well as data for various sample cables in [3].

6.2 Solution Strategy

The initial discussion focussed on the consideration of possible frictional forces that must occur between the layers. Newton's Third Law however says that these will occur in equal and opposite pairs and both the resultant force and resultant moment of such pairs will be zero.

More subtle effects of the frictional forces will be the redistribution of the forces and moments on the ends of the cable where the boundary conditions are those of fixed displacement. A simple thought experiment for the case of wires of different materials suggested that load would be shifted from the stiffer to the less stiff material by the friction. A naive adjustment based on this idea was made for the programme "corse" and the result checked for one of the examples. A change of the correct sort occurred but clearly any further effort in this direction would require more detailed work and would not solve the problem in the case of materials of the same kind such as all steel cables. This idea was left for possible later consideration.

The method which was finally used amounted to the introduction and distribution of "body couples" between the layers. These body couples were taken to be proportional to the normal force between layers with a coefficient of friction as the constant of proportionality. The coefficient of friction μ was assumed to depend only on the materials in the two adjacent layers; that is, aluminum–aluminum, aluminum–steel, or steel–steel. The values of these coefficients of friction became free parameters in our numerical experiments. Because the theory and experiment agreed in the all aluminum cable we made $\mu = 0$ in this case.

The formula and its derivation are given in Section 6.3. An adjustment was made to the programme "conse" and numerical experiments were carried out .

First we treated the all steel cable so that only one free parameter was available. A value was calculated to fit one set of data and the same parameter was used with other data sets. This method was repeated for the aluminum–steel wire "Peace." Results of these numerical experiments are given

in Section 6.4.

6.3 Changes in the Theoretical Model

In [1], Lanteigne and Akhtar obtain the following formulae for the incremental changes δF, δT in the axial force and the torque due to the incremental change $\delta\left(\frac{u}{l}\right)$ in the strain (these are adjusted here to the special case when there is no external twist)

$$\delta F = (AE)\delta\left(\frac{u}{l}\right) \tag{3.1}$$

$$\delta T = C\delta\left(\frac{u}{l}\right) \tag{3.2}$$

where

$$AE = \sum_{n=0}^{N} K_n A_n \cos^3\alpha_n \left(\frac{d\sigma_n}{d\varepsilon_n}\right) \tag{3.3}$$

$$C = \sum_{n=0}^{N} K_n A_n R_n \sin\alpha_n \cos^2\alpha_n \left(\frac{d\sigma_n}{d\varepsilon_n}\right). \tag{3.4}$$

In these formulae, K_n is the number of strands in the layer n, A_n is the cross sectional area of a strand in layer n, R_n is the radius of the helix for a strand in layer n, α_n is the lay angle for layer n and $\frac{d\sigma_n}{d\varepsilon_n}$ is the slope of a specified stress strain curve at the current value of the strain $\varepsilon_n = \left(\frac{u}{l}\right)\cos^2\alpha_n$. The increment of the force f_n along the axis of the helix is $df_n = A_n\left(\frac{d\sigma_n}{d\varepsilon_n}\right)d\varepsilon_n$. To account for the observed discrepancies between theory and experiment the value of C should be increased.

Note that α_n the lay angle is positive when the helix is anticlockwise and negative when it is clockwise. This means that the value of C is positive when the outer layer is anticlockwise and negative when it is clockwise.

To find the normal force per unit length between layers we use the normal component of the force due to a tension for the helix. This is $\kappa_n f_n$ directed toward the central axis of the helix where κ_n is the curvature. In terms of the lay angle α_n and the radius R_n of the helix the curvature is given by $\frac{1}{R_n}\sin^2\alpha_n$, thus the normal force is $\frac{f_n\sin^2\alpha_n}{R_n}$ per unit length.

This force would represent the normal force due to the n^{th} layer on the $n-1^{st}$. If we denote the outer layer as the N^{th} then the total normal force between the $n+1^{st}$ and n^{th} layers is $\sum_{k=n+1}^{N}\frac{f_k\sin^2\alpha_k}{R_k} = G_n$. (Equivalently, $G_{n-1} = G_n + \kappa_n f_n$, with $G_N = 0$).

We now consider a strand in the n^{th} layer, as shown in Figure 6.3. It is assumed that the normal forces G_n give rise to frictional forces $F_n = \mu_n G_n$ as shown. The contribution of these forces to the total torque is estimated as follows.

The resultant tangential force is taken to act at the centre of the layer and its moment is then

$$(F_{n-1} - F_n)R_n = (\mu_{n-1}G_{n-1} - \mu_n G_n)R_n.$$

The contribution to the torque for the whole layer is thus $(\mu_{n-1}G_{n-1} - \mu_n G_n)K_n R_n$. These contributions are summed over all layers to give the total torque.

The incremental change in the torque is then related to the incremental change in strain $\delta\left(\frac{u}{l}\right)$ by $\delta T = (C + C')\delta\left(\frac{u}{l}\right)$ where C is the constant defined in Equation 3.4

$$C'\delta\left(\frac{u}{l}\right) = \sum_{n=1}^{N}(\mu_{n-1}\delta G_{n-1} - \mu_n\delta G_n)K_n R_n$$

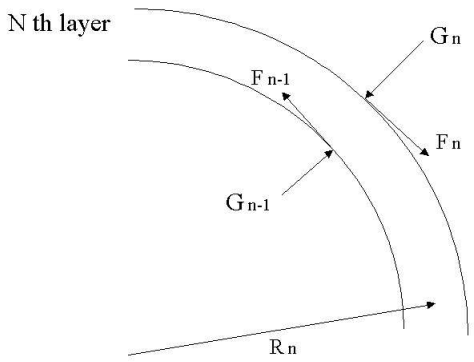


Figure 6.3: Single layer of strands

and

$$\delta G_n = \sum_{k=n+1}^{N} \frac{\sin^2 \alpha_k A_k}{R_k} \left(\frac{d\sigma_k}{d\varepsilon_k} \right) \cos^2 \alpha_k \delta \left(\frac{u}{l} \right).$$

The friction coefficients μ_n are parameters which depend on the nature of the n and $n+1^{st}$ layers.

For all aluminum cables all μ_n are taken to be zero, as the results for such cables seem to be satisfactory. For the other examples μ_n is given one value for all steel cables and another for steel aluminum interfaces. The values are chosen to give a good fit for one set of data and these values are used in other sets of data to examine the credibility of the model.

This model is equivalent to the introduction of resultant body couples of magnitude $\mu_{n-1} G_{n-1}(K_n R_n - K_{n-1} R_{n-1})$ at the n^{th} interface (n^{th} interface between layers $n-1$ and n) since

$$\sum_{n-1}^{N} K_n R_n (\mu_{n-1} G_{n-1} - \mu_n G_n) = \sum_{n=1}^{N} \mu_{n-1} G_{n-1} (K_n R_n - K_{n-1} R_{n-1}).$$

Note $G_N = 0, R_0 = 0$ so that $\sum_{n=1}^{N} K_n R_n \mu_n G_n = \sum_{n=0}^{N-1} K_n R_n \mu_n G_n = \sum_{n=1}^{N} K_{n-1} R_{n-1} \mu_{n-1} G_{n-1}$. The second form of the correction shows clearly that $C' > 0$ as the experiments in [2] suggest.

6.4 Numerical experiments

These were carried out using the data from 5/8 Ground wire,"gwire18" (all steel), Curlew (steel-core, aluminum) and Peace (steel core-aluminum) [3]. For the ground wire the friction μ_G coefficient was chosen to fit the data for the three layer wire and then tested against the two layer case.

The four layer curlew was fitted by using the previous value of μ_G for the steel–steel interface and choosing a new coefficient for the steel–aluminum interface. The aluminum–aluminum interface coefficient was taken to be zero because of the already excellent agreement obtained in [1, 2, 3] for all aluminum cables. These values were then used for a comparison with the three layer Peace results.

The approximate values for the torque in the two examples were

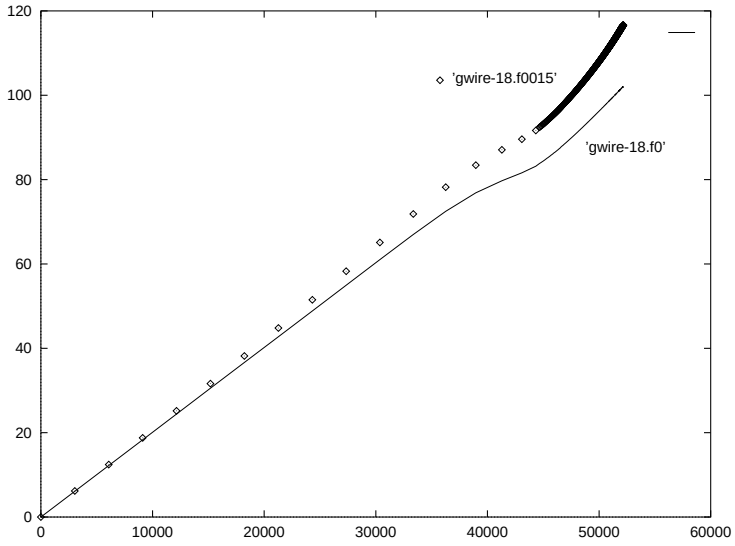


Figure 6.4: Torque vs Force in gwire 18

	Old Theory	New Theory	Exp.
5/8 cable, 2 layer	−76	−63	−53
Peace, 3 layer	−93	16	−70

Both results give corrections in the right direction. The first is not too bad, the second correction is perhaps a little to vigorous. Below we show graphs of Torque versus Force for the 5/8 cable, 2 layer ,"gwire 18" (Figure 6.4) and the "Peace" (Figure 6.5).

6.5 Conclusions

We were unable to devise a convincing new model and consequently resorted to numerical experiments with a feasible model allowing ourselves the luxury of choosing a parameter to fit a chosen set of data. To test whether the method was credible this choice was then used with other data.

The results were unconvincing and although they suggest that the changes were at least not in the wrong direction they are not recommended as a solution to the problem. A more detailed investigation of the changes might yield better results. For example our model did not take into account the number of contact points between two layers. The direction of the forces $\mu_n G_n$ is also rather arbitrarily specified and is perhaps not correct. These are matters which we hope to address in the future.

6.6 References

[1] J. Lanteigne and A. Akhtar, "Evaluation of Tensile Strength of Multistrand Conductors — Part I: Theoretical Basis", *ASME Journal of Engineering Materials and Technology*, Vol. 120, No. 1, 1998, pp. 33–38.

[2] 2 A. Akhtar and J. Lanteigne, "Evaluation of Tensile Strength of Multistrand Conductors — Part II: Experimental Results", *ASME Journal of Engineering Materials and Technology*, Vol. 120, No. 1, 1998, pp. 39–47

[3] 3 A. Akhtar and J. Lanteigne, "Tensile Strength of Stranded conductors in Relation to Properties of Constituent Wires", Report 84 J 637, December 1993.

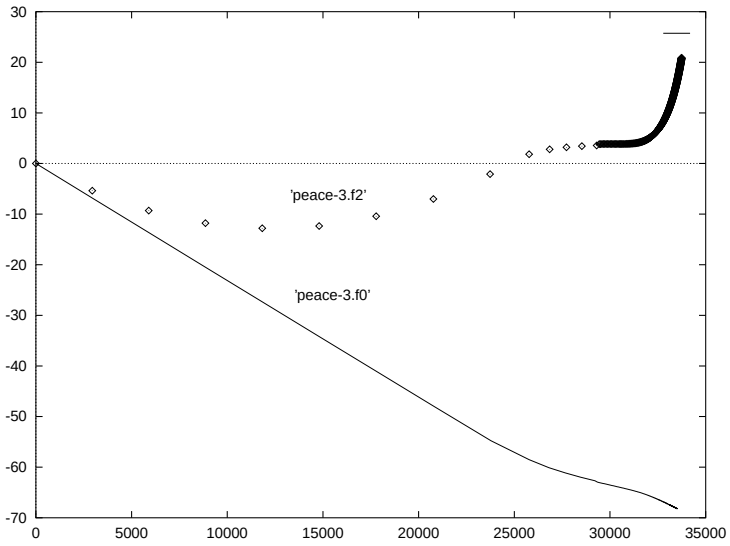


Figure 6.5: Torque vs Force in Peace

Chapter 7

Automatic Detection of Egg Shell Cracks

Rachel Kuske[1], David Lyder[2], Jonathan Samuel[3], Chris Soteros[4], David Wolfe[5]

7.1 Introduction

Eggs are natural objects that possess very complex surfaces, namely their shells, characterized by such features as pinholes, "starry nights", cage marks or thin shells, stretch marks, strains or discolouration, as well as internal flaws such as blood in the yolks, etc. At present the candling of eggs for these features is done by two means:

[1] Manual, which is

- labour intensive, and may cause some mechanical intrusion;
- reliable, if the eggs are of good quality; i.e. detection rates greater than 90% with false positives less than 1%;
- flawed by high false positives if eggs of poor quality.

[2] Mechanical, which is

- effective but intrusive;
- prone to failure after very little usage; this method requires maintenance and may be unreliable if eggs of very poor quality.

This problem has been considered in some detail and at considerable cost ($1.5 million) over the last three years. The method of looking for cracks is based upon the Batelle patent (U.S. patent 4,161,366) with modifications introduced by Weichman *et. al* 1997 (U.S. patent 5,615,777). The principle of this method is to scan the surface of the egg with a laser which will produce a significant amount of diffuse light, known as egg glow, due to multiple internal scattering when the laser beam penetrates the surface of the egg. In practice this method has been used with some success, yielding a crack detection rate of greater than 60% with a false positive rate of less than 3%. However it still does not meet industry standards which require a crack detection rate greater than 80% with a false positive rate less than 1%.

[1]University of Minnesota
[2]Villetard Systems Group Ltd.
[3]University of Victoria
[4]University of Saskatchewan
[5]University of Minnesota

The problem for the study group was to find a reliable, non-intrusive means of detecting cracks in eggs. Intensity data from eggs were collected by Vision Smart for the group to analyse. Given the short time period of the workshop, the group focused on three main questions:

[1] Is there a feature of the intensity data which detects, and discriminates between, pinholes, cage marks, and cracks

[2] Are there ways to improve the current data collection process so that the answer to question (1) would be yes.

[3] Alternatively, are there other data collection methods which should be tried?

An attempt to answer question one is in Section 7.2; we found that there was some promise of a positive answer, but many problems arose. Based on our analysis of the data, we give some answers to question two in Section 7.3. Sections 7.3 and 7.4 give some answers to question three.

7.2 Data Analysis

In order to determine distinguishing features of cracks vs. other markings such as cage marks and starry nights, we analyzed intensity data provided for one dozen eggs. This data was collected by shining a laser on the egg, rotating the egg on end, and measuring the intensity as a function of position during this rotation. This gives one dimensional data at a given distance from the end of the egg, which we refer to as a "slice". The data on each slice contained measurements over 4 revolutions of the egg. This measurement was repeated on 27 different slices of each egg, which covered approximately two-thirds to three-quarters of the egg away from its ends.

An example of this data is shown in Figure 7.2. Since an egg has a different cross-sectional radius depending on the distance from the end, and an egg can wobble as it is being scanned, a black stripe was placed on each egg to aid in locating position on the egg in the data analysis. As can be seen from the data in Figure 7.2, there are four places where the intensity drops to 0, which indicates the location of the black stripe. Also, one can see spikes in the data at 4 places, corresponding to increases in intensity due to either a crack, cage mark, or pin hole. We analyzed this data to see if it allows us to distinguish between these features.

From our candling observations we noted that cracks were, in general, composed of one or more narrow linear segments. Therefore we would expect cracks to show up on several consecutive slices at adjacent positions. We would also expect to see such a feature on each of the four revolutions. Therefore knowing the position in the data is necessary for determining whether there is a crack, and we needed to eliminate deviations due to cross-sectional diameter variation and wobble. We found that a linear transformation, a shift and scaling of position, was sufficient for lining up data from the different slices, using the black stripes. We also found that the deviations due to diameter variation and wobble were about 3% on average (see Figure 7.2), and therefore concluded that these variations were not causing significant errors in the data.

Based on some preliminary observations from candling the eggs and considering the data, we used a simple scheme based on windowing and thresholding to pick out significant features in the data. For example, large spikes in the intensity are generally narrow, and are usually followed by a drop below the mean. (This is the phenomenon of "ringing", due to the circuit response in the measuring device, as seen in Figure 7.2. Indeed, this ringing was a clue to use that the recorded data we were analyzing had been clipped to 256 levels by the software.) Then, if there was a significant variation in intensity, above a certain threshold, in a narrow window of the data, we would mark that location with a 1. Otherwise we would mark it with 0. That is, the data from each slice was converted to a string of 1's and 0's, 1's indicating the location of a significant variation in intensity. If there is a crack in the egg, then one would expect that this marking would yield a series of 1's on consecutive slices at adjacent positions, indicating the increased intensity of light on the crack.

We used this procedure on several of the sample eggs which had significant features. In particular we show the results for five eggs: one with a crack, one with a strong cage mark, one with a weak cage mark, one with a starry night, and one with almost no markings. The windowing/thresholding

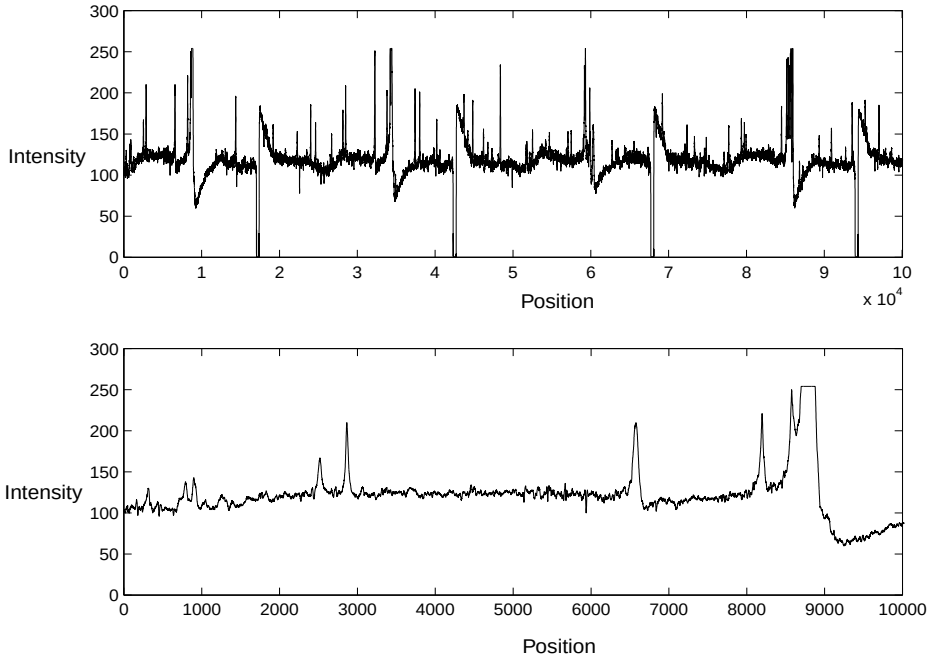


Figure 7.1: Intensity data collected for slice 11 of Egg13. The top figure shows the data collected for 4 revolutions of the egg, on slice 11. The drops in intensity indicate the location of the black stripe placed on the egg. The bottom figure zooms in on this data during the first rotation. Note that the data has been clipped at 256, so that the full height of the intensity spike is not captured. Note also the drop in intensity following the spike, caused by instrument response to rapid changes in intensity.

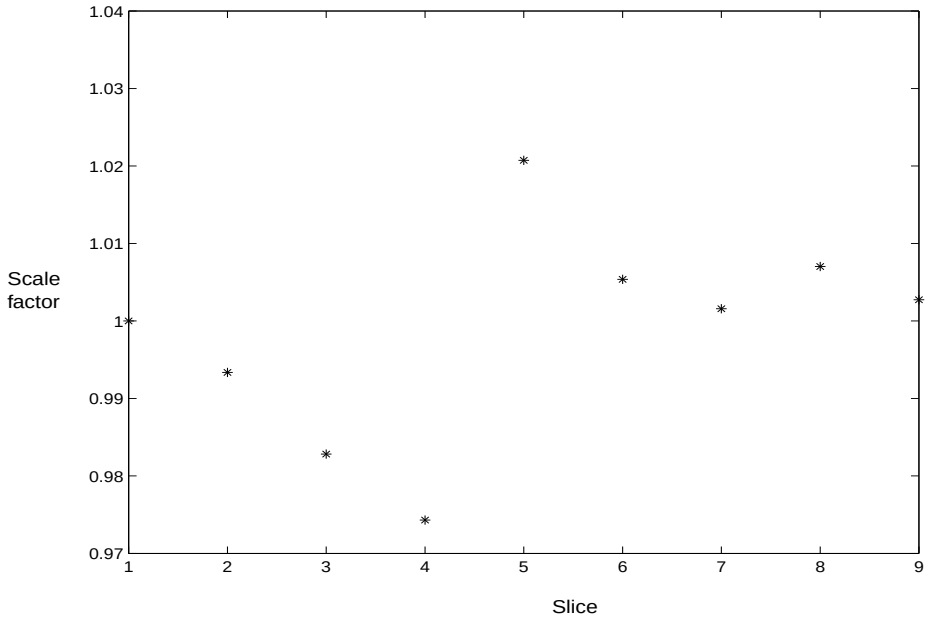


Figure 7.2: Scale factors used to line up data from 9 slices of Egg13. (Slice used as reference value has factor 1, which is the first factor shown.) Note that the scale factors vary on the order of 3%. We found this to be true for the other eggs as well.

scheme applied to the laser data resulted in the following plots, where dark spots indicate the 1's, and the white spaces indicate the 0's (see Figure 7.2). We used approximately the same thresholds and window size for each set of data, with the exception of the last plot (Egg 15), where the threshold was reduced to enhance the picture. The data is repeated 3 times in the horizontal direction, indicating 3 revolutions on the egg. The fourth revolution is not shown, since some parts of this data were not used after the data was lined up using the linear transformation for comparing position.

Summary of the plots:
Egg 13: This egg has a crack, which shows up as a linear sequence of dark spots at the same position on each slice, and it appears on each of the three revolutions. Note that there are also some other isolated spots of intensity variation, due to variations in the readings and the egg shell thickness.
Egg 16: This egg has a strong cage mark, which also shows up on each of the three revolutions as a linear sequence of dark spots at adjacent positions on neighboring slices. Therefore we would have some difficulty in distinguishing this cage mark from a crack, as in Egg 13. We make some comments on distinguishing cage marks, which are generally wider than cracks, in Section 7.3.
Egg 9: This egg has a weak cage mark, which appears on each of the revolutions for not more than 2 or 3 of the slices. Since these markings are isolated from the other variations, which are also isolated from each other, this mark does not have the linear nature of a crack and thus can be distinguished from the crack as in Egg 13.
Egg 6: This egg has a starry night, which is a speckled pattern of spots which show increased light intensity. Since these spots are so close together, it would be difficult to find a crack from this data if there was one. We note again that this data was clipped, so we can not comment on the possibility of finding a crack with unclipped data.
Egg 15: This egg had very few markings, and a significantly lower threshold was used to indicate that there was some variation in intensity. Note that the variations do not compose any linear feature on all three revolutions of the egg, so that we would conclude correctly that there is no crack.

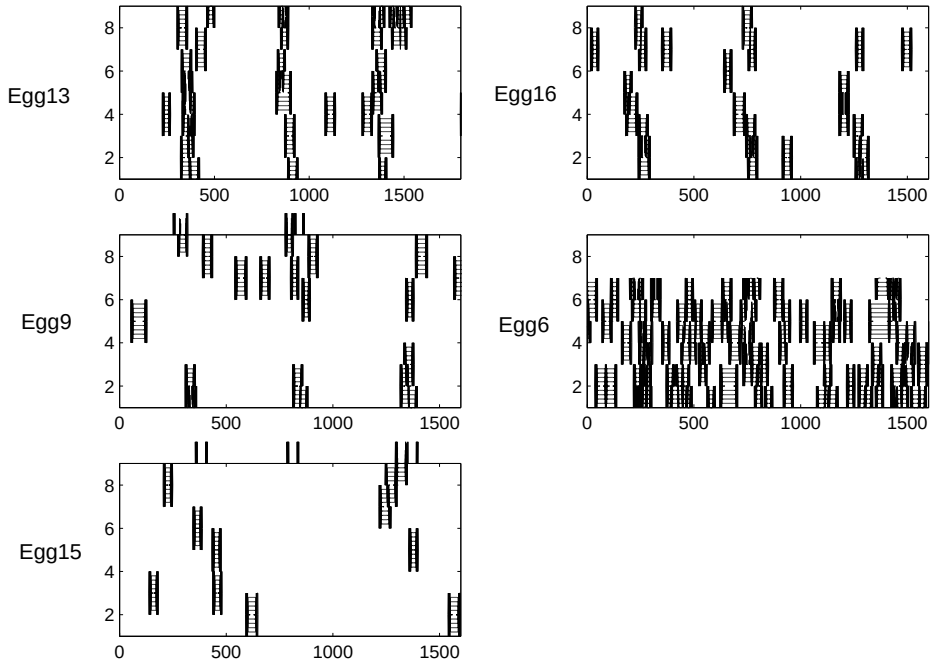


Figure 7.3: Each of these plots is a color coding of the laser data, as a function of position on the circumference (horizontal direction), and slice (vertical direction). The dark areas indicate where there was increased intensity variation.

7.3 Suggested approaches

We discussed a number of alternative investigations to pursue should the present plan prove unfruitful. We've restricted our discussion primarily to approaches which are minor variations on the original proposal.

7.3.1 Beam width

The data collected by VisionSmart used a 100 micron beam width. Since one is measuring intensity of the light penetrating the egg, an ideal situation would be to have the diameter of the laser beam smaller than that of a crack. This way, when the crack rolls past the laser, there would be a point when all the light from the beam would penetrate the shell; one could expect from this a significant increase in the intensity. Moreover, when the crack rolls past a narrower laser beam, the intensity would drop very quickly, much more quickly than when the beam passes through a cage mark, which is wider than a crack. In conversation with Dr. Lyder, it was mentioned that the average width of a crack is about 10 microns. So a beam size of 10 microns (smaller if possible) may help distinguish cracks from cage marks.

7.3.2 Registration

A minority opinion held by some members of our group is that the problem of *registration*, converting the one dimensional data to two dimensional information by aligning the slices of data, is very challenging. As discussed in Section 7.2, the slice diameter varies from the previous by about 3% due to egg wobble and cross-sectional variation in egg diameter. Note also that although we speak of *slices* as though the egg were sliced uniformly, the actual trajectory of the laser along the egg wobbles since the egg wobbles. The reader might think of wrapping a thread around an egg; although one

might hope to keep the spacing even, sometimes the threads will overlap and sometimes there will be wide gaps.

If features are reasonably clear, this problem of determining where the laser beam is on the egg may not be insurmountable. However, features are typically unclear. On an egg such as starry night, the features are quite blurry, making registration virtually impossible.

We propose two possible approaches to deal with the problem of registration:

[1] Avoid the problem of registration altogether by diagnosing cracks in one-dimensional data. With a 10 micron beam, this may work well enough. A *possible* crack found within the same region of several consecutive slices may be sufficient to determine that it is a *probable* crack without precise registration.

[2] Hold the egg steady while moving the laser beam. The position of the beam can be more precisely measured than the rotational position of the egg, making registration much easier.

7.3.3 Hairline cracks

Some hairline cracks are not visible by the untrained naked eye. This suggests that a direct image processing approach may be inadequate for observing these cracks. It may be necessary to apply some pressure to the egg to temporarily widen the cracks. To minimize risk to the equipment and to the egg, we propose that pressure be applied either through bursts of air (avoiding touching the eggs altogether) or with soft spring cushioned walls (the cushioning avoids the need for precise calibration of the motors which move the walls).

If applying pressure to the egg is feasible, then an image-differencing approach (taking an image of the egg before and after pressure is applied) should prove more fruitful than the laser-beam approach.

7.3.4 Vary beam wavelength

If we restrict ourselves to laser beam approaches, there has still been little investigation done concerning which wavelengths are best. We recommend trying several different wavelengths to determine if some wavelengths penetrate better than others.

7.3.5 Imaging techniques

Given the recent improvements in digital imaging and processing power, imaging techniques which were much too slow and cumbersome a decade ago may in fact be fruitful. Digital cameras with high resolution ($1248 \times 1024 \times 12$ bits) are currently able to output images at upwards of 30 frames per second. Actual algorithms for detecting cracks might range from thresholding to fractal or wavelet analysis.

7.4 Literature Review

A literature search, using the on-line databases available at the University of Calgary and the University of Saskatchewan, was performed in order to determine whether any of the questions which arose during our discussions had previously been addressed. Some papers were found which partially address some of the alternate approaches we have suggested; a brief review of these papers is given here.

One alternate approach to the laser experiment is to use an approach which attempts to more closely simulate the backlighting and human eye image analysis of existing egg candling systems. Two papers [1],[2] discuss image analysis of eggs based on approaches of this kind.

In the first paper, the objective of the study was to develop a global analysis procedure in the frequency domain for the detection of cracked egg shells. Direct images of backlit eggs were taken and the images were analyzed using two-dimensional Fourier transform analysis. A global analysis procedure was developed and using various inspection models they obtained an 88% success ratio.

They observed that while 88% is not good enough to replace a commercial egg candling system where manual candling can classify up to 98% of cracked eggs, their global analysis procedure showed better performance than the crack detection algorithm [3] which uses edge detection and contour finding methods. They also suggest that a frequency domain analysis coupled with a neural network may lead to an inspection system which would do better than their current frequency domain analysis approach.

In the second paper, the objective of the study was to develop an automated procedure for detecting egg freshness. An egg was backlit using a cold white light source. The yolk and air-sack of the egg then cast shadows on the opposite wall of the egg and a CCD camera was used to take a photograph of the illumination profile on this wall. The image obtained was pre-processed and then the data was passed to neural networks. They used both multilayer perceptron and modular multilayer perceptron neural network approaches. The work reported here is somewhat preliminary as they only tested 120 eggs, however, they report a 100% classification rate. The features relevant for detecting egg freshness are much less complex than those needed for detecting egg cracks so it is unclear whether the results of this work will be helpful for the egg crack problem.

Cork planks, used for making cork stoppers for example, exhibit a variety of features such as holes, cracks, and insects which have some similarity to the type of features found in egg shells. Several papers [4],[5] indicate that success has been obtained for classifying cork quality using image analysis techniques combined with a neural network approach.

Another alternate approach to the problem is to try to better understand whether the current industry standards are economically justified based on product and market research. Some of the issues relevant to this were discussed by A. Oosterwoud [6] in 1987. With regard to product research, he concludes that the most important factor controlling the proportion of cracked eggs was the type of cage in which the hens were kept. He also discusses consumer attitudes towards eggs. He notes that "It has been standard procedure in marketing technology to establish quality characteristics without asking consumers for their opinions." Although he seems mainly to be talking about internal qualities of the egg, one would suspect that the same is true for external qualities as well. He concludes that more research is needed with regard to consumer attitudes towards eggs.

7.5 References

[1] Y.J. Han and Y. Feng, Egg Shell Inspection Using Global Image Analysis, Appl. Eng. Agr., **10** 109-114 (1994).

[2] C.K Yong, Intact Egg Freshness Quality Inspection Using Neural Networks, Proceedings of the 1995 IEEE International Conference on Neural Networks, **2** 1080-1085 (1995).

[3] A. J. Bourely, Investigation of a robotic egg candling system, Unpublished M.S. thesis, Department of Electrical and Computer Engineering, University of California, Davis (1985).

[4] H. Pereira, F. Lopes and J. Graça, The Evaluation of the Quality of Cork Planks by Image Analysis, Holzforschung, **50** 111-115 (1996).

[5] J. Chang, G. Han, J. Valverde, N.C. Griswold, J.F. Duque-Carrillo and E. Sánchez-Sinencio, Cork Quality Classification System using a Unified Image Processing and Fuzzy-Neural Network Methodology, IEEE Transactions on Neural Networks, **8** 964-974 (1997).

[6] A. Oosterwoud. Effect of egg handling on egg quality, *in* Egg Quality – Current Problems and Recent Advances, Poultry Science Symposium Number Twenty, eds. R.G. Wells and C.B. Belyavin, Ch. 17, pp. 283-291 (1987) (Butterworths; London)

Chapter 8

The Organization of an Industrial Problem Solving Workshop

8.1 Introduction

History

The "Study Groups" with Industry started in Oxford in 1968, with the aim of creating a mutually beneficial link between researchers in industry and academic applied mathematicians. The formula of setting aside a week for intensive study of a few real world problems proved so successful that the Study Group became an annual fixture in many countries around the world. The success of such workshops can be attributed to a number of reasons, for example:

- they help foster contacts between academia and industry, sometimes leading to research contracts,
- they frequently lead to challenging new research areas, which have a direct bearing on physical problems,
- they are an excellent source of research topics for graduate students (as well as academics). Further, they allow companies to become acquainted with students and evaluate them for future employment.

Then there are the obvious reasons: the company wanting a problem solved and the pleasure of working with enthusiastic colleagues.

Set-up of the meetings

The usual format of the meetings is:

- Monday morning: problem presentations by industrial representatives,
- Monday afternoon to Thursday: intensive workshop sessions,
- Friday: summing up by academics.

Some meetings also have a regular summing up of the day's progress.

It is considered very useful to have at least one company representative present during the whole week, to answer any questions that may arise. Although, if this is not possible they should at least be present on Monday and Friday and be easily contactable throughout the week.

8.2 How to contact industry

Personal contacts are clearly the best way to get in touch with a company. These may come about from previous workshops (many companies become regulars), or in the case of PIMS, from the industrial working seminars. Many academics have their own contacts, particularly from graduate students that have moved into industry. As a groups experience with industrial modeling increases so does its reputation and number of contacts.

To foster new contacts a letter, fax or phone call to a companies research department may work. This should outline:

- how the meeting is set up,
- the benefits for the company and how much (or little) it will cost them,
- the previous item may make them suspicious and question the worth of the meeting, so explain the benefits to academics as highlighted in the history section,
- in a fax or letter you should include a list of problems tackled at previous meetings, to indicate what can be done (feel free to bias this list to the companies interests).

A selection of problems investigated at recent PIMS workshops is given in the Appendix.

8.3 Screening the problems

Once problems start to flood, or trickle, in, they require filtering. Industrial researchers may not be familiar with recent academic work and so the problems they submit may be already solved, trivial, reasonable or insoluble (however, even insoluble problems may be of interest, *i.e.* a different approach may be suggested). Presuming the coordinator is not an expert on every subject, they should try to enlist expert help to determine which category a problem falls into. Familiarity with, or knowledge of, experts in different fields is essential for a coordinator. Screening may also be carried out by searching the literature, using a science citation index.

Even if a problem appears tractable to the coordinator, it must be suitable for the participants. However, experts can usually be enlisted (lured) if they know a problem of specific interest will be presented.

The number of problems chosen should be based on the the number of participants. Typically ten people per problem is a realistic number. Bear in mind that only a small proportion of participants will be knowledgeable in any given field. Many people attend to learn new topics or through curiosity.

8.4 Report/proceedings

By Thursday it should have been decided who is to present the results the following day. This person will also be in charge of the writing up, whether they do it all themselves or coordinate efforts by group members is up to them, provided the report gets completed. Two months is the standard time for this task; the more time drags on, the less committed people are to dealing with it. The companies should be allowed to screen the report before it is finalized.

A member of the organization committee will be in charge of seeing that the deadlines are obeyed and also to put the final report together.

8.5 Organization

Most requirements for organizing a Problem-Solving Workshop are similar to those of a normal academic meeting. The following highlights a number of differences that should be kept in mind.

Location

Such a workshop must be held in the vicinity of a university, as both library and computer access are essential. For the problem presentation on the first day and the summing-up on the last day, a big room is required, with overhead projector facilities and a board. For the workshop's discussions, several small rooms are preferred with boards. All rooms should be close together, as some participants will want to flit between the different problem sessions.

Accommodation

Academics should be housed close together, allowing the discussions to keep going into the evening (and frequently in the bar). Note, industrial people will probably prefer hotel accommodation to university dorms; these should be close to the university so the representatives are not 'left out'.

Fees

Most meetings require companies to pay a fee, to help 'defray' costs for the academics (the fee from approximately 8 companies is unlikely to cover travel and accommodation costs for everybody). Typical charges range from $1600 to $4000. Government and other types of research grants can also be useful; emphasizing the benefits to be gained from interaction between universities and industry, such as increased competitiveness, can help in obtaining funds. Academics, and their universities, may also be willing to pay their own costs.

Facilities

The library should be accessible for all participants to consult journals and make xerox copies. If applicable, arrange for copy cards or access codes and ask library members attending the meeting to help with the borrowing of books.

Computer access is a must for e-mail junkies, but also to provide facilities such as telnet, maple, mathematica, matlab, fortran etc., plotting & print-outs.

Food and drink

Accommodation and meals should be provided to all participants. Coffee and tea breaks are invaluable for regular summing up and to get honest opinions and fresh ideas from people who are working on other problems. It's nice to point out some restaurants, pubs and bars in the vicinity of the workshop rooms, so that the discussions can continue in the evening.

8.6 References

Useful world-wide contacts on the study group meetings:
Canada: Dr. Nassif Ghoussoub *Director*, Pacific Institute for the Mathematical Sciences.
UK: Dr John Ockendon *OCIAM Research Director* University of Oxford.
UK: Dr Tim Myers *Industrial Post-doc* Cranfield University.
Australia: Dr Kerry Landman *MISG Director* Melbourne University.
US: Dr Ellis Cumberbatch *Professor* Claremont Graduate School.
US: Dr Avner Friedman *Director* IMA Minnesota, Minneapolis.
US: Donald Schwendeman *Associate Professor* Rensselaer Polytechnic Institute (RPI).

Web sites:
http://www.pims.math.ca
http://macserver.maths.mu.oz.au/misg/
http://www.maths.bath.ac.uk/ESGI97
http://www.maths.ox.ac.uk/ociam (see newsletter in particular)
http://www.math.rpi.edu

http://www.indmath.uni-linz.ac.at
http://www.mat.dtu.dk/ECMI
http://www.siam.org

8.7 A selection of problems investigated at recent meetings at PIMS

- Inversion for Anisotropic-Velocity Parameter (Petro-Canada)
- Fingerprint Identification (Kinetic Sciences Inc)
- Modeling Bronchial Epithelial Lesions (BC Cancer Labs)
- Optimally Cutting Logs (MacMillan Blodel)
- Modeling Stress Intensity in a Thermoroll (MacMillan Blodel)
- Measuring the Stress Intensity of a Composite Vessel (Powertech Labs)
- Evaluating Computational Methods in Fuel Cell Systems (Ballard Powersystems)
- Rapid Thermal Processing of Semiconductors (Vortek Industries Ltd)
- Optimal Trading Strategies for Electricity Trading (Powerex Ltd)

Chapter 9

Participant List

Company Representatives

Ainul Akhtar	Powertech	ainul.akhtar@bchydro.bc.ca
Dan Kenway	VisionSmart Inc.	vsmart@planet.eon.com
David Lyder	Villetard Systems Group Limited	vsmart@planet.eon.com
Avygdor Moise	Itres Research Limited	avy@itres.com
Greg Robel	Boeing Corporation	gregory.f.robel@boeing.com
Peter Sammon	Computer Modelling Group Limited	peter@cmgroup.com
Michael Slawinski	The Geomech Project	mslawins@enme.ucalgary.ca
Bill Vetter	Geospace Research, Inc.	vetterb@cadvision.com

Foreign Participants

Chris Budd	University of Bath	cjb@maths.bath.ac.uk
Alistair Fitt	University of Southamption	adf@maths.soton.ac.ak
Rachel Kuske	University of Minnesota	rachel@math.umn.edu
David Wolfe	Gustavus Adolphus College	wolfe@gustavus.edu

Carleton University

Markus Orasch	morasch@math.carleton.ca

McGill University

Shane Jensen	jensen@math.mcgill.ca

Queen's University

Andrew Irwin	irwin@mast.QueensU.ca

Simon Fraser University

Janez Ales	janez@sfu.ca
Bradley Bart	bbart@cs.sfu.ca
Daniel Chertok	chertok@cs.sfu.ca
Roger Coroas	coroas@cs.sfu.ca
Daya Gaur	gaur@sfu.ca
Arvind Gupta	arvind@cs.sfu.ca
Shabnam Karousian	skavousi@math.sfu.ca
Bruce Rout	routfam@island.net
John Stockie	jms@math.ucalgary.ca

University of Alberta

Andreea Amariei	amariei@math.ualberta.ca
Richard Archibald	rarchiba@math.ualberta.ca
Joseph Modayil	jmodayil@math.ualberta.ca
Edo Nyland	edo@phys.ualberta.ca
Jack Macki	jmacki@math.ualberta.ca
Sherman Riemenschneider	sriemens@math.ualberta.ca
Akbar Rhemtulla	akbar@malinda.math.ualberta.ca
J.G. Timouria	jtimouri@vega.math.ualberta.ca
Valdislav Agapov	vagapov@math.ualberta.ca

University of British Columbia

John Anderies	anderies@math.ubc.ca
Adriana Dawes	atdawes@math.ubc.ca
Kelly Kwok	kelly@stat.ubc.ca
Jian Liu	liu@stat.ubc.ca
Joshua Madden	jmadden@iam.ubc.ca
Rafael Sanegre	sanegre@aecb.ubc.ca
Bryan Seymore	seymour@math.ubc.ca
Michele Titcombe	titcombe@math.ucalgary.ca
Ryan Wai Tse	ryan@stat.ubc.ca
Huaxiong Huang	hhuang@math.ubc.ca
Vlad Shapiro	vshapiro@triumf.ubc.ca

University of Calgary

Ben Aggarawla	aggarwal@math.ucalgary.ca
Rita Aggarwala	rita@math.ucalgary.ca
R. Ait-Haddou	haddou@math.ucalgary.ca
Larry Bates	bates@math.ucalgary.ca
Bruce Bauslaugh	bauslaug@math.ucalgary.ca
Dan Calistrate	calistra@math.ucalgary.ca
Charles Cuell	cuell@math.ucalgary.ca
Ernest Enns	enns@math.ucalgary.ca
Jeff Grossman	grossman@math.ucalgary.ca

Chris Jessop	cjessop@acs.ucalgary.ca
David Kemppainen	dwkemppa@math.ucalgary.ca
Gerry Labute	labute@math.ucalgary.ca
Claude Laflamme	laf@math.ucalgary.ca
Michael Lamoureux	mikel@math.ucalgary.ca
Nancy Morrison	morrison@math.ucalgary.ca
Marc Paulhus	paulhusm@math.ucalgary.ca
Rob Petry	rpetry@math.ucalgary.ca
Arunas Salkauskas	arunas@math.ucalgary.ca
Kes Salkauskas	ksalkaus@math.ucalgary.ca
David Scollnik	scollnik@math.ucalgary.ca
Karen Seyffarth	kseyffar@math.ucalgary.ca
Gordon Sick	sick@acs.ucalgary.ca
Mohammadreza Simchi	simchi@math.ucalgary.ca
Jedrez Sniatycki	sniat@math.ucalgary.ca
Satoshi Tomoda	tomoda@math.ucalgary.ca
Tony Ware	aware@acs.ucalgary.ca
Paul Webster	pswebste@acs.ucalgary.ca
Rex Westbrook	westbroo@math.ucalgary.ca

Université du Montréal

Miro Powojowski	powojo@dms.umontreal.ca

Université du Québec, Sherbrooke

Khalid El Yassini	Khalid.El.yassini@DMI.USherb.ca

University of Saskatchewan

Alan M. Duffy	amd129@snoopy.usask.ca
Michael Elia	eliam@plasma.usask.ca
Eugene Fourkal	efurkal@plasma.usask.ca
Chris Soteros	soteros@math.usask.ca

University of Toronto

David Saunders	saunders@math.ucalgary.ca
Andrij Sheshnyov	avs@math.toronto.edu

University of Victoria

Sean Bohun	bohun@math.uvic.ca
Ian Putnam	putnam@math.uvic.ca
Jonathan Samuel	samuel@math.uvic.ca
Cristina Stoica	cstoica@math.uvic.ca

University of Western Ontario

Antonio Cabal cabal@apmaths.uwo.ca