Proceedings of the Eighth

PIMS-MITACS Industrial Problem Solving Workshop

PIMS-MITACS IPSW 8

Co-sponsored by:

The Natural Science and Engineering Research Council of Canada

Editor: C. Sean Bohun, Pennsylvania State University

Proceedings of the Eighth Annual PIMS-MITACS Industrial Problem Solving Workshop

Editor: C. Sean Bohun, Pennsylvania State University

Co-sponsored by:

The Natural Sciences and Engineering Research Council of Canada

July, 2004

Foreword by the PIMS Director

The PIMS Industrial Problem Solving Workshops have been held annually since 1997. The 2004 edition was co-sponsored by MITACS (Mathematics of Information Technology and Complex Systems), a Network of Centres of Excellence, which we thank for its support.

Approximately 68 participants worked intensely on four problems posed by industrial companies from across North America. Each of them shows the need for correct mathematical modelling, and the industrial payback from a good solution. The problems were provided by Bill Mawby (Michelin Tire Corporation), Surrey Kim (Random Knowledge), John Ceko (Husky Injection Molding Systems), and Kenneth J. Hedlin (Husky Energy) with Gary Margrave (University of Calgary).

I wish to thank Sean Bohun from Penn State University who carefully edited these proceedings, and Rex Westbrook and Kes Salkauskas from the University of Calgary for the difficult task of providing the industrial contacts and organizing the meeting.

Dr. Ivar Ekeland, Director
Pacific Institute for the Mathematical Sciences

Contents

List of Figures

List of Tables

Preface

The University of British Columbia was the site of the eighth annual PIMS-MITACS Industrial Problem Solving Workshop (IPSW). Hosted from May 17 through May 21, 2004, the workshop was cosponsored by MITACS (Mathematics of Information Technology and Complex Systems and saw the bringing together of participants from around the world.

Many first time participants of these workshops are surprised by the intensity of the work. Based on the Oxford Study Group Model, the morning of the first day consists of the initial problem presentations and the beginning of focused discussions. Work continues throughout the next few days culminating in a summary presentation on the morning of the fifth day. These presentations form the basis of the proceedings. I would like to personally thank the authors for their dedication and prompt response so that these proceedings could appear in a timely fashion. These individuals were:

- Michael Lamoureux: The Michelin Tire Problem;
- Randall Pyke: The Random Knowledge Problem;
- Greg Lewis: The Husky Injection Molding Systems Problem;
- Nilima Nigam: The Husky Energy Problem.

To ensure the smooth and efficient operation of the workshop, many individuals are needed behind the scenes. I would like to thank the organizing committee consisting solely of Rex Westbrook and Kes Salkauskas both from the University of Calgary. Without the considerable efforts of these two individuals, this event would not have been possible.

Two other essential ingredients for the success of the workshop were the industrial representatives and the industrial experts. While the representatives are certainly an asset and provide a grounding for each of the problems, it is the academic experts that are responsible for each of the groups moving along productive lines. For this year the academic experts were:

- Petra Berenbrink, Simon Fraser University;
- C. Sean Bohun, Penn State University;
- Ian Friggard, University of British Columbia;
- Huaxiong Huang, York University;
- Michael Lamoureux, University of Calgary;
- Tim Myers, University of Cape Town;
- Nilima Nigam, McGill University;
- Anthony Peirce, University of British Columbia;
- Randall Pyke, University College of the Fraser Valley;
- Tobias Schaefer, University of North Carolina at Chapel Hill;

- John Stockie, Simon Fraser University;
- Brian Wetton, University of British Columbia.

A special thanks goes to the industrial contributors and their representatives who include:

- Bill Mawby, Michelin Tire Corporation;
- Surrey Kim, Random Knowledge;
- John Ceko, Husky Injection Molding Systems;
- Kenneth J. Hedlin, Husky Energy.

At the end of this monograph there is a listing of the workshop participants and the various institutions where they can be found and I would like to take this opportunity to apologize for any mistakes or omissions therein.

In closing I would like to thank Heather Jenkins at the PIMS central office for her dedication and efficiency in dealing with the production of these proceedings.

C. Sean Bohun, Editor,
Department of Mathematics,
Penn State University.

Pimslips

When people are working intensely, they sometimes say things that, in retrospect, are amusing or downright hilarious. Here are some from IPSW 8:

- (Preamble: ”Love” is the author of a standard textbook in this field)
  “I guess wherever Love is, that’s a good place to look.”
  “You don’t go to where “Love” is.”
  “No, Rex did. I went to where the money is.”
- A recommendation to industry on coolants:
  “Invest money in better pumps and lager pipes instead”
- “Saying it out loud makes me not sure if I agree with what I just said.”
- “21 is a prime, and it’s a great prime, but let’s pretend it isn’t.”
- “The blue line is clearly different from the red line.”
- (While arguing about the singularity of a matrix whose determinant is about 10^{-2243}):
  “Still, 10^{-2243} is bigger than zero.”
- “In Vancouver you can get the skiing without the snow!”
- “_____ This line is of course vertical.”
- “51 is a prime—no, wait, I forgot my 17 times table!”
- “The more I leak, the less bouncing I expect.”
- “Catherine is more jittery here that Gilles. Gilles looks smooth.”
  (with reference to denoising of seismic data).

Chapter 1

General Statistical Design of an Experimental Problem for Harmonics

Problem Presented By: Bill Mawby (Michelin Tire Corporation)

Mentors: Sean Bohun (Penn State Mont Alto); Peter Gibson (York University); Michael Lamoureux (University of Calgary)

Student Participants: Benjamin Akers (University of Wisconsin - Madison); Andreas Hofinger (Johannes Kepler University, Linz); Jason Lobb (Carleton University); Malcolm Roberts (University of Alberta)

Report prepared by: Michael Lamoureux (mikel@math.ucalgary.ca)

1.1 Introduction

Each year, the Pacific Institute for the Mathematical Sciences organizes a one-week long Industrial Problem Solving Workshop (IPSW), where researchers from academia and industry work for five intense days to make progress on solving specific mathematical problems that arise in real industrial environments. Typically, many new, interesting mathematical results are obtained that provide concrete value to the companies proposing the problems.

Four years ago, the Michelin Tire Corporation proposed a problem on experimental design, to improve the manufacturing process for their tires. The idea is basically to determine the effects of placements for various layers built up in the construction of a tire, to allow the design of a smooth tire with a smooth ride. A highly success solution was developed, and it has been reported that this method introduced savings of over half a million dollars in their test processes. This year, Michelin returned to the workshop with an extension to the original problem, to address specific refinements in the testing method. This report summarizes the work completed in course of the five day workshop.

1.2 Problem Description

Tires are subjected to a variety of force measurements that are stored as periodic waveforms. Harmonic components of these waveforms are related to tire performance characteristics such as noise and comfort and hence the control and reduction of the amplitudes of these harmonics is an important activity of manufacturing. Technicians may choose to perform designed experiments on their production processes to understand better their impact on the resulting force harmonics. It could be advantageous to have a general design of experiment methodology which allows technicians to choose optimal designs for their studies.

To make this more concrete consider two types of forces (F and G). F is characterized by 5 harmonics F1-F5 and G is characterized by 10 harmonics (G1-G10). Practically, the technician might have 20 different process elements (P1-P20) that can be rotated within the construction of the tire and which can affect the force measurements. It is assumed that rotation of a production process will result in the equivalent rotation of the force measurement and that superposition of P1 through P20 will result in a corresponding superposition of resulting Fi's. In general the movement of any process element such as P1 may affect all harmonics and forces (F1-F5 and G1-G10). The general problem is to choose the angles of rotation for a set of Ps so that the harmonic effects are well estimated and the cost of experimentation is minimized. Note that the variance of the estimates is related to the angles chosen (for example choosing 180 degrees prevents the estimation of the even harmonics) and that the cost of a study is proportional to the number of angles that are used in the design. Other features of interest include the reparability/extensibility of designs, identifying sets of competitors/surrogates and allowing different precision for different harmonics.

The original problem in 2000 was proposed as follows: Develop a method that allows estimation of n harmonics on m production steps from sampled waveforms on tires. This method should be flexible, robust and easily constrained to meet operating conditions. For example we measure a waveform of force variation on a cured tire sampled at 256 equally-spaced points made with 5 products with joints at fixed angular positions. Then we change these relative angular positions, construct a new tire and measure its waveform. We would like to decompose the overall effect represented in each tire into contributions due to each product. We can do this with each harmonic of the waveform but would like to ensure good estimation of all effects for all of a specified set of harmonics.

The new problem for 2004 would be to extend the previous PIMS IPSW 2000 results on Statistical Design in any of several directions that are discussed below.

1. Develop fully the method called the Good Lattice Points (GLP) method presented in the proceedings [1] so that it could be implemented in practice to allow estimation without the prime number restriction.

2. Include the fitting of a few non-harmonic frequencies with the harmonics and find good designs for these (assumes that all harmonics are not fitted) such as a frequency that passes through the signal exactly 0.62 times. The harmonics are relative to the tire circumference in the old problem, but sometimes effects such as extrusion or measuring devices put sinusoidal patterns into the overall waveform but these effects are not harmonics (the periods are not integral divisors of the tire circumference) of the tire but rather have periods that are fractional parts (like 0.62). We use multiple linear regression to estimate non-harmonic frequency effects and harmonic frequency effects simultaneously (with some correlation between the estimates). The problem is to provide

an optimal design strategy for this situation given that we can provide some information like number of non-harmonic frequency effects and possible ranges for their frequencies.

3. Determine the best design for a function of the harmonics of different types such as $f(\text{type 1 harmonic } 1, \text{type 2 harmonic } 1)$ where the experiment is performed on rotation of tire components as before and f is of a specified class or form. In this case we measure two or more type of waveforms on each tire as before. We then combine these multiple outputs into single derived output (often linearly by summation etc. but it could be non-linear). We want to ensure adequate estimation of the product effects for each of a selected set of harmonics for this derived output.

4. Expand the concept to a two dimensional Fourier transform or equivalent where the surface could be considered flat or as the surface of an inflated tire (semi-toroidal).

1.3 Methodology

It was clear early in the workshop that this problem could be handled quickly by reviewing the analysis which was done in 2000, and extending those ideas to the new problems at hand. We reviewed the required Fourier techniques to describe the harmonic problem, and statistical techniques to deal with the linear model that described how to accurately measure quantities that come from real experimental measurements. The "prime method" and "good lattice points method" were reviewed and re-analysed so we could understand (and prove) why they work so well. We then looked at extending these methods and successfully found solutions to problem 1) and 2) above. Matlab code was written to test and verify the algorithms developed. We have some ideas on problems 3) and 4), which are described below.

1.4 Fourier Series Formulation - a review

The tire's force characteristic curve can be described as a continuous function on the interval $[0, 1]$ with periodic boundary conditions. As the problem was originally posed, it was suggested that these periodic functions $f \in C[0, 1]$ be expanded in terms of sines and cosine functions: that is, one writes

$$f(t) = \sum_{n=0}^{\infty} \left[a_n \cos(2\pi nt) + b_n \sin(2\pi nt) \right],$$

where a_n, b_n are real coefficients encoding the magnitude and phase of the corresponding harmonic. Algebraically, it is more convenient to use complex exponentials to expand the periodic function in a Fourier series, as

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi int}, \qquad c_n = \int_0^1 f(t) e^{-2\pi int}\, dt.$$

Since f is real, $\bar{f} = f$ with the bar denoting the complex conjugate. As a result, the c_n must satisfy the condition

$$\bar{c}_n = \int_0^1 \bar{f}(t) e^{2\pi int}\, dt = \int_0^1 f(t) e^{2\pi int}\, dt = c_{-n}.$$

More important is the redundancy due to the fact that the continuous function $f \in C[0,1]$ is observed only at finitely many equally spaced points $t_k = k/2N$, for $k = 0, 1, \ldots, 2N-1$ and we can represent f as

$$f(t) = \sum_{k=0}^{2N-1} f(t_k)\delta(t - t_k).$$

Using this representation the Fourier coefficients take the form

$$c_n = \sum_{k=0}^{2N-1} f(t_k)e^{-2\pi i n t_k}$$

and as a result

$$c_{n+2N} = \sum_{k=0}^{2N-1} f(t_k)e^{-2\pi i(n+2N)k/2N} = \sum_{k=0}^{2N-1} f(t_k)e^{-2\pi i n t_k}e^{-2\pi i k} = c_n.$$

We see that observing f at only finitely many points introduces a redundancy in the Fourier coefficients effectively collapsing the Fourier series expansion to the finite sum

$$f(t_k) = \sum_{n=-N+1}^{N} c_n e^{2\pi i n t_k}.$$

It is worth noting in passing that the Fourier coefficients c_n are quickly calculated using an FFT software routine, and so this formulation of the problem does not introduce any additional complexity in the problem.

In the problem at hand, a function f is considered the signature of a given tire component the will affect the final force profile. If this component is rotated by an angle θ, the signature function f is shifted and the corresponding Fourier coefficients change. Introducting the notation S_θ for the shift operator, one obtains

$$(S_\theta f)(t) = f(t-\theta) = \sum_n c_n e^{2\pi i n(t-\theta)} = \sum_n (e^{-2\pi i n\theta}c_n)e^{2\pi i n t} = \sum_n c_n^\theta e^{2\pi i n t}.$$

That is, the Fourier coefficients transform under the shift by θ as a linear transform $c_n \mapsto e^{-2\pi i n\theta}c_n$. Equivalently, the vector of coefficients c_n transforms as

$$\begin{pmatrix} \vdots \\ c_n \\ \vdots \end{pmatrix} \mapsto \begin{pmatrix} \vdots \\ c_n^\theta \\ \vdots \end{pmatrix} = \begin{pmatrix} \ddots & & \\ & e^{-2\pi i n\theta} & \\ & & \ddots \end{pmatrix} \begin{pmatrix} \vdots \\ c_n \\ \vdots \end{pmatrix}$$

which is just multiplication by a diagonal matrix D_θ whose entries are complex exponentials.

1.5 The Component Problem - a review

A tire is built up from a number of layers (tread, cords, airtight inner rubber, etc.), usually on the order of $m = 20$ components. Each layer contributes a signature $f^k(t) \in C[0,1]$ to the observed force profile

$$F(t) = \sum_{k=1}^{m} f^k(t),$$

where the assumption (based on Michelin's experience) is that the contribution is additive. The component functions f^k cannot be measured directly; however, a factory technician may modify the construction of the tire by changing the positioning of individual layers within the test tire. Each layer may be shifted independently by some angle θ. Applying a vector of shifts $\Theta = (\theta_1, \theta_2, \ldots, \theta_m)$, where θ_k is the rotation angle for kth layer, gives an operation on the observed force profile as

$$(S_\Theta F)(t) = \sum_{k=1}^{m} (S_{\theta_k} f^k)(t).$$

In Fourier components, this becomes

$$\sum_n C_n^\Theta e^{2\pi i n t} = \sum_k \sum_n \left(e^{-2\pi i n \theta_k} c_n^k\right) e^{2\pi i n t} = \sum_n \left(\sum_k e^{-2\pi i n \theta_k} c_n^k \right) e^{2\pi i n t},$$

and by equating terms in the Fourier expansion one obtains the transform directly on the coefficients as

$$C_n^\Theta = \sum_{k=1}^{m} e^{-2\pi i n \theta_k} c_n^k.$$

In particular, one observes there is no mixing of harmonics: that is, the nth harmonic of the observed (transformed) force curve is a weighted sum of the nth harmonics of the contributing layers.

The analysis problem is to determine the coefficients c_n^k from the observed C_n^Θ, using some choice of the vector of angles Θ. Since the observed spectra are real, it is enough to consider only non-negative n in determining the harmonics, and the constant term ($n = 0$) is irrelevant. In practice, only a small number of low harmonics is of interest (eg. $n = 1, 2, \ldots, 5$), but the coefficients must be determined for all layers (eg. $k = 1, 2, \ldots, 20$). Also note this is becomes a statistical problem, as the measured coefficients include measurement error and statistical deviations due to variations in the construction of these real tires.

1.6 The Linear Model - a review

The problem is to determine individual coefficients c_n^k, for all layers $k = 1, \ldots, m$, from observations of the lumped coefficients C_n^Θ, where the experimental design involves choosing some appropriate vectors of angles $\Theta = (\theta_1, \theta_2, \ldots, \theta_m)$. The design also should find how many vectors $(\Theta^1, \Theta^2, \ldots, \Theta^R)$ are required to accurately determine the coefficients c_n^k. These coefficients must be determined for a range of harmonics, say $n = 1, \ldots, q$, and it will be convenient to design the experiments to work for all these harmonics simultaneously.

It is natural to group the coefficients into vectors, as

$$\boldsymbol{c}^k = \begin{pmatrix} c_1^k \\ \vdots \\ c_q^k \end{pmatrix}, \qquad \boldsymbol{C}^\Theta = \begin{pmatrix} C_1^\Theta \\ \vdots \\ C_q^\Theta \end{pmatrix}.$$

The linear model encompassing all layers, and the range of harmonics, can thus be written in block form as

$$\begin{pmatrix} \boldsymbol{C}^{\Theta^1} \\ \boldsymbol{C}^{\Theta^2} \\ \vdots \\ \boldsymbol{C}^{\Theta^R} \end{pmatrix} = \begin{pmatrix} D_{\theta_1^1} & D_{\theta_2^1} & \dots & D_{\theta_m^1} \\ D_{\theta_1^2} & D_{\theta_2^2} & \dots & D_{\theta_m^2} \\ \vdots & & & \vdots \\ D_{\theta_1^R} & D_{\theta_2^R} & \dots & D_{\theta_m^R} \end{pmatrix} \begin{pmatrix} \boldsymbol{c}^1 \\ \boldsymbol{c}^2 \\ \vdots \\ \boldsymbol{c}^m \end{pmatrix} + \boldsymbol{\epsilon}$$

where $\Theta^1, \dots, \Theta^R$ is the choice of vectors of angles set in the experiment, the $D_{\theta_k^j}$ are $q \times q$ diagonal matrices with entries $\exp(-2\pi i n \theta_k^j)$ on the diagonal, and $\boldsymbol{\epsilon}$ is the statistical measurement error.

More succinctly, the linear model is represented by $\boldsymbol{C}^{\boldsymbol{\Theta}} = D\boldsymbol{c} + \boldsymbol{\epsilon}$, with D a matrix in block form, each block a diagonal matrix as above. If these were real matrices, the solution via least squares is clear; it is a simple exercise to verify that even for complex matrices the least square solution (using complex inner products) is straightforward. Namely, one solves for $\boldsymbol{c}$ as

$$\hat{c} = (D^\dagger D)^{-1} D^\dagger \boldsymbol{C}^{\boldsymbol{\theta}},$$

where $D^\dagger$ indicates the complex conjugate transpose of the matrix D. Similarly, the variance estimates for the inversion will depend on the properties of matrix $(D^\dagger D)^{-1}$.

Noting that $D^\dagger D$ is also in block form, it is convenient to permute rows and columns (essentially grouping terms by layers, rather than harmonics) to obtain the matrix $X = \mathrm{Perm}(D)$ so that $X^\dagger X$ is in block diagonal form, with

$$X^\dagger X = \begin{pmatrix} (Z_1) & & & 0 \\ & (Z_2) & & \\ & & \ddots & \\ 0 & & & (Z_q) \end{pmatrix},$$

where each $m \times m$ block (Z_n) has entries

$$(Z_n)_{jk} = \sum_{r=1}^{R} e^{2\pi i n(\theta_k^r - \theta_j^r)}.$$

This greatly simplifies the analysis, since each block (Z_n) may be examined separately. Notice each such block corresponds to a separate harmonic.

The problem becomes that of estimating the regression coefficients in the multiple regression model

$$\boldsymbol{C}^{\boldsymbol{\Theta}} = X\boldsymbol{c} + \boldsymbol{\epsilon}$$

where for simplicity we have used $\boldsymbol{C}^{\boldsymbol{\Theta}}$, $\boldsymbol{c}$ and $\boldsymbol{\epsilon}$ to denote the reordered versions of these vectors. A standard assumption is that

$$\mathrm{Var}(\boldsymbol{\epsilon}) = \sigma^2 I_{qm}$$

for some (unknown) σ. It was pointed out that this assumption essentially says the R tires and q harmonics act independently, and different measurements have equal error; this may be a gross over-simplifcation worth further investigation. For instance, there may be some bias in the way the tires are constructed for the test, or trends reflected in the sequence that the tires are built. On the other hand, the harmonics are orthogonal measures in a large dimensional space, and at least some of us were convinced that the first few harmonics would act independently, with similar measurement error. In any case, we proceed with this assumption.

The least-square estimator of $\boldsymbol{c}$ is

$$\hat{c} = (X^\dagger X)^{-1} X^\dagger \boldsymbol{C}^{\boldsymbol{\Theta}}$$

with variance

$$\mathrm{Var}(\hat{c}) = (X^\dagger X)^{-1} X^\dagger \,\mathrm{Var}(\boldsymbol{\epsilon}) X (X^\dagger X)^{-1} = \sigma^2 (X^\dagger X)^{-1}.$$

Thus the equation of finding an optimal design boils down to finding a matrix X such that $X^\dagger X$ is "good".

Some possible optimality conditions ("goodness" of X) include minimizing the determinate of matrix $(X^\dagger X)^{-1}$ (D-optimality), minimizing the spectral norm of $(X^\dagger X)^{-1}$, or minimizing the maximum eigenvalue of $(X^\dagger X)^{-1}$. It turns out these three conditions are equivalent, since the matrix $X^\dagger X$ has a trace independent of the choice of angles (equal to mqR) and thus the minimum occurs when all eigenvalues are equal, and $X^\dagger X$ is R times the identity matrix. Generally speaking, the closer $X^\dagger X$ is to diagonal, the better.

Another optimality condition is to fix some vector $\boldsymbol{w}$ and minimize the variance $\mathrm{Var}(\boldsymbol{w}^{\mathrm{T}}\hat{c})$, which is a weighted sum of the entries of $\hat{c}$. This would be of interest to the manufacturer when some harmonics, or some layers, are deemed to be more important than others.

1.7 The Prime, GLP, and SLP Methods - a review

A key result of our earlier project in 2000 was the development of three methods to pick a matrix Z_n for the nth harmonic in such a way that the off-diagonal terms of the covariance matrix $Z_n^\dagger Z_n$ are all zero. These three methods were called, respectively, the Prime Method, the Good Lattice Point Method, and the Simplified Lattice Point Method. They all depended on the idea of picking the entries of matrix Z_n to be suitable (complex) roots of unity, in such a way that the corresponding sums of roots always came out to zero on the off-diagonal. Moreover, this optimal choice of design angles can be chosen in such a way that works for a range of harmonics.

As a simple example, for a design matrix of size $m = 5$, we use the prime method to choose design angles of the form

$$(\theta_k^j) = \frac{2\pi}{5} \cdot \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 1 & 3 \\ 0 & 3 & 1 & 4 & 2 \\ 0 & 4 & 3 & 2 & 1 \end{pmatrix}$$

and set the matrix entries for the nth harmonic matrix Z_n as

$$(Z_n)_{jk} = e^{in\theta_k^j}.$$

Indeed, we recognize matrix Z_1 as the matrix for a discrete Fourier transform of dimension 5, which is of course orthogonal, and hence $Z_1^\dagger Z_1$ is diagonal, as desired. Moreover, since $\mathbb{Z}/5$ is a field, it is easy to see that for the nth harmonic, the action of n on entries $n(\theta_k^j)$ just permutes the rows around ($n \neq 0 \bmod 5$), so this design works equally for all harmonics n which are not multiples of 5.

In general, for m equal to any prime, the Prime Method prescribes the design matrix of angles to be an $m \times m$ matrix with entries

$$\theta_k^j = \frac{2\pi}{m}\{(j-1)(k-1) \bmod m\}.$$

Again, this design gives optimal Z_n for any prime m and any harmonic n which is not a multiple of m.

Theorem 1 (THE PRIME METHOD). *For integer m prime, n not a multiple of m,and design angles chosen as*

$$\theta_k^j = \frac{2\pi}{m}\{(j-1)(k-1) \bmod m\},$$

then the corresponding design matrix Z_n with entries

$$(Z_n)_{jk} = e^{in\theta_k^j}$$

is optimal.

Proof. A proof is provided in reference [1], but it is easy enough to observe that matrix Z_n is a DFT matrix of size $m \times m$, with rows permutated around by multiplication by n. □

The Prime Method is a powerful method both in that it gives explicitly an optimal solution and it works for a range of harmonics. However, it has some disadvantages. It is somewhat inflexible in the number of layers m, as m must be prime. It requires using as many angles as there are layers. This can be an expensive, if not impossible, construction in some tire plants. In addition, there is the problem of setting angles exactly.

The second method is inspired by the Good Lattice Point (GLP) method of Fang and Wang [5], which uses a careful choice of lattice points in an m-dimensional hypercube to accelerate integration over a multidimensional Riemann sum. The basic insight is to look for angle combinations which will lead to sequences $\{\theta_k^r - \theta_j^r\}_{r=1}^R$ which will allow for fast convergence of the sum

$$\sum_{r=1}^{R} \frac{1}{R} e^{ni(\theta_k^r - \theta_j^r)} \to \int_0^1 e^{2\pi i n x} dx.$$

With the Monte Carlo method, random sampling of the hypercube produces random sequences on $[0,1]$, but with a convergence of order $R^{-1/2}$. The GLP method will exhibit convergence at the faster rate of $R^{-1}\log^m(R)$. We tested the GLP method to see if we would obtain a good sequence of angles, and came up with a surprising conjecture.

First, recall the definition of a lattice point set and a GLP set. Let $(R; h_1, h_2, \ldots, h_m)$ be a vector of positive integers satisfying i) $m < R$, ii)$1 \leq h_j < R$, iii) $h_j \neq h_k$, for all $j \neq k$, iv) (h_j, R) are coprime, for all j.

The *lattice point set* of the generating vector $(R; h_1, h_2, \ldots, h_m)$ is the set of vectors $\{(x_{r1}, \cdots, x_{rm}), r = 1, \cdots, R\}$, with values

$$x_{rj} = \mathrm{frac}\left(\frac{2rh_j - 1}{2R}\right),$$

where "frac" denotes the fractional part of the given real number. If this set has the smallest discrepancy (defined in Fang and Wang), it is called a GLP set.

The principle behind GLP sets is that generating vectors can always be constructed so that a GLP set is created, whose points are uniformly distributed about the hypercube. Fang and Wang tabulate many different choices for a range of R and m, corresponding in our case to numbers of tires R in the experimental design, and number of layers m per tire.

In our tire example, the vector of angles are obtained from the lattice point sets by scaling by a factor of 2π, so $\theta_k^j = 2\pi x_{jk}$. We tested a number of the GLPs from the book to determine how close they are to being optimal. In every instance the GLP set was optimal. We have the following:

Conjecture 1 (GLP METHOD). *Every GLP set produces an optimal design. That is, for each harmonic n as above, the design matrix Z_n satisfies*

$$(Z_n^\dagger Z_n)^{-1} = \frac{1}{R} I_m$$

exactly. Moreover,

- *m can be chosen arbitrarily* (*not necessarily prime*)
- *R can be chosen arbitrarity* (*although prime is a popular choice*)
- *the same design is optimal for all harmonics co-prime with R.*

In the workshop, there was not enough time to explore how GLP sets were constructed in the literature, so it was not clear to us how optimal designs were resulting from these choices. A quick review of work in the area indicates some number theoretical results are being used to construct the charts of Fan and Wang. However, a simple examination shows the differences $x_{rj} - x_{rk} = \mathrm{frac}(r(h_j - h_k)/R)$, so as in the prime method, the sum in the covariance matrix will cycle around a subset of the R roots of unity. For a good choice of the h_j, this subset will always sum to zero. Thus, while this is short of a proof verifying the GLP method works, there is the basis for another useful technique, the Simplified Lattice Method.

The third method we developed in the 2000 workshop is called the Simplified Lattice Method, which is also inspired by the GLP method of Fang and Wang, but for which we could provide a proof of optimality.

Theorem 2 (SLP METHOD). *Fix an integer $m > 0$ and fix $\mathbf{N}$ a subset of $\{1, 2, \ldots\}$. Suppose $(R; h_1, h_2, \ldots, h_m)$ with $m < R$ is a vector of positive integers such that R is not a divisor of $n(h_j - h_k)$ for any $n \in \mathbf{N}, j \neq k$. Then the vectors of angles* (*scaled lattice points*) *defined by*

$$\theta_k^j = \frac{2\pi}{R}(j \cdot h_k \bmod R), \qquad 1 \leq j \leq R, 1 \leq k \leq m$$

gives an optimal design for all harmonics $n \in \mathbf{N}$. That is, the design matrix Z_n satisfies

$$(Z_n^\dagger Z_n)^{-1} = \frac{1}{R} I_m.$$

Proof. For harmonic n in the set $\mathbf{N}$ and $j \neq k$, the integer $n(h_j - h_k)$ is not divisible by R and hence the map $r \mapsto n(h_j - h_k)r \bmod R$ defines an endomorphism on the ring $\mathbb{Z}/R$ which has more than one element in its range, which is a subring of $\mathbb{Z}/R$. Thus when scaled by $2\pi i$ and exponentiated, one obtains some r roots of unity, for some divisor $r > 1$ of R. Thus the sum

$$\sum_{r=1}^{R} e^{2\pi n r(h_k - h_j)/R}$$

simply cycles around these r roots of unity, and so sum to zero. Hence the off-diagonal terms of the covariance matrix $X^\dagger X$ are zero, the diagonal terms are R, and thus the optimal design is achieved. □

These criteria are easy to fulfill in any situation of tires, as shown in the following.

Example 1. *With m the number of layers in the tire, and $\mathbf{N}$ a finite set of harmonics, let R be any prime number strictly bigger than m and all integers $n \in \mathbf{N}$. Then the integer vector $(R; 0, 1, 2, \ldots, m-1)$ generates lattice points yielding an optimal design.*

This example gives a method much like the original prime method described above. However, the number of layers m need not be prime, and one can select any finite set of harmonics, yet still obtain an optimal design. The number of tires R need not be prime: one could choose a composite number with some prime factor bigger than m and all n. There remains the disadvantage that almost every layer on almost every tire must be set to a non-zero angle, and the angles must be set to accuracies on the order of $2\pi/R$.

1.8 Problem 1: Extending the Simplified Lattice Method

All the methods developed above involve choosing design angles θ_k^j as simple fractions scaled by 2π, which then determine a matrix Z_n with entries that are roots of unity. During the current workshop, we experimented with a number of other choices for design angles and roots. For instance, we have this sense that distributing the roots randomly, but more or less uniformly around the unit circle might be a good idea. We tried taking multiples of an irrational for the design angles, say

$$\theta_k^j = j \cdot k \cdot \frac{1+\sqrt{5}}{2}$$

as multiples of the golden ratio, or multiples of $\sqrt{2}$, and so forth. The combinatorist in our group tried assorted permutations of roots of unity to try to obtain maximum orthogonal columns for the matrix.

However, our numerical tests suggested nothing we did was better than taking a simple DFT matrix, and modifying for harmonics. That is, for a given matrix size R, we pick design angles

$$\theta_k^j = 2\pi \frac{(j-1)(k-1)}{R}, \qquad 1 \leq j, k \leq R$$

which results in the nth harmonic matrix Z_n with entries

$$(Z_n)_{jk} = e^{in\theta_k^j}.$$

Thus Z_1 is the DFT matrix, and the other Z_n matrices have rows (and columns) which are just a permuted subset of those of Z_1. If it is a full subset, we obtain an orthogonal matrix, and hence an optimal design.

This full subset is obtained precisely when the action of n on the fractions $(j-1)(k-1)/R$ is just a permutation. A moments thought on divisors of R and n shows the following is true.

Theorem 3 (DFT MATRIX METHOD). *Fix an integer $R > 0$ and choose design angles*

$$\theta_k^j = 2\pi \frac{(j-1)(k-1)}{R}, \qquad 1 \le j, k \le R$$

and corresponding nth harmonic matrix Z_n with

$$(Z_n)_{jk} = e^{in\theta_k^j}.$$

Then the design matrix is optimal ($(Z_n^\dagger Z_n)^{-1} = I/R$) if and only if n and R are relatively prime.

So, for instance, for an experiment with $R = 20$ tires, and $m = 20$ layers under control, we obtain an optimal design precisely for those integers n that have no common divisors with 20. For instance, that would be

$$n = 1, 3, 7, 9, 11, 13, 17, 19, 21, 23, 27, \ldots .$$

In particular we can determine all the odd harmonics which are not a multiple of 5. Notice the harmonic need not be smaller than $R = 20$, just relatively prime to it.

As another example, with $R = 16$ tires, and $m = 16$ layers under control, we obtain optimal designs for harmonics

$$n = 1, 3, 5, 7, 9, \ldots .$$

That is, for all odd harmonics (including those larger than 20.)

In the original problem, we were asked to design for harmonics $n = 1, 2, 3, 4, 5$. Trying to find a small number R that is relatively prime to these 5 numbers (simultaneously), and also small, typically gives R as a prime number. For instance, the possible R relatively prime to $1, \ldots, 5$ are

$$R = 1, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 49, 53, \ldots$$

and we see the first non-prime solution (other than 1) is $R = 49$. This explains why the Good Lattice Point method we described above often degenerated into the Prime Method.

However, if one is willing to look for only a select set of harmonics (such as just the odds), it is easy to make designs with small R (such as $R = 8$ or $R = 16$).

It also makes sense, then, to design two experiments, say one to recover odd harmonics (eg $R = 8$ recovers harmonics $n = 1, 3, 5, 7, 9, \ldots$ and another to recover most even harmonics (eg. $R = 9$ recovers harmonics $n = 1, 2, 4, 5, 7, 8, 10, 11, 13, \ldots$ which include the even harmonics co-prime to 3. In this way, two experiments with a small number of tires, and a small number of controlled layers, can recover all the harmonics of interest.

We also pursued possible designs where the number of tires R is larger than the number of layers m which are being analysed. Once again, the $R \times R$ DFT matrix is the key idea, and so we select a subset of m columns from that DFT matrix for Z_1 to get an optimal design for the first harmonic. When n is co-prime to R, the nth harmonics matrix Z_n is also optimal, so again we can solve for a range of harmonics n which have no common divisors with R. (The factorization of m itself is irrelevant.) Is this better? Not really, as we have more tires to experiment with than number of layers to determine. But, it might be useful for the technicians designing the tests, as it tells them the range of angles they select in an experiment is determined by the number of tires, not by the number of layers.

Also, there is the possibility of selecting a subset of integers $h_1, h_2, \ldots, h_m$ from the full set $0, 1, \ldots, R-1$ as in the Simplified Lattice Method in such a way that the action of n on the differences $h_j - h_k \mod R$ is non-trivial. In such a case, the design matrix Z_n obtained from corresponding design angles will again be optimal. However, a few examples convinced us, that this was unlikely to lead to very efficient designs (i.e. number of tires R close to the number of layers m under control), and we abandoned this approach.

1.8.1 Extension by blocking

As noted above, there is a germ of an idea on how to cover harmonics of interest using a series of smaller test runs, rather than one large test. The advantages of this method include the simplification in angle setting for the test under consideration. That is, for a test with R tires, the angles to be set will be multiples of $360/R$ degrees, so for R small, we get a manageable set of test angles to work with. As an example, with $R = 8$, our test angles are multiples of 45 degrees, and the experiment will recover harmonics $1, 3, 5, 7, \ldots$. Repeating an experiment with $R = 9$ give angles as multiples of 40 degrees, and will recover harmonics $1, 2, 4, 5, 7, 8, 10, \ldots$. Thus with two experiments, using a total of 17 tires, we recover most the the first ten harmonics (we miss $n = 6$).

True, we could have run just one large test with 17 tires (17 is prime), but this would require test angle setting that are multiples of $360/17$ or 21.18 degrees. This is not a convenient number for real factory settings.

We explored this idea by considering a number of test cases: various layers, various numbers of tires, etc. Unfortunately, we misplaced our records of the cases we studied, but concluded we could often recover a wide range of harmonics by breaking up a large test (with unusual test angles) into two or three smaller tests, with better angles. The same, or slightly larger, number of test tires is all that is required in the sum of the tests. A tabulation of possibilities would be useful, which we could generate with some direction from Michelin researchers on what would be convenient tests.

1.9 Problem 2: Non-harmonic components

The second part of the problem described was to account for non-harmonic components in the measured waveform. We had some difficulty with this concept at first, as we understood the measurement to be cyclic (they represented forces around the circle of the tire) which by definition is periodic. Moreover, from the standard Fourier theory, any periodic waveform can be represented as a sum of harmonics, so there are no “non-harmonic” components. That is, we felt there would never be such a

thing as a non-harmonic component, since the harmonic components were enough to describe everything.

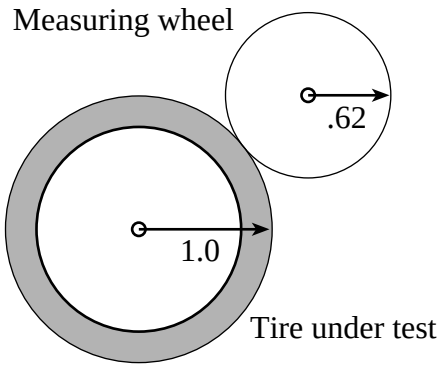


Figure 1.1: The force measuring wheel as it contacts the test wheel. Note the difference in dimension for the wheels.

However, this is a case of knowing too much mathematics for our own good! After a careful description of the construction and measurement methods used by Michelin for its tires, it became clear physically that such "non-harmonic" components could exist. For instance, as shown in Figure 1.1, the forces around the tire are measured by rolling a smaller wheel around the edge of the tire. Since the size of the measurement wheel might not be the same size as the tire under test, nor a simple ratio, the measurement process is not truly periodic! For example, the measurement wheel may not be perfectly round (neither is the test tire), and as the two roll against each other, the imperfections line up at different places, which may not be a periodic function of the test tire's circumference.

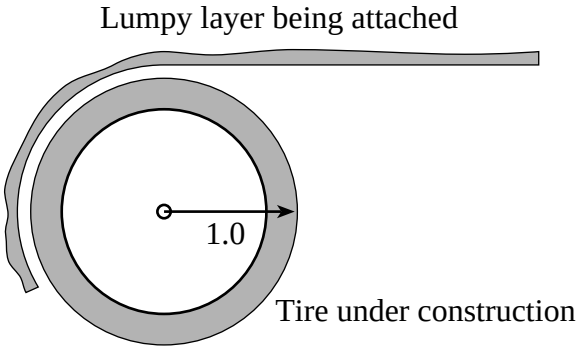


Figure 1.2: A lumpy, extruded layer being wrapped around the wheel.

Similarly, when the tire is being constructed, the component layers (rubber, cords, steel) may have some regular imperfections along their length – they may be "lumpy" for some reason, such as because of the extrusion process by which these layers are manufactured. The lumps, although regular, may not line up exactly along the circumference of the tire, and thus their forcing function is not really periodic. This is illustrated in Figure 1.2.

Although the harmonics are complete, we asked the question "what is the effect of analysing a non-harmonic component using only harmonics?" More precisely, we experimented with taking the usual discrete FFT of a signal that had one non-harmonic component, and observed its effects on the harmonic decomposition. For instance, in Figure 1.3, we see the first 10 Fourier components for a signal of the form $f(t) = \sin(2\pi t/.62)$ which can be thought of as a signal of period .62, or frequency 1.61 times the fundamental. Observe that, rather than a single spike (which we would expect of a signal with integer frequency multiple), we obtain a smear of energy in the first five or so low frequency components. For our tire problem, this says that a perfectly smooth tire, with a wobbly measurement wheel, would falsely exhibit large harmonics for $n = 1, 2, 3, 4, 5$. This is exactly the problem we wish to avoid.

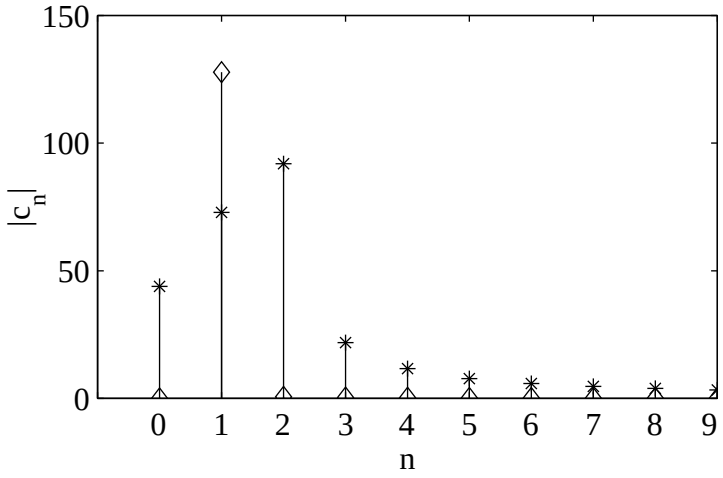


Figure 1.3: A stem plot of the modulus of the components of the Fourier spectrum for a non-harmonic signal with period 0.62. Rather than a single spike at frequency 1.61, we see it contaminates nearby frequency components $n = 0, 1, 2, 3, 4$. The diamonds correspond to the components of the waveform $\sin(2\pi t)$.

We also observed similar effects for the extrusion-type non-periodic variations, with contamination centred around frequencies close to the base frequency of the extrusion regularities.

This problem then has two parts: how do we detect and measure the strength of a non-harmonic component, and how do we design our experiments to account for these effects. Measuring the strength of the non-harmonic component is rather straight-forward. Once we know to expect such a component, we simply include it in our Fourier series decomposition of the force signature. For instance, in the case where the measurement wheel has a diameter .62 times the diameter of the tire under test, we expect a non-harmonic component at frequency 1/.62 times the fundamental. For $2N = 256$ samples the series expansion for $f(t)$ is taken to be

$$f(t_k) = \sum_{n=-127}^{128} c_n e^{2\pi i n t_k} + c_* e^{2\pi i t_k/.62} + c_{**} e^{-2\pi i t_k/.62}, \qquad t_k = \left\{ \frac{k}{256} \right\}_{k=0}^{255}.$$

The coefficients c_*, c_{**} represent the strength of the non-harmonic component. Of course, this is now an over-determined system of equations (256 equations in 258 unknowns), so to select a useful

solution, we pick the one that minimizes the L^1 norm of the coefficients

$$\sum_{n=-127}^{128} |c_n| + |c_*| + |c_{**}|.$$

The choice of the L^1 norm forces the minimization routine to select a solution that is sparse; i.e. it concentrates the energy of the decomposition into as few components as possible. This is a well-known approach in frame theory.

We did some experiments in MATLAB with this approach and the relatively small dimensions ($d = 258$) gives a problem that is within the capabilities of the software. But, it is rather slow. Another approach is to use an undertermined system, by trying to express the force signature $f(t)$ as a sum of just a few harmonic and non-harmonic terms. For instance, using just the first 5 harmonic components, and one non-harmonic component, one can attempt to minimize

$$\left\| f(t_k) - \sum_{n=-5}^{5} c_n e^{2\pi i n t_k} - c_* e^{2\pi i \frac{t_k}{.62}} - c_{**} e^{-2\pi i \frac{t_k}{.62}} \right\|_2 + \alpha \left(\sum_{n=-5}^{5} |c_n| + |c_*| + |c_{**}| \right).$$

Here, we use the usual L^2 norm to measure the distance between the force f and its representation as a sum of harmonic and non-harmonic components and include a penalty term (scaled by α) which aims for sparseness of the representation.

When there are several non-harmonic components, we simply add on addition terms in the decomposition above. Note that this discussion assumes we know the frequency of the non-harmonic component. This would be the case of a measurement wheel of known dimension, or adding layers to the tire that have regular undulations (say due to extrusion) of known period. However, it is also possible to account for non-harmonic components of unspecfied frequency, by including terms like $e^{2\pi i \omega_n t}$ where ω_n is a parameter to be varied under the minimization step above. Of course, this is a much more challenging problem numerically.

Finally, we observed at the workshop that the usual Fourier decomposition of a signal can be generalized to a decomposition of any choice of basis functions with (non-harmonic) frequencies. We select some range of frequencies $\omega_n, n = 1, 2, \ldots, 256$ and solve the equation

$$f(t) = \sum_n c_n e^{2\pi i \omega_n t}.$$

For regularly spaced t, this is a Vandermonde matrix equation, for which there are fast linear solvers, and thus we need not shy away from inverting this 256×256 system of equations.

There remains the question of how to incorporate this non-harmonic component in the experimental design. For the problem with the measuring wheel, we are of the opinion that the non-harmonic component arising from measurement only serves to contaminate the true harmonics present in the tire under test. Thus, the approach should be to estimate the non-harmonic component using the techniques described here, subtract it from the signal, then perform the experiment as before. For the problem of non-harmonic components coming from extrusions, it seems these parts really do contribute to the performance of the tire – they are not simply a part of the measurement process. To design an experiment in this case, one must keep track not only of the position of the layer on the tire, but also the relative placement of the extruded irregularities. We did not have enough time to explore this in detail.

1.10 Problem 3: Combining waveforms

The third problem essentially asks: what happens when you measure several force waveforms? This would be the case, say, when several measuring wheels are used on one tire, perhaps one measuring radially, another perpendicular to the tread. Or, one might measure forces along three or four paths around the tread of the tire. The measured waveforms may be recorded separately, or they may be combined, say as a sum, or as a sum of squares.

After briefly considering this problem, we realized it had been essentially solved by the work above. The experimental design asks how to build test tires in order to determine the contribution of each component layer to some force profile $f(t)$. The analysis above generalizes from the case of a real-valued function (a single force), to the case of a vector-valued function (several forces, recorded simultaneously). There is no interaction between components, so we proceed with test tires as before. The analysis is then performed component-by-component. That is, we just look at each measured force as an independent result, and apply the methods discussed above.

In the case where the forces are summed together before analysis, we are back to the simple case of a single (real-valued) force function $f(t)$, and once again the earlier analysis works.

The complications occur when we take a sum of squares, for instance $f^2(t) + g^2(t)$. Of course, this results in a single real-valued function to analyse, but we have to observe that our Fourier decomposition in $f(t), g(t)$ will combine non-linearly. If we do a Fourier analysis of just the combined function, we expect some problems, since the underlying physical problem is now non-linear, and there should be interaction between the harmonics. In particular, this says our basic assumption of the linear model described in the earlier section will fail, and so we do not expect great success by this method.

On the other hand, we found it difficult to understand why one might combine force measurements in such a way. Since the analysis works for vector functions, we would recommend to Michelin to simple recover all measured force waveforms and work with them independently, rather than as some combined, lump sum result.

1.11 Problem 4: Extending to 2 dimensions

The original problem of 2000 looked at variations in the tire causes by translation of layers around the circumference of the tire. Problem 4 suggested we also consider variations caused by lateral displacement of the layers; that is, what happens when a layer is placed in different positions, varying in a direction perpendicular to the circumference of the tire. Michelin explained this could happen when laying down the steel belts - this layer could be position slightly to the left of centre-line, or slightly to the right. It is also possible to put down a layer at an angle, so the belt is to the left on part of the circumference, to the right on another part.

Several members of our groups spent quite a bit of time trying to extend our linear analysis described above to the two dimensional case. Namely, data would be recorded and a two-dimensional FFT would be attempted. At some point, we realized the 2D FFT model was really not appropriate, as the two dimensions are fundamentally different. For, around the tire, we have periodicity around the full circle. But, in the perpendicular direction, the layers can only be moved small amounts, and certainly never all the way around the cylinder of the tire. More precisely, we think of the tire as a two dimensional torus, or donut. We can move layers all the way around the large circumference of the

torus, but only slightly in the cross-sectional direction.

A more appropriate model then would be to treat the circumferential direction via a 1D FFT, and model the variations in the transverse direction via a small order polynomial. That is, we expect small transverse movements in a layer to have a small effect on the harmonics, which could be modeled by a simple polynomial. Indeed, we could use a couple of small parameters to model two effects for each later: the distance from centre-line for the layer, and a "torque" amount, representing the amount of transverse twisting for the layer. Other effect due to small movements could be modeled similarly.

With this idea in mind, we see the experimental design utilizing FFT matrices in one dimension, and Vandermonde matrices (arising from the powers of x in the polynomial approximations) in the other dimension. We expect the two modes to be decoupled, so solving them would be straightforward.

However, we lacked sufficient time in the workshop to pursue all the details of this method. We believe the method is viable and worth pursuing.

1.12 Conclusions

Michelin presented us with four problems in the experimental design of a statistical test on manufactured tires. By extending the linear, Fourier model developed at the IPSW in 2000, we were able to make remarkable progress on the four problems, essentially resolving the first three, and suggesting a promising direction to resolve the fourth problem. It was a challenging and rewarding set of problems for our group to attack.

1.13 Acknowlegements

We would like to recognize the support from the Pacific Institute for the Mathematical Sciences, which organizes and funds these Industrial Problem Solving Workshops, and the University of British Columbia for hosting this year's workshop. We also would like to thank the people at the Michelin Corporation, and in particular Bill Mawby, for participating in the workshop, proposing the problems, and supporting our research.

Bibliography

[1] Aggarwala, R., Lamoureux, M. & Powojowski, M. (2002). *Statistical Design in Tire Manufacturing*. Canadian Applied Mathematics Quarterly, 10(2), pp. 179-198.

[2] Fang, K. T. & Wang, Y. (1994). *Number-theoretic Methods in Statistics*. Chapman & Hall: London.

[3] Körner, T. W. (1988). *Fourier Analysis*. Cambridge University Press: Cambridge.

[4] St. John, R. C. & Draper, N. R. (1975). *D-Optimality for Designs: A Review*. Technometrics, 17(1), pp. 15-22.

[5] Winker, P. & Fang, K. T. (1997). *Applications of Threshold-Accepting to the Evaluation of Discrepancy of a Set of Points*. SIAM Journal of Numerical. Analysis, 34(5), pp. 2028-2042.

Chapter 2

Analyzing Network Traffic for Malicious Hacker Activity

Problem Presented By: Surrey Kim (Random Knowledge Inc.)

Mentors: Hongwei Long and Weiguang Shi (University of Alberta); Randall Pyke (University College of the Fraser Valley); Lang Wu (University of British Columbia)

Student Participants: Stanislava Peker (Concordia University); Benjamin Chan, Radu Haiduc and, Andrei Maxim (Cornell University); Vilen Abramov (Kent State University); Bo Zeng (Purdue University); Pengpeng Wang (Simon Fraser University); Robert Liao, Yury Petrachenko, Yulia Romaniuk, Mengzhe Wang, Zhian Wang and Mohammad Ali Yassaei (University of Alberta); Gabriel Mititica, Shijun Song and Xuekui Zhang (University of British Columbia); Song Li and Tzvetalin Vassilev (University of Saskatchewan); Nancy Azer, Mahin Salmani and Jiaping Zhu (University of Victoria)

Report prepared by: Randall Pyke (pyke@math.toronto.edu)

2.1 Problem Statement

Network security is still at its infancy. Existing intrusion detection and prevention solutions lack accuracy, broad attack coverage, speed, performance, and scalability. They do not provide reliable protection to today's vital networks.

Random Knowledge Inc.'s approach to intrusion detection is to apply Mathematically Optimal Detection that outperforms other methods, including pattern matching, neural networks and statistical techniques. This detection system, Portscan Detection System (PDS), detects and localizes traffic patterns consistent with possibly stealthy forms of attacks from within hoards of legitimate traffic. With the network's packet traffic stream being its input, PDS relies on high fidelity models for normal traffic from which it can critically judge the legitimacy of any substream of packet traffic.

In this modelling workshop, we try to characterize normal traffic which involves:

- Defining all the different types of connection sessions.

- Verification of a Poisson measure model for the incoming connection sessions, i.e. if the connection session types are labelled $1, \ldots, n$, determining if $N(A \times (0, t])$ is Poisson distributed for any subset A of $\{1, \ldots, n\}$, where N is the Poisson measure.

- Determining the rates for $N(A \times (0, t])$ or equivalently its mean measure if the session generation indeed conforms reasonably to the Poisson measure model, otherwise suggesting other suitable models.

- Verification for self-similar processes and heavy tailed distributions within connection sessions (for example the transmission time), and the estimation of its parameters.

Hitherto, there has been much study of traffic characterization that focuses on the implications for improved network performance. Random Knowledge's approach is the study of traffic characterization for the implications of detecting malicious hacker activity.

2.2 Glossary

- Normal traffic: Traffic generated by legitimate users.

- Mark: The IP address and port number pair. Each mark is a server with a specific service.

- Session: The analog of a phone call between a client (source IP) and a mark (the HTTP server). It starts when the first packet from a client arrives at a mark and ends if the silent (no packet) time is longer than a preset time-out value. Note: One session could include transmission of multiple packets.

- Interarrival time: Time between the arrivals of two consecutive sessions.

2.3 Introduction

2.3.1 What is the Problem?

Since the Internet came into life in the 1970s, it has been growing more than 100% every year. On the other hand, the solutions to detecting network intrusion are far outpaced. The economic impact of malicious attacks in lost revenue to a single e-commerce company can vary from 66 thousand up to 53 million US dollars [1]. At the same time, there is no effective mathematical model widely available to distinguish anomaly network behaviours such as port scanning, system exploring, virus and worm propagation from normal traffic.

PDS proposed by Random Knowledge Inc., detects and localizes traffic patterns consistent with attacks hidden within large amounts of legitimate traffic. With the network's packet traffic stream being its input, PDS relies on high fidelity models for normal traffic from which it can critically judge the legitimacy of any substream of packet traffic. Because of the reliability on an accurate baseline model for normal network traffic, in this workshop, we concentrate on modelling normal network traffic with a Poisson process.

2.3.2 Data Description

The dataset is a record of network traffic of the University of Auckland, New Zealand in March 2001. The arrival times are recorded to the nearest nanosecond and are collected with IP addresses and port numbers. Each combination of IP and port denotes a network service provided by a specific server of the University. The pair of IP address 5122 and port 80, for example, denotes a server with IP 5122 providing HTTP web service. In this workshop, we analyzed the traffics of IP addresses 5122, 5226 and 5264 and services ftp (ports 20, 21), telnet (port 23) and HTTP (port 80).

The record also contains both inbound and outbound network traffic of the University. For our purpose of modelling normal traffic from outside, only the inbound traffic is relevant and considered. Note that a relatively small portion, about 7% of half a day, of the traffic data of IP 5264 is missing. Data analysis of this server was conducted on the reduced dataset.

2.3.3 Goals

The primary motivation of this work is to justify the modelling of normal session traffic with a non-homogeneous Poisson process. We justify this by noting that i) service usages of different clients are expected to be independent events and ii) the number of packets having arrived by time t is a counting process $\{N(t), t \geq 0\}$ whereby the number of events in an interval of length t is $E(N(t)) = \lambda(t)t$, $\lambda > 0$. The actual recorded session traffic will be used to model the normal network traffic which may include some malicious network behaviour.

One method used to identify vulnerable ports of a network service system is to send a sequence of probing packets to all available ports over a relatively short period of time. This reconnaissance behaviour identifies which ports of a network are open and which services have been made available. In the traditional network traffic model using packets, port scanning takes up a tiny portion of the traffic and is difficult to detect. By grouping the packets of each session together a probing session will violate the assumption of independence of the arrival times across the ports of the network. This violation of independence allows one to identify of this type of malicious behaviour much more efficiently.

To justify the usage of the Poisson process model we note that the sessions representing different service requests are independent events. However it is known that the arrival rate can be considered constant only over a relatively short (five minutes) interval. Extensions beyond this short interval do not model modern servers very well. A rather comprehensive critique of the use of a single homogeneous Poisson process in modelling network arrivals can be found in [6]. Rather than a single Poisson process, we assume that the overall session traffic, namely session arrivals, can be modelled as a sequence of homogeneous Poisson processes. Alternatively this can be thought of as a nonhomogeneous Poisson process where the arrival rate $\lambda(t)$ is piecewise constant.

We can break up the model characterization into the following tasks:

1. Analyze the arrival patterns of normal traffic of different service types (marks) of connection sessions. (A mark is defined as the IP address and port pair. It can be also thought of as a specific service running on a server.)

2. Test if the arrivals are nonhomogeneous Poisson. If they are indeed nonhomogeneous Poisson, then estimate the relevant parameters of the models such as mean and standard deviation, also maximize time interval within which the arrival rate is constant.

3. Otherwise, suggest other suitable model and estimating its parameters.

2.4 Methodology

In this section we will describe the statistical framework for testing the nonhomogeneous Poisson model. Because of the close relationship between the exponential and Poisson distributions (detailed below) we test if the session arrivals are Poisson distributed by checking that the interarrival times are i) exponentially distributed and ii) independent. In the following we begin by briefly reviewing the definitions of the Poisson and exponential distributions and their relationship. To test for a non-homogeneous Poisson distribution both a Goodness-of-Fit test and a test for independence are then described. More sophisticated nonparametric techniques for estimating the model parameters of a nonhomogeneous Poisson model are described in detail in [4].

2.4.1 Poisson and Exponential Distributions

A Poisson distribution is typically used to model the number of events happening per time interval, such as the number of customers arriving at a store per hour, or the number of visits per minute to an internet site. A random variable (r.v.) N that takes values $0, 1, 2, ...$ has λ-Poisson distribution if

$$P(N = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots . \tag{2.1}$$

The mean and variance of the Poisson distribution both equal λ.

A continuous random variable X is said to have an exponential distribution with parameter λ, $(\lambda > 0)$, if its probability distribution is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

Equivalently the cumulative distribution function $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$; $F(x) = 0$, $x < 0$. The mean and variance of the exponential distribution are $1/\lambda$ and $1/\lambda^2$ respectively.

Recall that a counting process $\{N(t), t \geq 0\}$ is said to be a Poisson process having rate λ, $\lambda > 0$, provided

1. $N(0) = 0$.

2. The process has independent increments.

3. The number of events in an interval of length t is Poisson distributed with mean λt.

A nonhomogeneous Poisson process allows λ to be a function of time and as a result, allows for the possibility that events may be more likely to occur at certain times during the day.

Consider a Poisson process and let T_1 denote the time of the first event. In addition let T_n denote the time elapsed between the nth and $(n-1)$st events. To determine the distribution of the sequence

of interarrival times $\{T_n\}$ note that the event $(T_1 > t)$ occurs only if no events of the Poisson process occur over the interval $[0, t]$. As a result,

$$P(T_1 > t) = P(N(t) = 0) = e^{-\lambda t}$$

so that T_1 is exponentially distributed with mean $1/\lambda$. The distribution of T_2 is obtained by computing

$$\begin{aligned} P(T_2 > t | T_1 = s) &= P(0 \text{ events in } (s, s+t] | T_2 = s) \\ &= P(0 \text{ events in } (s, s+t]) = e^{-\lambda t}. \end{aligned}$$

Repeating the same argument we see that the $\{T_n\}$ are independent identically distributed exponential random variables with mean $1/\lambda$.

2.4.2 Goodness-of-fit Test

To test that interarrival times between sessions are λ-exponentially distributed, we use Anderson-Darling (A-D) test, which checks if a given sample is drawn from a population with a specified distribution [2, 3]. There are other methods to test for goodness-of-fit, such as Kolmogorov-Smirnov (K-S) or chi-square tests. However, A-D test is more appropriate in our case, as it doesn't require the true population parameters, but uses those estimated from available data [2, 3]. The test statistic is $A^2 = -N - S$, where N is the sample size, and

$$S = \sum_{i=1}^{N} \frac{2i-1}{N} \left[\ln(F(Y_i)) + \ln(F(Y_{N+1-i})) \right]. \tag{2.2}$$

In the above equation, Y_i are sample values (sorted in ascending order), and F is the cumulative distribution function of the specified distribution (in our case, exponential with $\lambda = 1/\bar{Y}$, $\bar{Y} = \sum_{i=1}^{N} Y_i/N$). As the A-D test uses an estimated mean, A^2 has to be multiplied by a correction factor, so that the actual used statistic is $A_*^2 = A^2 \cdot (1 + .6/N)$. The null hypothesis that the sample is drawn from a given distribution is rejected if $A_*^2 \geq 1.341$ at the 95% significance level [7].

The data files contained session interarrival times for a six-hour trace of internet traffic. The sessions were determined based on the time-out values. Each six-hour interval was subdivided into 5-minute subintervals to test whether λ was constant during this subinterval (A-D test).

2.4.3 Independence Test

Next, we test whether the interarrival times were independent within each time interval, as well as between the first lag of the 5-minute subintervals. For this, we used the autocorrelation function: given measurements $Y_1, Y_2, ..., Y_n$ at times $t_1, t_2, ..., t_n$,

$$r_k = \frac{\sum_{i=1}^{N-k} (Y_i - \overline{Y})(Y_{i+k} - \overline{Y})}{\sum_{i=1}^{N} (Y_i - \overline{Y})^2}, \qquad \overline{Y} = \frac{1}{N} \sum_{i=1}^{N} Y_i. \tag{2.3}$$

Autocorrelation is significant in two ways: it can be too strong in magnitude or too frequently positive/negative. Note that for a time series of N samples from an uncorrelated white noise process, $P(r_k > 1.96/\sqrt{N}) = .05$.

We use binomial probability distribution with parameter $p = .95$ (the probability of success) to check for independence of interarrival times. The density of the binomial r.v. X, representing the number of successes in N trials, is:

$$P(X = k) = \binom{N}{k} p^k (1-p)^{N-k}, \qquad k = 0, 1, \ldots, N. \tag{2.4}$$

Within each 5-minute subinterval, we calculated autocorrelation values and checked (using binomial distribution) whether sufficiently many of them were below $1.96/\sqrt{N}$ to claim that the interarrival times were uncorrelated. So if $P(X = k) \leq .05$, we reject the hypothesis that the interarrival times are independent. Similarly, we performed the binomial tests to check for the significance of the autocorrelation sign (using $p = .5$ and a rejection region of .025) within each subinterval and between the lags for the subintervals of the entire six hour trace.

Alternatively we could have tested the null hypothesis that the arrivals are independent in a given time series using the Box-Pierce or Ljung-Box tests [5]. The advantage of the A-D test lies in its robustness and the fact that it does not assume the arrival times are normally distributed.

2.5 Results

Our test results are based on the three servers that have the highest traffic volume, namely, web servers with IP addresses 5122, 5226 and 5264. This is a starting point about which we can frame a set of algorithms for model diagnostics. Later, these algorithms can be systematically extended to the other servers.

2.5.1 Arrival Pattern Suggesting Strong Periodicity in Interarrival Time

Figure 2.1 shows strong periodicity in session interarrival times. This periodicity suggests that a piecewise constant interarrival rate is a reasonable first assumption. (It would be of interest to use the periodicity to predict future interarrival times.)

2.5.2 Statistical Tests Reasonably Confirming Nonhomogeneous Poisson

Based on periodicity, our test procedures assume that interarrival times of a given server will depend on time of day. Recall that our null hypothesis is that the arrivals follow a nonhomogeneous Poisson process. It is a convincing indicator of a nonhomogeneous Poisson process if the interarrival times follow an exponential distribution with a moderately time varying rate. For convenience we used intervals of equal length and started with 5 minutes as basic unit.

Table 2.1 summarizes the results of K-S and A-D tests that the interarrival times are distributed exponentially on the time scale of five minutes. As a result there is no strong evidence to reject the claim that the arrival of session requests follows a nonhomogeneous Poisson process with a piecewise constant arrival rate.

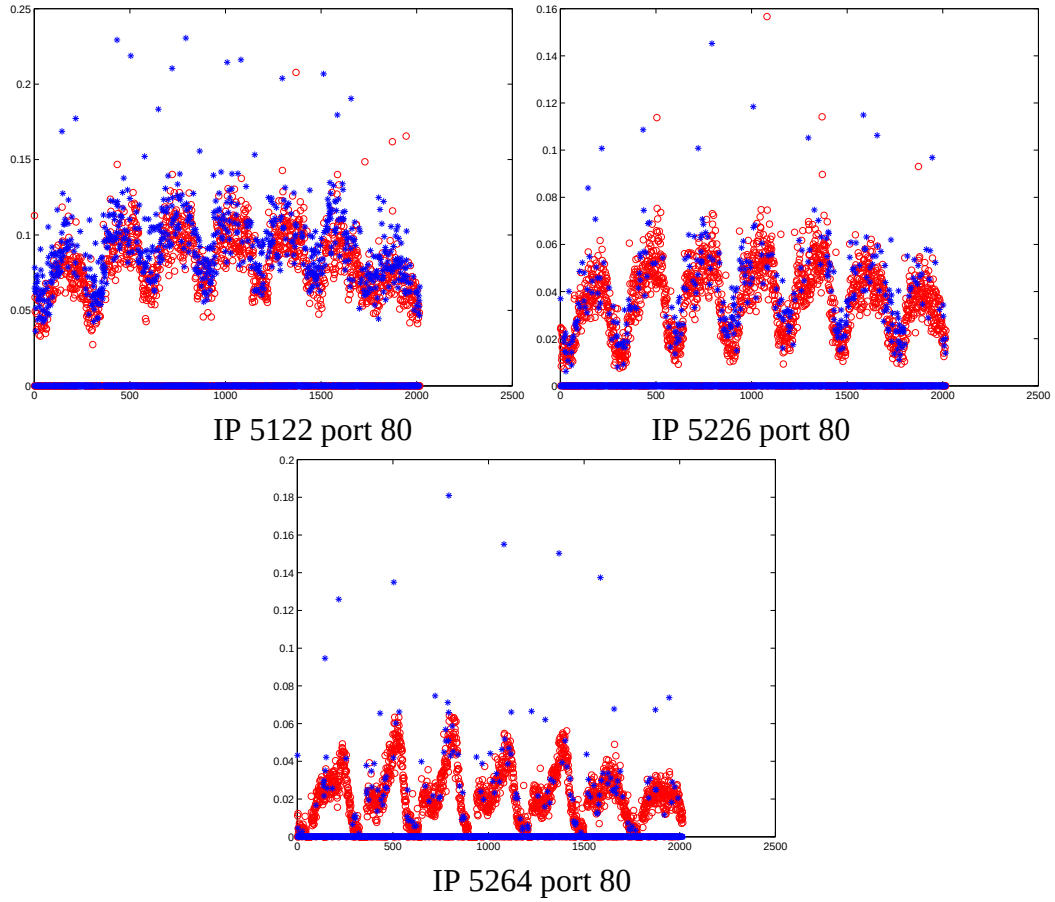


Figure 2.1: The data plot of the session interarrival time of different marks.

To further validate independence amongst interarrival times for all of the tested servers, autocorrelations, shown in Figure 2.2, are not significant either in magnitude or in too frequently positive/negative. These figures can be viewed as an analog of a confidence interval depicting how wide the autocorrelations spread out from the zero line.

In addition to visual validation, the tests on individual points by Box-Pierce statistics does not show strong evidence against the independence hypothesis. Table 2.1 provides a percentage of points that pass the independence test.

2.6 Conclusion and Future Work

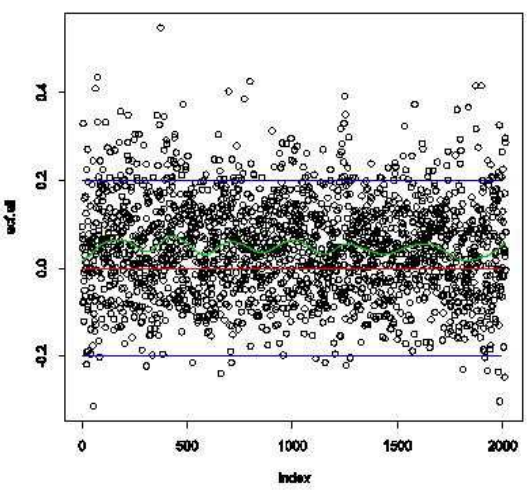
In this report, we have tested the nonhomogeneous Poisson process by validating if the interarrival times were independently exponentially distributed with time varying rates. In most respects the results of three tested servers do not differ noticeably from each other. Both visual and numerical tests have reasonably confirmed an independent exponential distribution.

At this workshop we investigated the modelling of normal session traffic with a nonhomogeneous Poisson process in the specific situation when the arrival rate is piecewise constant. However, there are a few other questions to be answered as to further this anomaly detection study such as: how to

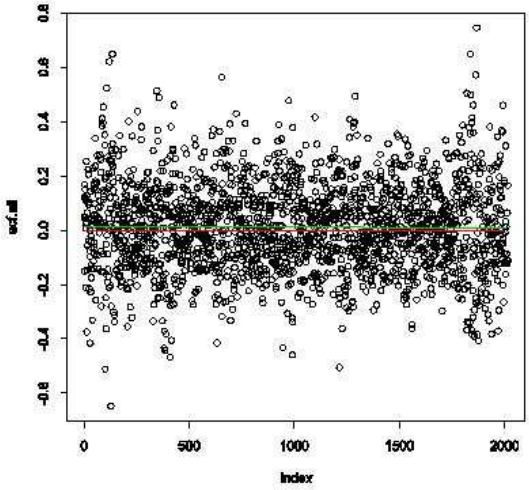
Table 2.1: Percentage of sessions requests that pass.

	IP 5122	**IP 5226**	**IP 5264**
K-S ($\alpha = 1\%$)	84%	93%	97%
A-D ($\alpha = 5\%$)	83%	63%	82%
Independence	96.5%	99.2%	98.2%

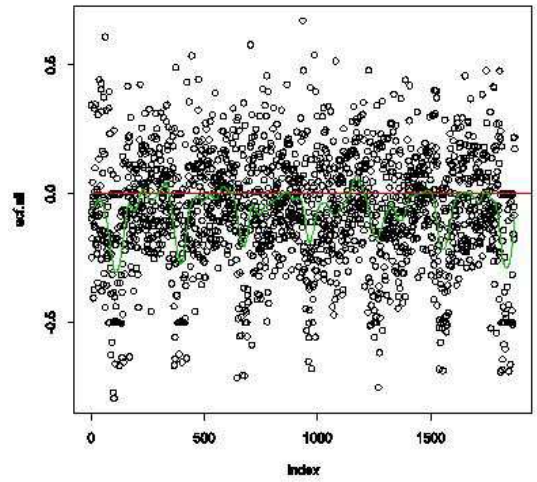
extend the nonhomogeneous model on the other servers; to develop an algorithm by periodicity so as to forecast future arrival patterns; and to apply a method to group marks by dependence. As a result of joint efforts of PIMS and Random Knowledge Inc., some of workshop team members are participating in an ongoing project to develop anomaly detection algorithms.



IP 5122 port 80



IP 5226 port 80



IP 5264 port 80

Figure 2.2: The autocorrelation of the session interarrival time of different marks.

Bibliography

[1] *The Return on Investment for Network Security*, (2002). Cisco Systems Inc.
`http://www.cisco.com/warp/public/cc/so/neso/sqso/roi4_wp.pdf`

[2] Anderson, T. W. and Darling, D. A. (1952). *Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes*. The Annals of Mathematical Statistics, 23(2), pp. 193-212.

[3] Anderson, T. W. and Darling, D. A. (1954). *A test of goodness of fit*. Journal of the American Statistical Association, 49(268), pp. 765-769.

[4] Leemis, L. M. (1991). *Nonparametric estimation of the cumulative intensity function for a nonhomogeneous Poisson process*. Management Science, 37(7), pp. 886-900.

[5] Ljung, G. M. and Box, G. E. P. (1978). *On a measure of lack of fit in time series models*. Biometrika, 65(2), pp. 297-303.

[6] Paxson, V. and Floyd, S. (1995). *Wide-Area Traffic: The Failure of Poisson Modeling*. IEEE/ACM Transactions on Networking, 3(3), pp. 226-244.

[7] Stephens, M. A. (1974). *EDF Statistics for Goodness of Fit and Some Comparisons*, Journal of the American Statistical Association, 69(347), pp. 730-737.

Chapter 3

The Effects of Impact on Design Features

Problem Presented By: John Ceko (Husky Injection Molding Systems)

Mentors: Ian Frigaard (University of British Columbia); Huaxiong Huang and Jinbiao Wu (York University); Greg Lewis (University of Ontario); Tim Myers (University of Cape Town); John Stockie (Simon Fraser University); Rex Westbrook (University of Calgary)

Student Participants: Boyan Bejanov (University of Alberta); Mariana Carrasco-Teja (University of British Columbia); Hooman Javidnia, Olga Krakovska and Azar Shakoori (University of Western Ontario)

Report prepared by: Greg Lewis (`Greg.Lewis@uoit.ca`)

3.1 Introduction

The system of interest is an injection molding machine consisting of a hydraulic piston that forces some molten metal into a mold; see Figure 3.1. The injection piston itself consists of a flange (the tail end), a piston rod, and a screw that is attached to the tip of the piston rod. A variable hydraulic force, that is applied to the piston, attempts to keep the piston moving at a constant velocity, where a maximum force of $F_{\max}$ can be applied. A molten metal (e.g., a magnesium alloy) is injected into a mold, which is pictured at the right side of Figure 3.1. In normal operation the piston screw, travelling at the prescribed velocity, will impact the molten metal that has been injected into the mold and force the molten metal to completely fill the mold. The whole machine is symmetric under rotation about the lengthwise axis of the piston.

It is desired that the machine be designed for (essentially) infinite life. Therefore, design features must be specified so that the machine can withstand the repeated strain on the piston due to the impact of the piston screw on the molten metal. In addition, in the event that there is an insufficient amount of material in the mold, the piston may "bottom out". That is, the flange of the piston may impact the housing at full velocity. The machine must also be designed to withstand such impacts.

In order to efficiently engineer the machine, it is necessary to understand the effects of impact on design features. At present, these effects are verified using a transient finite-element analysis (FEA).

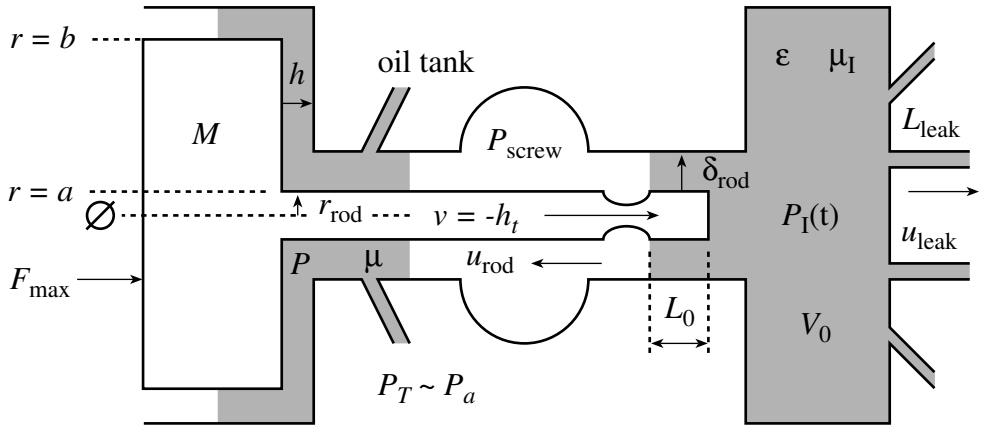


Figure 3.1: Schematic diagram of injection molding machine. See Section 3.4 for definitions of all variables and parameters.

However, an FEA is not only time consuming but the company's FEA resources are limited. Therefore, a simplified model, that could be used by a designer to obtain a first pass type of analysis, is desired. Once appropriate design features are obtained using the simplified model, they can be verified using an FEA.

Husky was interested in finding a simplified model of the forces involved in the impact of the screw on the molten metal, and the impact of the piston flange on the injection housing (i.e., when the piston "bottomed out"). For the impact of the screw, information is desired regarding the effect of the molten metal's bulk modulus (alternatively, its compressibility), and regarding the effect of the leakage flow rate of the molten metal past the impacting screw tip created by clearances between the screw and the housing (see $u_{\rm rod}$ in Figure 3.1). Currently, the company's analysis of the "bottom out" problem assumes a dry contact between the piston flange and injection housing, when, in fact, there is a thin film of hydraulic fluid between the impacting bodies that will significantly reduce the strain on the system. Information regarding the extent of this reduction is desired.

The company also wished to include the effects of material damping in the simplified models. However, we decided that an investigation of this effect would best be approached in an FEA. Therefore, given the limited time of the workshop, it was decided that we would postpone this investigation.

The model currently used by Husky ignores all effects except the deformation of the piston and housing, which is assumed to be elastic. The impact is modelled using a mass-spring system, where the force of impact is obtained for a given impact velocity, system mass and pre-calculated "spring" constants for the impacting bodies. In order to determine the importance of the presence of the hydraulic fluid, the compressibility of the molten metal, and the leakage past the screw tip, we will assume that the deformation of the piston and housing can be ignored. However, in order to make quantitative predictions of the impact forces, it is necessary to include the effects due to the compressibility of the machine parts. Thus, we also investigate a model that, like the current Husky model, uses a mass-spring system to incorporate the elastic deformation of the piston and the housing.

The goal is to derive simplified equations of motion for the piston that can be used to generate pressure profiles that occur during impact. The pressure profiles can be used to determine if the design features are sufficient to ensure that the impact is not likely to cause damage to the machine. In

the next section, we consider the "bottom out" problem. In particular, we discuss the effect of the hydraulic fluid on the impact. In section 3.3, we consider the impact of the screw on the molten metal. In particular, the effects of the leakage of the molten metal past the screw tip and the effect of the compressibility of the molten metal are discussed. In section 3.4 results from a model that ignores the effects due to the deformation of the machine parts are presented. In section 3.5, we discuss preliminary results that have been obtained on a simple model that assumes elastic deformation of the piston and housing.

3.2 The "bottom out" problem

It is well-known that lubricants can support high loads in sufficiently small gaps. A lubricant will act to prevent contact between moving components and reduce stress by spreading the load. This is the reason that automobile or machine parts exhibit very little wear when adequately lubricated. The process may be modelled by the lubrication approximation to the Navier-Stokes equations.

When a two-dimensional flat indenter is pressed onto an elastic surface, a stress singularity occurs at the edges of the indenter (in practice, some plastic deformation will occur to reduce this). For this reason it is common practice to round edges and so prevent excessively high stresses. When a fluid is placed between the indenter and the elastic body the fluid forms what is known as a squeeze film. The pressure in the squeeze film is highest at the centre and reduces to the ambient pressure at the edges of contact. Obviously the stress distribution is completely different from the case when there is no fluid present, with the lubricated contact being much less likely to exhibit stress related wear.

Due to the presence of hydraulic fluid between the piston flange and the housing, this lubrication approximation may be employed in the "bottom out" problem. During the workshop, a model of the oil squeeze film between the flange and housing was developed. The process occurs in two stages, described by constant velocity and constant force. During the constant velocity phase, the squeeze film model may be used to predict the pressure distribution in the fluid. When the load (integral of the pressure) reaches the maximum operating value, the model switches to a constant load situation. An equation of motion was derived for the evolution of the film thickness during this second stage. However, the main aim of the analysis was to determine the force that is exerted by the squeeze film on the piston, and thus, a relation between the force and the film thickness is required. When the flange and housing are flat (or conform), such a relation can be found in the form of a simple analytic expression that describes the force as a function of film thickness (and the rate of change of film thickness).

3.3 The piston screw – molten metal impact

In considering the impact of the piston screw on the molten metal, it is expected that both the compressibility of the molten metal and the leakage of the molten metal past the screw tip will decrease the strain on the system.

To investigate these effects, a simplified model of the impact was derived from the equation for the conservation of mass of the molten metal. That is, we obtain an equation indicating that the rate of change of mass within the mold itself is equal to the rate at which the mass leaves the mold. The mass is written as density ρ times volume V of the mold (which includes all volume on the mold side of

the screw tip). It is assumed that the density can be written as a function of the pressure in the mold, $P_{\rm I}$, only, and that the pressure $P_{\rm I}$ is spatially constant. As is the case for the squeeze film, the process occurs in two stages, the first is a constant velocity stage, while the second is a constant load. As for the squeeze film, we assume that the stage of interest is the second, and that at the beginning of this stage the mold has been filled (no holes) and the initial velocity of the piston is the velocity that is prescribed during the first stage.

3.4 Results

The problem now reduces to a coupled system of differential equations with dependent variables h (representing both the height of the squeeze film and the position of the piston) and $P_{\rm I}$ (the pressure in the mold). The resulting equations are as follows:

$$Mh_{tt} = -\frac{6\mu h_t}{h^3} I + P_{\rm I}\pi r_{\rm rod}^2 - \frac{P_a}{2}\left(b^2 - a^2\right) - F_{\rm app} \tag{3.1}$$

$$\frac{\partial P_{\rm I}}{\partial t} = \frac{1}{V_0 + \pi r_{\rm rod}^2 h}\frac{1}{\varepsilon}\left[-\pi r_{\rm rod}^2 h_t - \pi r_{\rm rod}\delta_{\rm rod} h_t - \frac{\pi r_{\rm rod}\delta_{\rm rod}^3\left(P_{\rm I} - P_{\rm screw}\right)}{6L_0\mu_{\rm I}}\right] \tag{3.2}$$

where h is the height of the squeeze film between the flange and housing, $P_{\rm I}$ is the pressure inside the mold, $V(t) = V_0 + \pi r_{\rm rod}^2 h$ is the volume of the molten metal inside the mold, ε is the compressibility of the molten metal, μ is the viscosity of the hydraulic fluid, $\mu_{\rm I}$ is the viscosity of the molten metal, M is the mass of the piston (including the screw), $F_{\rm app}$ is the (constant) force applied to the piston during the second stage (i.e., $F_{\rm app} = F_{\rm max}$), $P_{\rm screw}$ is the pressure between the screw and the housing, P_a is the ambient pressure (pressure in the oil tank, see Figure 3.1), $r_{\rm rod}$ is the radius of the piston (the radius of the piston rod and the piston screw are assumed to be equal), $\delta_{\rm rod}$ is the width of the gap between the screw and the injection housing, L_0 is the length of the screw, and

$$I = 2\pi\left[\frac{1}{2}b^4\ln\frac{b}{a} - \frac{1}{8}\left(3b^2 - a^2\right)\left(b^2 - a^2\right)\right], \tag{3.3}$$

that is, a constant that depends on b and a, the radius of the flange (or the flange width) and the radius of the housing along the length of the piston rod, respectively. Equation (3.1) describes the change in height h of the squeeze film, and depends on two unknowns, h and the pressure in the mold $P_{\rm I}$. The various terms on the right hand side show that the deceleration of the piston depends on the current film thickness, the pressure on the end of the rod, the ambient pressure and the applied force. Equation (3.2) describes the pressure within the metal. This depends on the rod motion (which depends on the squeeze film thickness) and the leakage between the rod and housing.

We non-dimensionalise using the following scaling factors:

$$h = h_0 x_1, \qquad h_t = \frac{h_0}{t_0}x_2, \qquad P_{\rm I} - P_{\rm screw} = P_0 x_3, \tag{3.4}$$

(i.e. x_1 is the scaled height of the squeeze film, x_2 is the scaled piston velocity, and x_3 is the scaled

pressure in the mold) to obtain the equations of motion

$$x_1' = x_2 \tag{3.5}$$

$$x_2' = -\alpha\frac{x_2}{x_1^3} + \beta x_3 - f_{\text{app}} + \beta\frac{P_{\text{a}}}{P_0} \tag{3.6}$$

$$x_3' = \frac{1}{1+\delta x_1}\left[-x_2\left(1+\frac{\delta_{\text{rod}}}{r_{\text{rod}}}\right) - \gamma x_3\right] \tag{3.7}$$

where t_0 is defined as h_0/v_{init}, v_{init} is the initial velocity of the piston (i.e., the velocity prescribed in the first stage of the impact). In addition,

$$P_0 = \frac{Mv_{\text{init}}^2}{h_0\pi r_{\text{rod}}^2} = \frac{1}{\varepsilon}\frac{\pi r_{\text{rod}}^2 h_0}{V_0}, \tag{3.8}$$

where the first equality describes the pressure required to stop the moving piston in a distance h_0, while the second equality describes the pressure induced by compressing the molten metal a distance h_0 (equating these can produce an expression for h_0). The system is controlled by six dimensionless parameters, the aspect ratio $\delta_{\text{rod}}/r_{\text{rod}}$, (i.e. the ratio of the gap width to the radius of the piston rod), and five others α, β, γ, δ, and f_{app}, which are related to the physical parameters as follows:

$$\alpha = \frac{6I\mu v_{\text{init}}/h_0^2}{Mv_{\text{init}}^2} \tag{3.9}$$

i.e. α is the ratio of the energy dissipated in the squeeze film to the initial energy,

$$\beta = \frac{P_0 h_0 \pi r_{rod}^2}{Mv_{\text{init}}^2} = \frac{\frac{\Delta V}{\varepsilon}\frac{\Delta V}{V_0}}{Mv_{\text{init}}^2} \tag{3.10}$$

(where $\Delta V = \pi r_{\text{rod}}^2 h_0$ is the change in the volume V of molten metal in the mold corresponding to a change in piston position h of h_0), i.e. β is the ratio of the energy required to compress the molten metal a distance h_0 to the initial energy, which by definition implies that $\beta = 1$,

$$\gamma = \frac{P_0 t_0 \delta_{\text{rod}}^3}{6r_{\text{rod}}h_0 L_0 \mu_{\text{I}}} = \frac{2\pi\delta_{\text{rod}} r_{\text{rod}}\frac{P_0\delta_{\text{rod}}^2}{12\mu_{\text{I}}L_0}}{\pi r_{\text{rod}} v_{\text{init}}} \tag{3.11}$$

i.e. γ is the ratio of the leakage flow rate at maximum compression to the initial flow rate,

$$\delta = \frac{h_0\pi r_{\text{rod}}^2}{V_0} \ll 1 \tag{3.12}$$

i.e. δ is the ratio of volume of compression of molten metal to volume of the mold, and

$$f_{\text{app}} = \frac{t_0^2 F_{\text{app}}}{h_0 M} \tag{3.13}$$

is the non-dimensionalised applied (constant) force. Because δ is very small, it may be neglected and therefore, there are only five parameters that control the motion of the piston. Thus, working with the

non-dimensional equations reduces the number of parameters that need be considered. For example, from the definition of the dimensionless parameter α it can be seen that increasing the viscosity of the hydraulic fluid μ will have the same effect as increasing I by increasing the flange width b. Alternatively, to investigate the effect of changing a physical quantity, we need only vary the dimensionless parameters accordingly. For example, we can increase the dimensionless parameter α if we wish to determine the effect of increasing the flange width b.

We now carry out some preliminary numerical calculations to demonstrate the various possible types of behaviour that can be observed. For example, we expect that if there is little leakage of molten metal past the screw tip, then the piston will be stopped by the molten metal. However, if there is significant leakage, then it will be the squeeze film that acts to stop the motion. In the first example, we choose the gap width between the screw and the injection housing $\delta_{\rm rod} = 50\mu$m, which is a reasonably large gap through which the molten metal may pass. The results of the numerical calculations, shown in Figure 3.2, indicate that the load of the impact is shared by both the squeeze film and the molten metal (although a large part of the load is taken by the squeeze film; see below). The squeeze film height h initially decreases rapidly and then slowly tends to a constant (non-zero) value. The velocity of the piston has a corresponding initial stage when it changes slowly, and then a stage when it rapidly tends to zero. The scaled pressure starts at a low value and then increases rapidly as impact is approached. As the piston velocity decreases, the pressure decreases until it reaches a low value sufficient to balance the applied force.

As the gap width $\delta_{\rm rod}$ decreases, there is less leakage and more of the load is taken by impact of the screw on the molten metal. Calculations for the case when the gap width $\delta_{\rm rod}$ is 25μm are shown in Figure 3.3. As in the previous case, the squeeze film height h initially decreases rapidly, however, the relaxation of the film height h to the constant value is slower, and the deceleration of the piston is not as sharp. Of particular interest is that an increase in the maximum of $P_{\rm I}$ (the pressure in the mold) is observed. For the calculations, the pressure parameter P_0, given in (3.8), is taken to be 3.13×10^8Pa, which implies that for the case of the gap width $\delta_{\rm rod} = 50\mu$m there is a maximum pressure of approximately 6×10^7Pa, while for the case of the gap width $\delta_{\rm rod} = 25\mu$m there is a maximum pressure of approximately 4 times this (2.4×10^8Pa).

If the gap width $\delta_{\rm rod}$ is taken to be as small as 10μm (very little leakage), there seems to be a qualitative change in the solution; see Figure 3.4. In this case, the initial kinetic energy of the piston goes almost fully into compressing the molten metal, i.e. the squeeze film takes little of the burden. This compression is assumed to be elastic, and thus, because there is very little leakage, the piston rebounds, and oscillates. Note that there is a very slight rebound in the case of the gap width $\delta_{\rm rod} = 25\mu$m, so this effect is already being seen here.

The definition (3.9) of the dimensionless parameter α indicates that as α increases, the effects of the squeeze film become more important. In particular, in the case when both the dimensionless parameter α and the gap width $\delta_{\rm rod}$ are large, the full impact of the piston is absorbed by the squeeze film, i.e., this corresponds to the situation when there is no molten metal in the mold. The parameter α can be increased by increasing the value of I (see equation 3.3) which in turn can be increased by increasing the flange width b. The results of increasing the dimensionless parameter α can be seen in Figure 3.5. Qualitatively, the behaviour is similar to the situation shown in Figure 3.2, except that there is a sharper deceleration of the piston on impact.

Figure 3.6 contains a plot comparing the pressure in the mold $P_{\rm I}$ to the spatial maximum pressure in the squeeze film as a function of time for the case when the gap width $\delta_{\rm rod} = 50\mu$m, and the flange

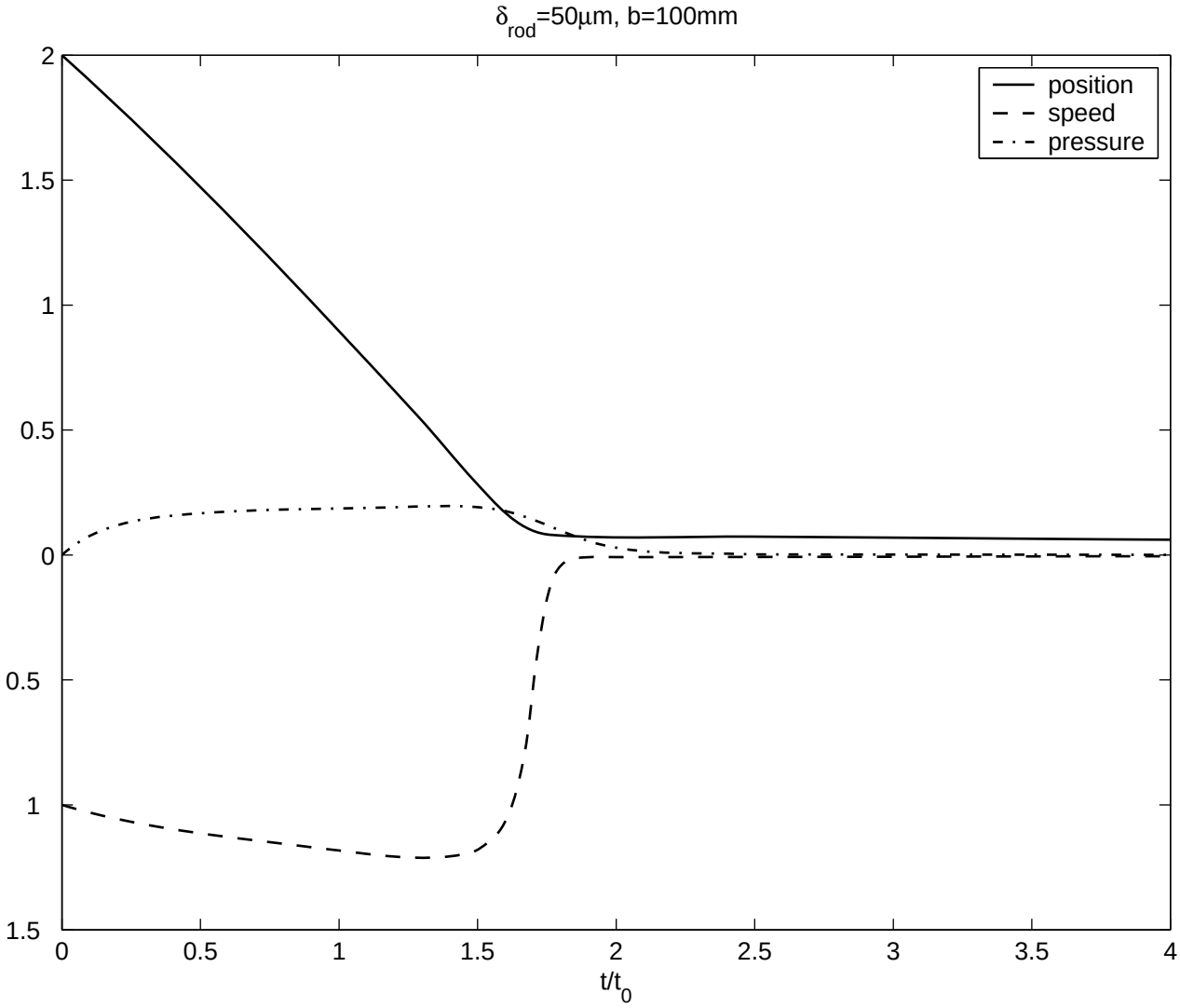


Figure 3.2: Results of simulation with moderate leakage: the gap width $\delta_{\rm rod} = 50\mu$m. Position, speed and pressure refer to the non-dimensionalised variables, x_1, x_2 and x_3, respectively. These related to the dimensional quantities via (3.4).

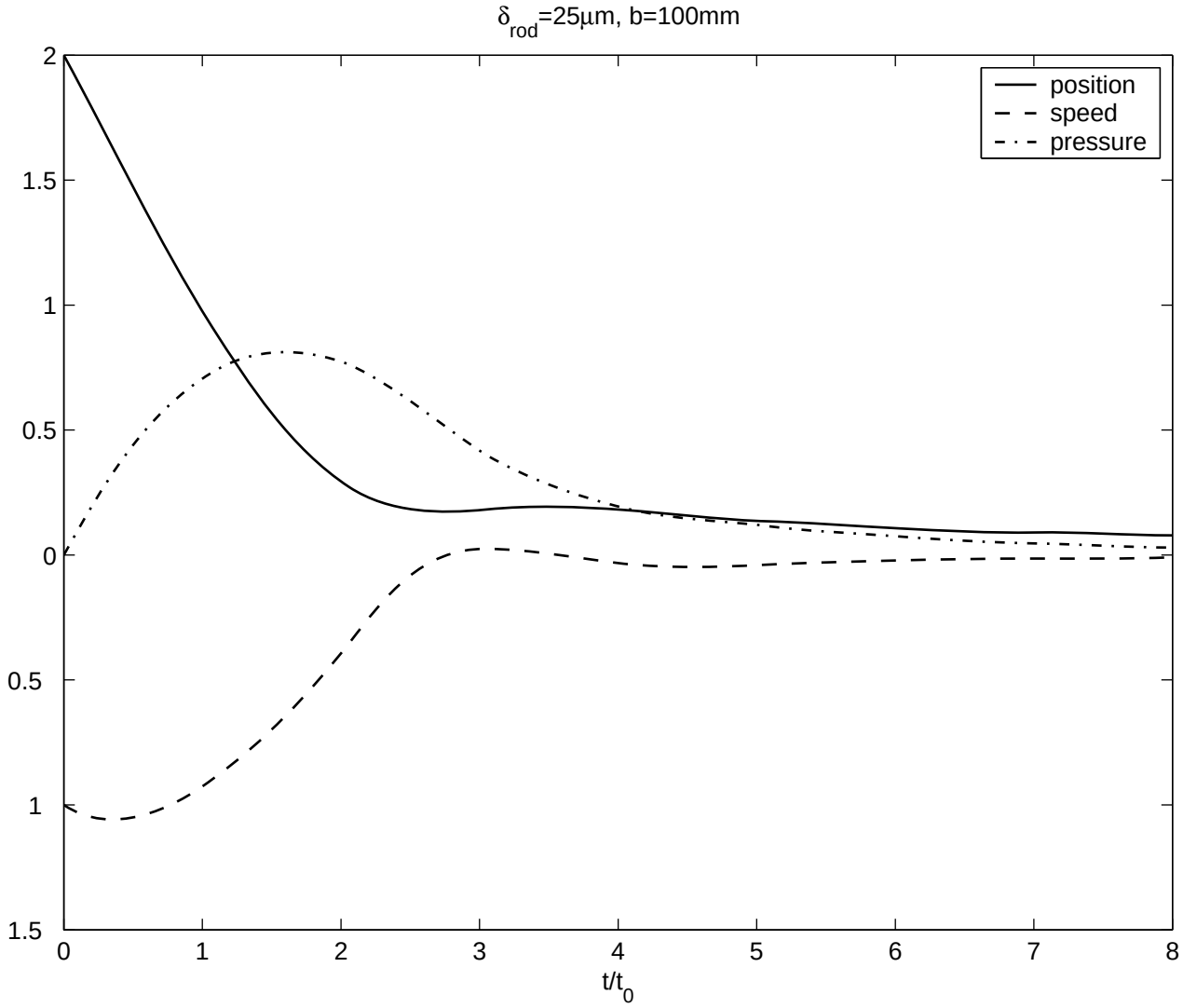


Figure 3.3: Results of simulation with small leakage: the gap width $\delta_{\rm rod} = 25\mu$m. Position, speed and pressure refer to the non-dimensionalised variables, x_1, x_2 and x_3, respectively. These related to the dimensional quantities via (3.4).

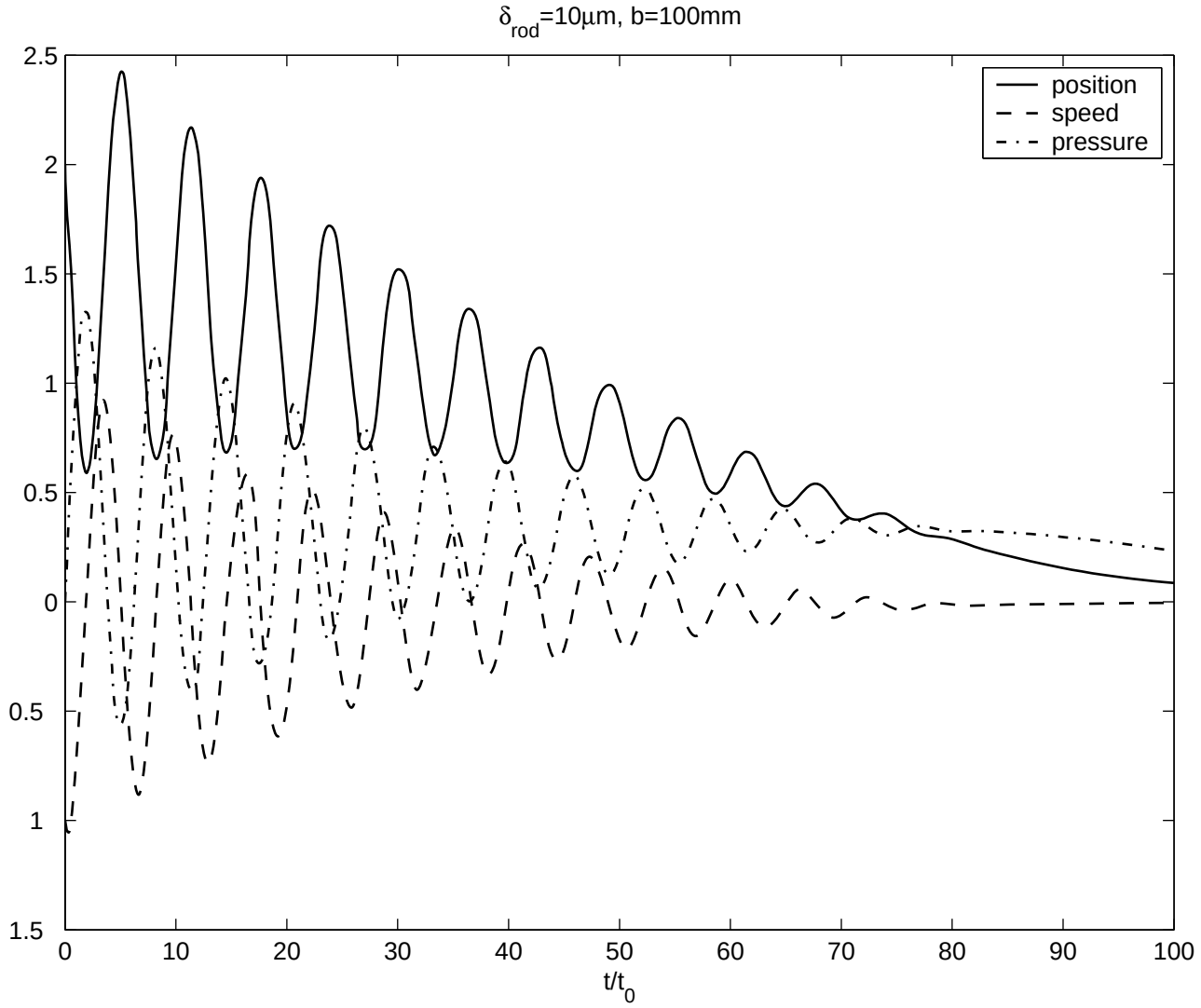


Figure 3.4: Results of simulation with extremely small leakage: the gap width $\delta_{\rm rod} = 10\mu$m. Position, speed and pressure refer to the non-dimensionalised variables, x_1, x_2 and x_3, respectively. These related to the dimensional quantities via (3.4).

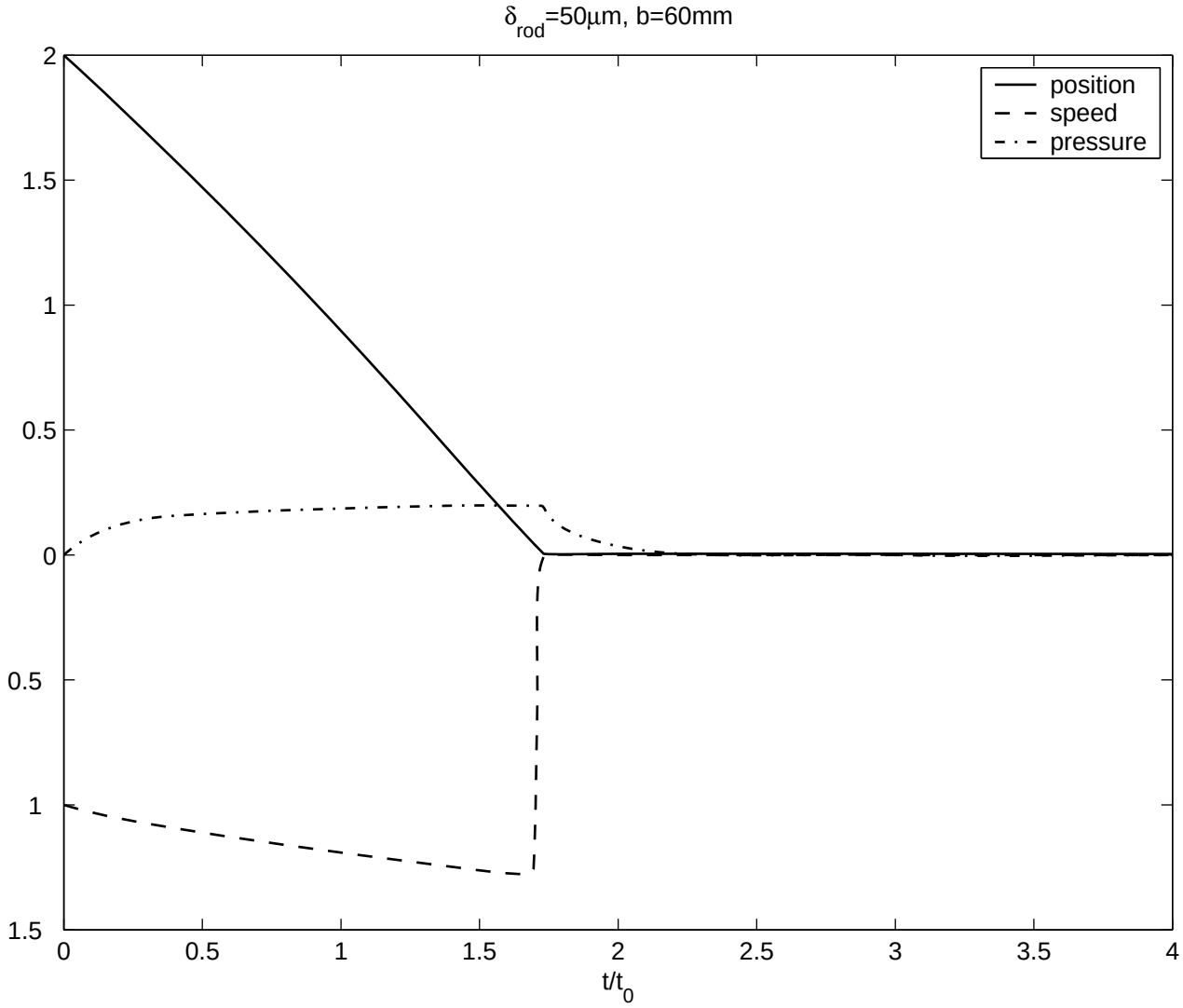


Figure 3.5: Results of simulation with moderate leakage: the gap width $\delta_{\rm rod} = 50\mu$m, and the flange width $b = 60$mm. Position, speed and pressure refer to the non-dimensional variables, x_1, x_2 and x_3, respectively. These related to the dimensional quantities via (3.4).

width $b = 100$mm. As mentioned above the maximum of P_{I}, the pressure in the mold, in dimensional units is approximately 6×10^7Pa. The maximum pressure in the squeeze film is nearly 10 times greater. This pressure is felt over only a small area near the outside of the flange; the pressure decreases rapidly toward the inner part of the flange. To illustrate this, the pressure in the squeeze film, as both a function of time and distance along the flange, is plotted in Figure 3.7.

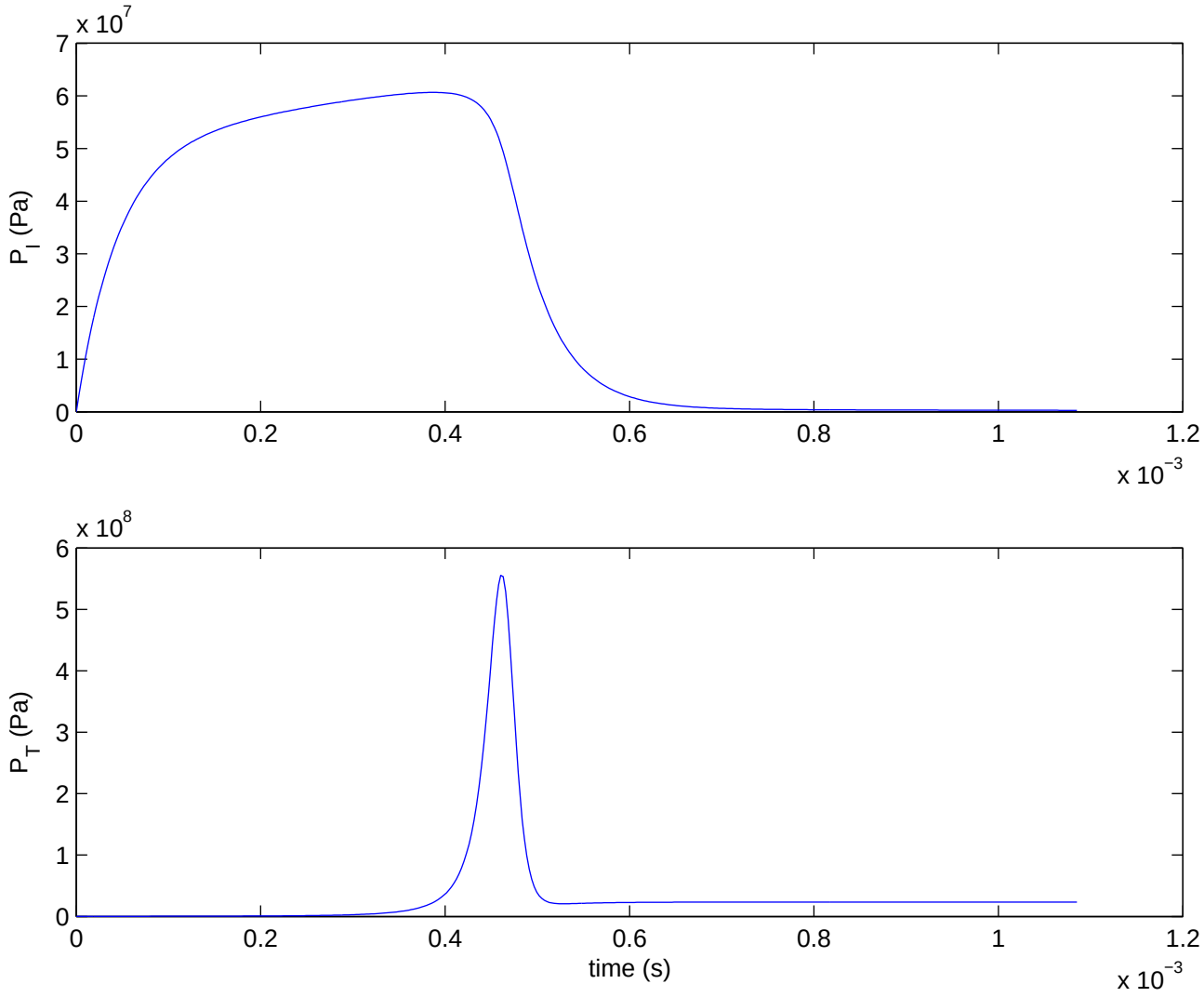


Figure 3.6: Pressure in mold P_{I} and pressure (spatial maximum) in squeeze film P_{T} vs. time for the parameter values the gap width $\delta_{\mathrm{rod}} = 50\mu$m, and $b = 100$mm.

3.5 Effects of compressibility of piston and housing

The results of the previous section indicate that depending on the circumstances, either the squeeze film or the molten metal can act to stop the piston, and therefore both should be included in any math-

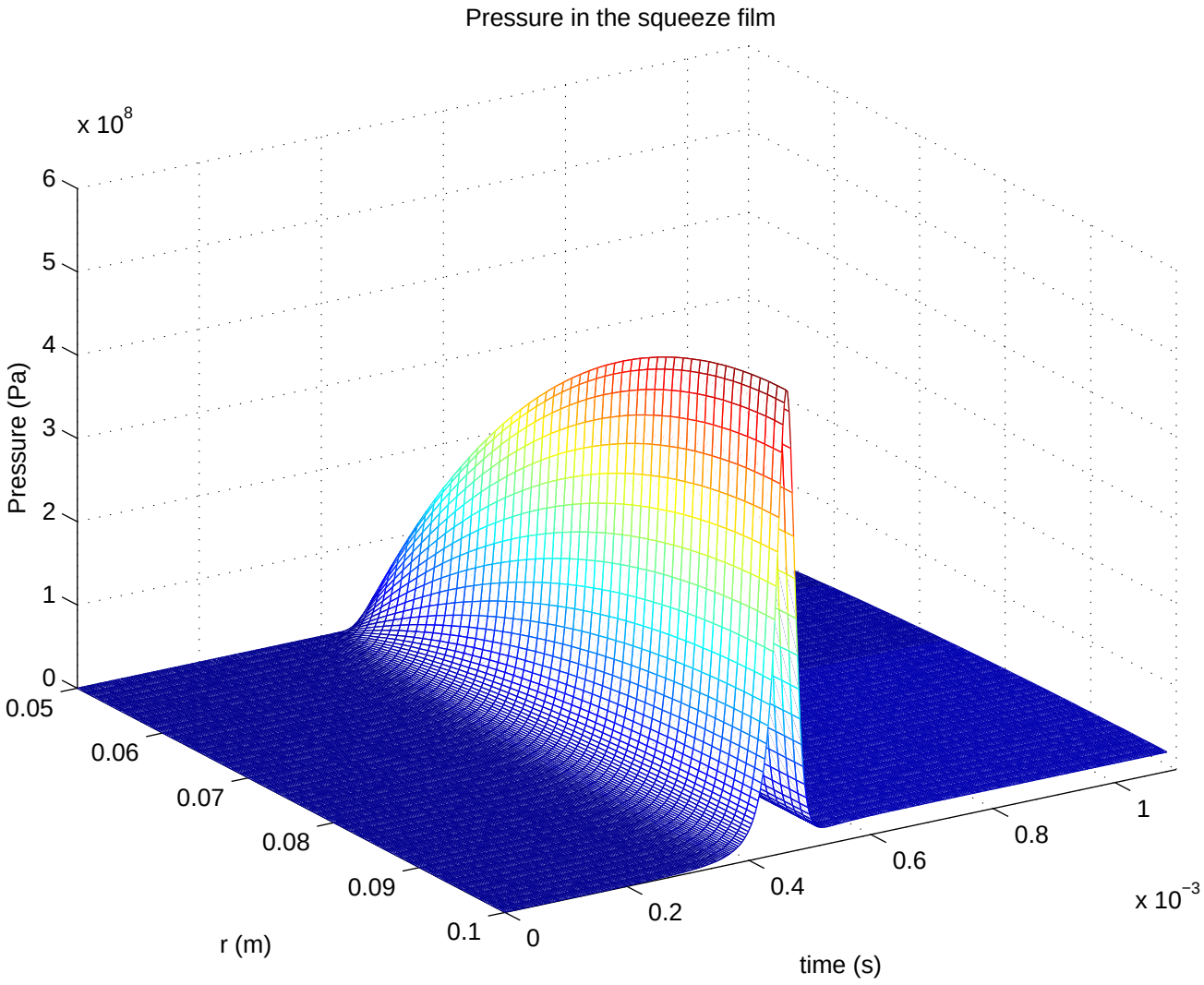


Figure 3.7: Pressure in squeeze film vs. time and radius.

ematical model of the process. The actual numerical values of pressure that were obtained, however, cannot be taken as the values that would occur in practice. At the pressures that are predicted, the compressibility of the piston and housing would likely have an effect.

Therefore, a second model was developed to investigate the effects of the elasticity of the machine parts on the impact of the piston flange on the housing. We consider a one dimensional model that is represented in Figure 3.8. The piston is modelled as two bodies of mass M_1 and M_2 coupled together by a spring with spring constant k_1 and uncompressed length l_1, and the housing is modelled by another body of mass M_3 that is attached to an immovable body (wall) by a spring with spring constant k_2 and uncompressed length l_2. The piston moves toward the housing, with the second body M_2 impacting the housing. It is assumed that hydraulic fluid is present between the two impacting bodies, and thus, a squeeze film is created during impact. As pressure is generated in the squeeze film, the springs associated with the piston and housing begin to compress. All interactions are assumed to be elastic.

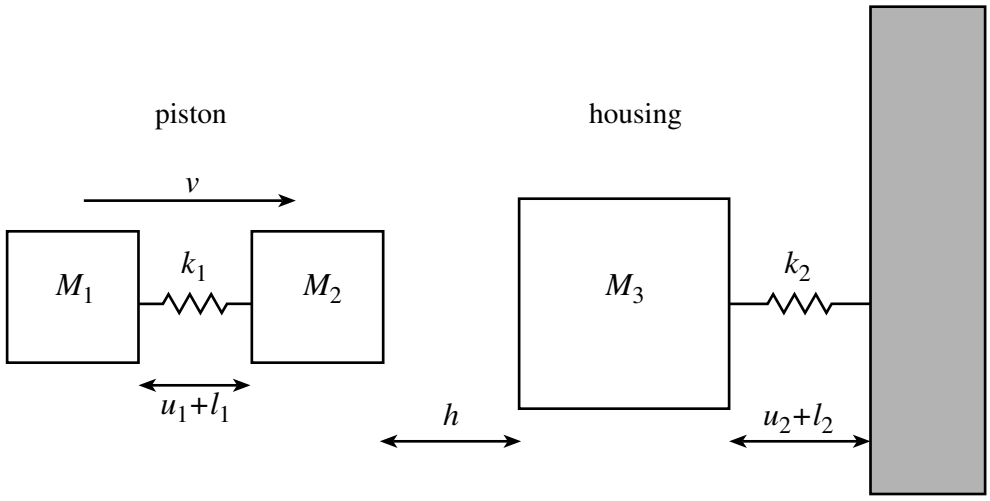


Figure 3.8: Model system of impact of piston and housing including both the effects of the squeeze film and the effects of the elasticity of the machine parts.

The model consists of the equations of motion for the three bodies where the spring forces, the force exerted by the squeeze film and the applied force are included. The equations describe the evolution of the three variables: h the distance between the piston and housing, u_1 the distance that the piston spring is compressed (that is, $u_1 + l_1$ is the distance between M_1 and M_2), and u_2 the distance that the spring of the housing is compressed. See Figure 3.8.

As in the previous section, we carry out some preliminary calculations. The case when $k_1 = 2 \times 10^9$kg/sec^2 and $k_2 = 7.2 \times 10^9$kg/sec^2 is plotted in Figures 3.9. Other relevant parameters are taken to be the same as the case plotted in Figure 3.2. Initially, the housing is stationary and the spring of the piston is not compressed. As the piston approaches the housing, the effects of the squeeze film begin to be felt and the spring of the piston begins to be compressed. As the piston impacts the housing, the squeeze film pressure spikes, with corresponding jumps in the rates of change of the variables. After impact, the piston and housing essentially become stuck together (very little variation of h, the distance between the piston and housing, from zero), and there is an oscillation of all other variables. Of particular interest is the peak pressure that is observed in the squeeze film. In this case,

the maximum pressure is less than 2×10^7Pa, while in the situation where the elasticity of the machine parts is not considered, the maximum pressure is more than 5×10^8Pa. That is, the elasticity reduces impact pressure by more than an order of magnitude. We have not yet investigated how this result changes as parameters are varied. This will be discussed further in the final report.

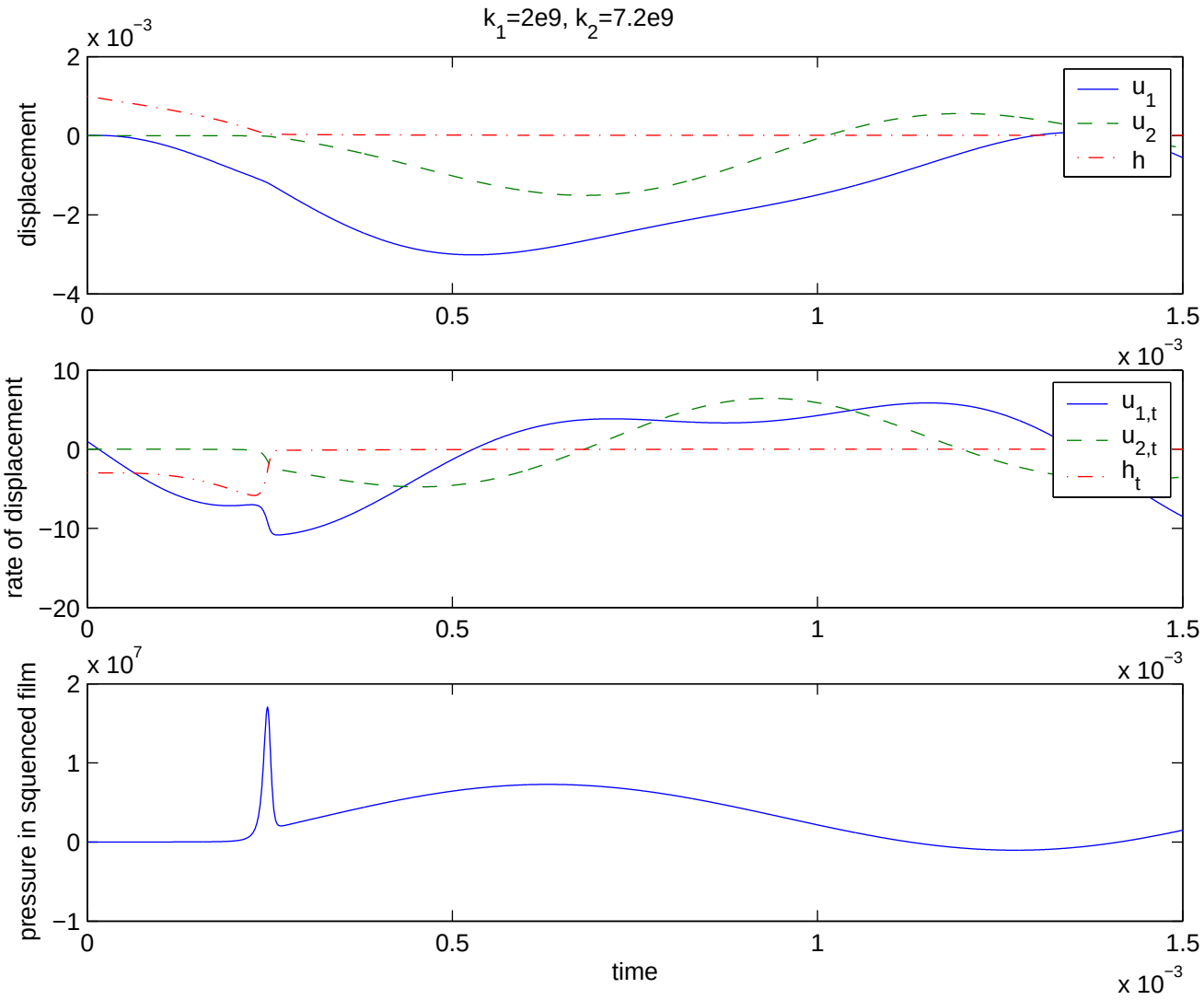


Figure 3.9: Results from model of impact of piston and housing including both the effects of the squeeze film and the effects of the elasticity of the machine parts. $k_1 = 2 \times 10^9$kg/sec^2 and $k_2 = 7.2 \times 10^9$kg/sec^2, and other constants are taken to be those for the case plotted in Figure 3.2.

3.6 Concluding Remarks

The results indicate that the presence of the hydraulic fluid, the compressibility of the machine parts and molten metal, and the leakage of molten metal past the screw tip can all produce significant effects

on the pressure profiles and therefore should all be included in any mathematical model of the process.

In the final report, the models and results that are briefly presented here will be described in detail. The discussion will include the derivation of the forces generated in the squeeze film and the mold, and will also include more details of the model that assumes elasticity of the machine parts. Any improvements in the models that are made will also be reported. In particular, an improved (continuous) model that incorporates all the effects discussed here, including the compressibility of the machine parts, is currently being developed.

Chapter 4

Seismic Attenuation Problem

Problem Presented By: Kenneth J. Hedlin (Husky Energy); Gary Margrave (University of Calgary)

Mentors: Nilima Nigam (McGill University); Tobias Schaefer (University of North Carolina at Chapel Hill)

Student Participants: Mohammad al-Khaleel (McGill); Linping Dong and Carlos Montana (University of Calgary); Wan Chen, Catherine Dupuis, Gilles Hennenfend, Felix Hermann and Peyman Poor Moghaddam (University of British Columbia); Heejeong Lee and Jinwoo Lee (Seoul National University); Joohee Lee (University of North Carolina at Chapel Hill); Namyong Lee (Minnesota State University); Yan Wu (University of Manitoba)

Report prepared by: Nilima Nigam (nigam@math.mcgill.ca)

4.1 Introduction

Seismic imaging, a technique in which the reflections of a source seismic wave are recorded as it passes through the earth, is a major tool for geophysical exploration. Seismic imaging can be used to reconstruct a profile of the material properties of the earth below the surface, and is thus widely used for locating hydrocarbons.

The problem presented by Husky Energy concerns seismic attenuation: the loss of energy as a seismic wave propagates through the earth. As an exploration tool, attenuation effects have only recently attracted attention. These effects can prove useful in two ways: as a means of correcting seismic data to enhance resolution of standard imaging techniques, and as a direct hydrocarbon indicator. Theoretically, a subsurface reservoir full of hydrocarbons will tend to be acoustically softer than a porous rock filled only with water; Kumar et al show that attenuation is highest in a partially fluid-saturated rock.

Many physical processes can lead to the attenuation of a seismic trace. In the present work, we ignore attenuation effects such as spherical divergence or scattering, and concentrate on intrinsic attenuation effects exclusively. The latter are caused by friction, particularly in porous rocks between fluid and solid particles, see [2, 7].

The goal of the workshop was to find a means of computing seismic attenuation from relatively short windows of seismic imaging data, and particularly be able to identify regions of anomalous attenuation.

The paper is organised as follows. We begin by a detailed description of the attenuation problem in Section 4.2, collecting important notation and assumptions for easy reference. The reader may find the simple numerical example presented in Section 4.2.3 useful to fix ideas.

In Section 4.3, we consider the use of *frequency-shift techniques* to identify anomalous attenuation; two different attributes are used on simulated and real data. In Section 4.4, different wavelet-based denoising techniques are used to identify the attenuation anomaly. In Section 4.5, we present the mathematical ideas behind an extension of a Wiener technique. We end the paper in Section 4.6 with ideas for future work, including the use of a constrained optimization problem for estimating the attenuation.

4.2 Problem Statement

The ability of a material to attenuate seismic waves is measured by a dimensionless quantity Q, called the *attenuation factor*, by

$$Q := \frac{\text{energy of seismic wave}}{\text{energy dissipated per cycle of wave}} = \frac{2\pi E}{\triangle E}$$

where E is the energy of the wave, and $\triangle E$ is the change in energy per cycle. Typical values of Q range from 5-20 (dirt) through 100 (rock) to 10,000 (steel). In what follows, we assume that this attenuation factor is independent of frequency ω in the useful seismic bandwidth. The attenuation of the wave is directly linked to the different layers that compose the Earth, so that whenever changes in the composition of layers occur, the attenuation changes too. This is why we would like to be able to detect changes in attenuation, as it would enable us to identify s change material properties.

Goal: *To estimate Q from given seismic data.*

In this paper, we restrict our attention to 1-D models, in which all geological layers are horizontal. In addition, we assume the source and receiver are effectively coincident. We assume that the receiver is positioned at the surface of the earth (in other words, we do not consider vertical seismic profiles).

For a medium with linear stress-strain relation, it is known that wave amplitude A is proportional to $\sqrt{E}$. Hence,

$$\frac{1}{Q} = -\frac{\Delta A}{\pi A} \tag{4.1}$$

from which we can obtain the amplitude fluctuations due to attenuation. That is, given initial amplitude A_0, let λ be the wave length given in terms of frequency ω and phase velocity c by $\lambda = 2\pi c/\omega$, then $\Delta A = \lambda(dA/dz)$. Hence, equation (4.1) becomes,

$$\frac{dA}{dz} = -\frac{\omega}{2cQ}A \tag{4.2}$$

with the exponential decaying solution

$$A(\omega, z) = A_0(\omega) \exp\left(-\frac{\omega z}{2cQ}\right). \tag{4.3}$$

Now, from observation of exponentially decaying values of $A(\omega, z)$, we can compute Q value. That is, from (4.3), we have

$$\ln\left(\frac{A}{A_0}\right) = -\omega\left(\frac{z}{2cQ}\right) = -\omega\left(\frac{t}{2Q}\right) \tag{4.4}$$

Here we assume that the phase velocity c does not depend on frequency, ie, that there are no dispersion effects. This has the added effect of correlating well the time of travel of the reflected wave with the depth of the layer from which the reflection occurs.

Hence, by recording the $\ln(A/A_0)$ versus ω graph, and then estimate the *average* slope, we can recover the value of Q. This idea is known as *log spectral ratio* method.

Another model, which includes the effect of reflections of the signal from various layers, is as follows. Let $s_k(t)$ denote the seismic trace obtained from layer k, received at time t_k at the receiver (assuming the source signal s_0 was emitted at time $t = 0$). Suppose the coefficient of reflection at the kth layer is r_k, and the the source signal is s_o. In general, r_k is unknown, as is s_o. Indeed, the source signal is usually generated by a denotation; characterizing this signal is not possible.

In the Fourier domain at frequency ω, we may write

$$|\hat{s}_k(\omega)| = r_k|\hat{s}_o(\omega)|e^{-p\omega t/Q} \tag{4.5}$$

where p is a constant, and r_k is the coefficient of reflectivity of the kth layer.

Suppose we have similar information about a seismic trace reflected from layer j, then the log spectral ratio method estimates the attenuation Q as:

$$\log\left(\frac{|\hat{s_k}|}{|\hat{s_j}|}\right) = \log|r_k| - \log|r_j| + \frac{p\omega}{Q}(t_j - t_k).$$

4.2.1 A Convolutional Model of Attenuation

Let $s(t)$ denote an unattenuated seismic trace received at time t at a receiver. If $w(t)$ was the source waveform and $r(t)$ is the reflectivity as a function of depth (equivalently time), then we may write

$$s(t) = w * r := \int_{-\infty}^{\infty} w(\tau)r(t-\tau)\,d\tau. \tag{4.6}$$

Two key assumptions are made regarding the source signal and the reflectivity:

- *the white reflectivity assumption*. The white reflectivity assumption simply means the reflectivity r satisfies

  $$\int_{-\infty}^{\infty} \tilde{r}(s)r(t-s)\,ds = \delta(t)$$

  where $\delta(t)$ is the Dirac measure.

- *the minimum phase assumption*: this means that the source signal $w(t)$ is causal, invertible, and possesses minimum phase in the sense that if we write the signal in the frequency domain

$$\hat{w}(\omega) = A_0(\omega)e^{i\phi(\omega)}$$

we can find the phase $\phi(\omega)$ by using a Hilbert transform.

With these assumptions, and in the absence of attenuation, we are able to recover the source signal w from a given trace s using the Wiener process on equation (4.6):

$$s * \tilde{s} = (w * r) * (\tilde{w} * \tilde{r}) = (w * \tilde{w}) * \delta.$$

Taking a Fourier transform of the above expression, we get the amplitude $A(\omega)$ of the source signal w; the minimum phase assumption now allows us to recover the phase.

Unfortunately, the Wiener process does not apply in quite such a straight-forward fashion to the case where the signal is attenuated. The process of attenuation is described by the action of a pseudo differential operator: the attenuated trace s_a is now

$$s_a(t) := \int_{-\infty}^{\infty} w_\alpha(\tau, t-\tau)r(\tau)d\tau \tag{4.7}$$

where

$$w_\alpha(u;v) := \int_{-\infty}^{\infty} \alpha(u,\eta)e^{i\eta v}\hat{w}(\eta)\, d\eta, \tag{4.8}$$

$$\alpha(u,\eta) = \exp\left(-\frac{\eta u}{2Q}\right)\exp\left(\frac{iu}{2Q}\int_{-\infty}^{\infty}\frac{e}{\eta - e}\, de\right). \tag{4.9}$$

In Section 4.5 we describe an extension of the Wiener technique to the case of attenuated waves.

We can also describe, in the setting of this convolutional model, the windowed log spectral ratio technique which is commonly used. Let Ω_1 and Ω_2 be two intervals of time over which the seismic trace s_a has been sampled. We expect that

$$\hat{s_a}(\Omega_1) \approx \hat{w}_{\text{eff}_1} \cdot \hat{r}_{\text{eff}_1}$$

where w_{eff_1} is the effective signal over window Ω_1, and r_{eff_1} is the effective reflectivity and the hats denote the Fourier transform.

We also expect that

$$|\hat{w}_{\text{eff}_2}| = |\hat{w}_{\text{eff}_1}|\exp(-\frac{\omega\Delta\tau}{2Q})$$

where $\Delta\tau$ is the window separation. Therefore, the log spectral ratio is

$$\log\left(\frac{|\hat{s}_a(\Omega_1)|}{|\hat{s}_a(\Omega_2)|}\right) = \frac{\omega\Delta\tau}{2Q} + \log\left(\frac{|\hat{r}_{\text{eff}_1}|}{|\hat{r}_{\text{eff}_2}|}\right).$$

4.2.2 A Discrete Model

In practice, seismic trace data is sampled at discrete time intervals, for a finite duration of time. We therefore describe a discrete version of the convolutional model above: suppose we know the initial source signal, as well as the attenuation and reflectivity properties of the medium being sampled. Let the data be sampled at times $t_1, t_2, \ldots t_n$. From this, we can construct a matrix W_α, and a vector of reflectivities $\boldsymbol{r} = (r_1, r_2, \ldots r_n)^T$, where r_i is the reflectivity of the layer at depth ct_i. Then, the discrete version of equation (4.7) is

$$W_\alpha \boldsymbol{r} = (\boldsymbol{w_1}|\boldsymbol{w_2}|\ldots|\boldsymbol{w_n})\,\boldsymbol{r} = \boldsymbol{s} := (s_1, s_2, \ldots s_n)^T.$$

The entries w_{ij} of matrix W_α have the following properties:

- If $t_i > t_j$, $w_{ji} = 0$ (causality assumption)
- If $t_i < t_j$, $w_{ji} = w_\alpha(t_i, t_j - t_i)$, where w_α was defined by equations (4.8) and (4.9).

Therefore, W_α is lower triangular, and the amplitude spectra of column vectors $\boldsymbol{w}_i$ attenuate by an exponential factor from left to right.

The *forward seismic problem* is: given W_α, $\boldsymbol{r}$, find the seismic trace vector $\boldsymbol{s}$.

The *inverse seismic problem* is: given $\boldsymbol{s}$, find W_α, $\boldsymbol{r}$. In our specific case, we have to find W_α, specifically the amount of attenuation between the amplitude spectra of the columns of W_α. As is easy to see, the inverse problem is quite ill-posed.

4.2.3 A Simple Model

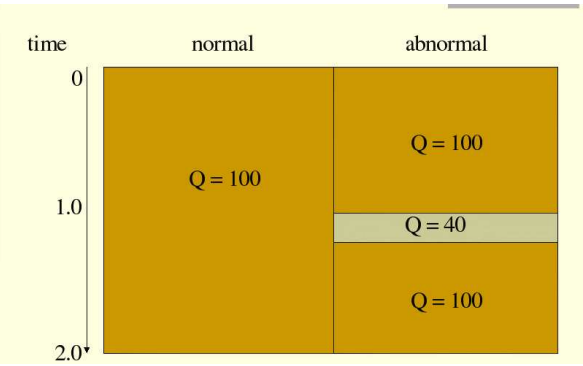
As a first approach to solving the attenuation retrieval problem, we began by creating some simulated data and solving the forward problem, where the Q-profile is known. The intermediate goal is to see how the presence of the anomaly changes a known signal.

The data is sampled for $t = 2$ s, at intervals of time 0.002 s. The seismic trace vector $\boldsymbol{s}$ thus has 1000 entries. The reflectivity vector $\boldsymbol{r}$ is drawn from a nearly-white probability distribution. The source signal w has dominant frequency 20 Hz, is generated once, and used repeatedly.

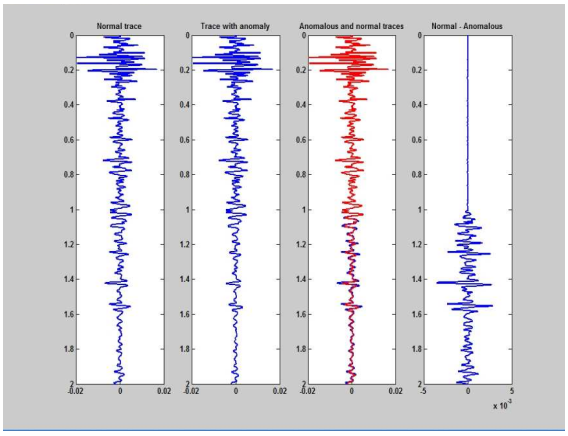
We consider the following simple 1-D models as shown in the picture Figure 4.1(a):

- The normal attenuation case, where the attenuation is a constant $Q = 100$ for all depths;
- The anomalous attenuation case, where Q changes at depth $ct = 1$ to $Q = 40$, and then changes back to $Q = 100$ at depth $ct = 1.1$ (we assume the speed of propagation has been normalized to $c = 1$).

We calculate the resultant seismic traces $s(t)$ as it propagates through the two media. In Figure 4.1(b) we show the computed seismic traces with and without the anomaly. As can be seen, the two traces appear nearly identical; only when we subtract them can we clearly spot the onset of the anomaly (at time $t = 1$).



(a) Profile of "normal" and "anomalous" attenuation.



(b) Left–right: normal trace, anomalous trace, traces superimposed, difference of traces.

Figure 4.1: The seismic trace corresponding to the attenuation anomaly is nearly identical to the normal attenuation case.

4.3 Anomaly Detection Using Moments of Frequency

With the model (4.5) of seismic attenuation that we are using, it is clear that the amplitudes of higher frequency components attenuate more over the same depth than do lower frequency components. The net effect of this phenomenon is that there is a *red-shift* in the signal as it propagates through the earth. For example, in Figure 4.3 we show the amplitude spectrum of the source signal in blue, and the post-attenuation spectrum in red. A clear shift in the mean frequency is seen. The strategy we propose in this section is to look at the amplitude spectrum of the seismic trace over many overlapping windows in time, and look at the changes in the mean frequency of these spectra. These changes should correlate well with changes in attenuation.

With a given seismic data $s(t)$, we can take a windowed Fourier transform (e.g. the Gabor transform) to see the local spectral property of the data:

$$\hat{s}_g(t,\omega) = \int_{-\infty}^{\infty} s(\tau) g(t-\tau) e^{-i\omega\tau}\, d\tau, \tag{4.10}$$

where g is a Gaussian function used as a window. We can then compute the average with respect to the frequency to obtain the centroid frequency, $f_c(t)$, by the formula:

$$f_c(t) = \frac{\int \omega |\hat{s}_g(t,\omega)|\, d\omega}{\int |\hat{s}_g(t,\omega)|\, d\omega} \approx \frac{\sum_k \omega_k |\hat{s}_g(t,\omega_k)|}{\sum_k |\hat{s}_g(t,\omega_k)|}. \tag{4.11}$$

In a similar fashion, we could also compute other moments. During the workshop, we computed the second moment, and expect an amplification of the red-shift phenomenon.

$$f_s(t) = \frac{\int \omega^2 |\hat{s}_g(t,\omega)|\, d\omega}{\int |\hat{s}_g(t,\omega)|\, d\omega} \approx \frac{\sum_k \omega_k^2 |\hat{s}_g(t,\omega_k)|}{\sum_k |\hat{s}_g(t,\omega_k)|}. \tag{4.12}$$

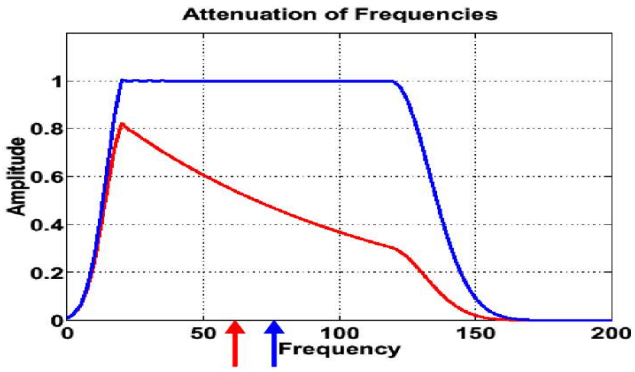


Figure 4.2: Shift in mean frequency: amplitude spectrum in red is for signal after attenuation.

Since we can expect a sudden decay of overall frequency amplitudes when sharp anomalies occur, we may observe lower values of $f_c(t)$ at the abnormality. We test this hypothesis by examining both the synthetic data from Section 4.2, as well as real data provided by Prof. Margrave. On both sets of data, we look at the trends in f_c and the second moment f_s.

4.3.1 Simulation with Synthetic Data

We tested the attributes f_c, f_s as predictors of the onset of attenuation anomalies on the synthetic data described in Section 4.2.3. In Figure 4.3 the centroidal frequency f_c of the anomalously attenuated signal is decreasing faster after $t = 1$ than that of the normal signal, acting as a good indicator of the abnormality.

In Figure 4.4 we show the variation in the second moment f_s for the seismic traces in the normal (blue) and anomalous (red) case. We also check the robustness of this attribute to noise; for very low levels of noise, the second moment is still a good predictor of the onset of the instability. As the noise increases, the amplitude spectra become too polluted in the high frequency regime to yield accurate predictions.

In real situations one does not have reference data to compare with. Thus, if we use these attributes, there should be intrinsic changes in the values of these attributes which allows us to identify anomalous behaviour.

Through many experiments we initially suggested that the graph of the centroid frequency f_c as a function of depth becomes concave near the onset of an anomaly. Unfortunately we found that while this concavity necessarily happens near abnormality, it also happens even in normal regions, which makes the criteria rather useless. Nevertheless we will see that the centroid frequency still acts as a good indicator for abnormality in the next section.

4.3.2 Simulation with Real Data

To test the ability of the centroidal frequency techniques to find attenuation anomalies in real data, we used two data sets provided by Prof. Margrave. These were seismic profiles taken at two geographical

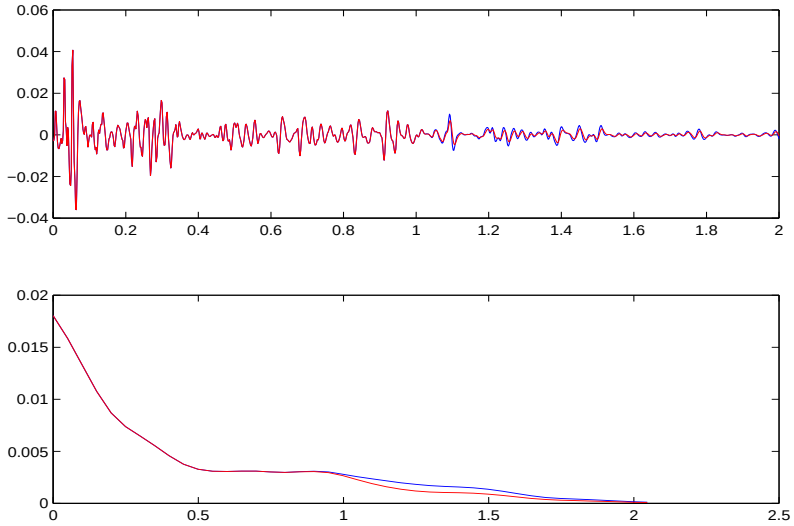


Figure 4.3: Above: Synthetic signals. Below: Centroid frequencies f_c for both traces over many windows. Note f_c values for "normal trace" are greater than the "anomalous trace".

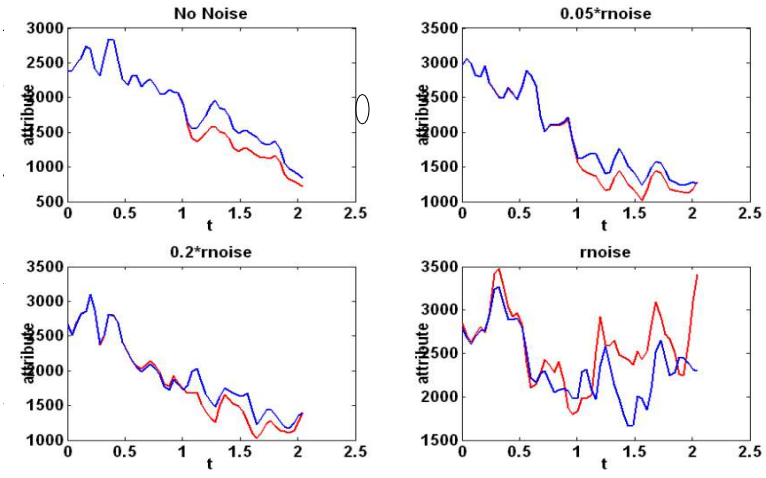
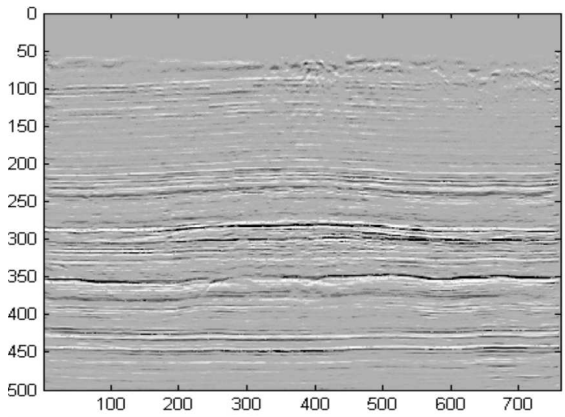
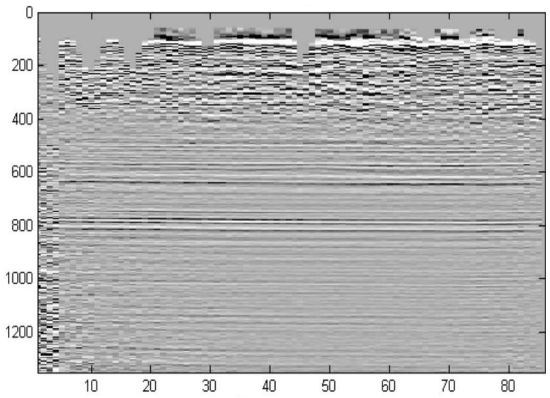


Figure 4.4: Synthetic signals and the second moment of frequency. Adding noise renders the detection of the anomaly less robust.

locations: Pike Peak and Blackfoot, respectively. In Figure 4.5 we see the *seismic cross sections* associated with the two data sets: in each, many source-receiver pairs were located along a horizontal line, and data was collected over a period of time. In these figures, the surface of the earth is on top; the y axis indicates depth (equivalently time of travel of the signal).

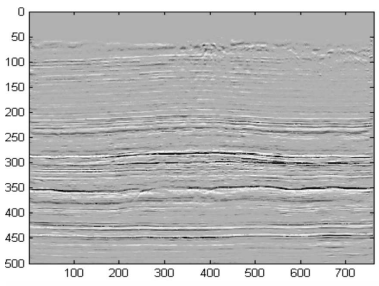


(a) Pikes Peak: anomaly at depth 350, location x=400.

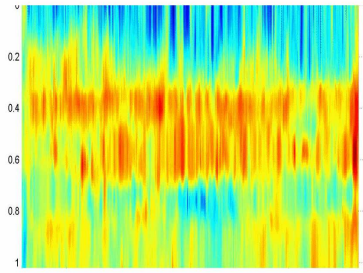


(b) Blackfoot reservoir: anomaly at depth t=600, location x=40-50.

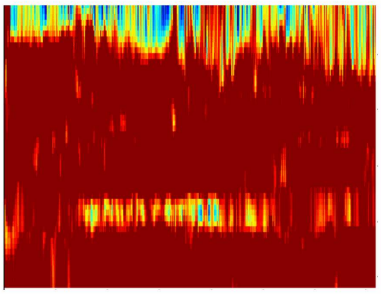
Figure 4.5: Actual seismic traces: Pikes Peak and Blackfoot.



(a) Pikes Peak: actual trace.



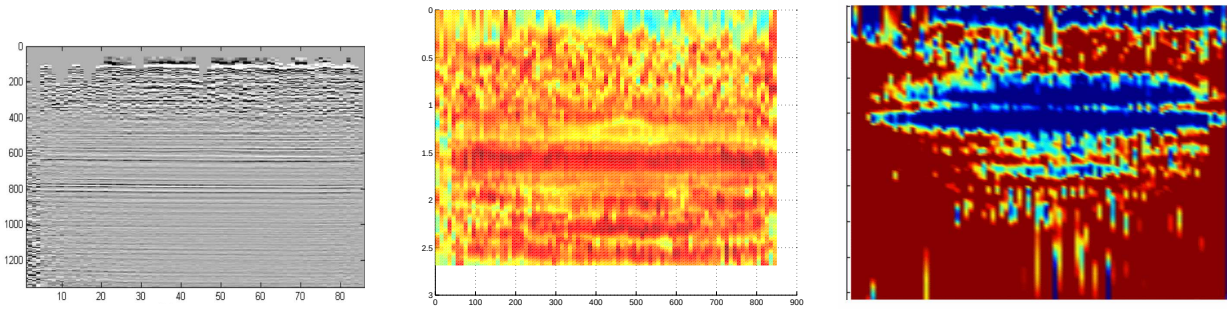
(b) Centroidal frequencies f_c.



(c) Second moments f_s.

Figure 4.6: The Pike Peaks data set: the actual section, and the attribute-derived sections. Both f_c and f_s predict the anomaly at depth 350 well.

In Figures 4.6(b) and 4.7(b), the red indicate that the centroidal frequency f_c over a given window is high, and blue indicates a low f_c. In Figure 4.6(b), one sees a region of sudden decay of f_c at depth approximately 350, and location 400. In Figure 4.7(b), one can clearly see a region in yellow (lower f_c at depth 600, location $45 - 50$, identifying the presence of an attenuation anomaly. This region corresponds well with a known gas reservoir.



(a) Blackfoot: actual trace. (b) Centroidal frequencies f_c. (c) Second moments f_s.

Figure 4.7: The Blackfoot data set: the actual section, and the attribute-derived sections. f_c captures anomaly at $t = 600$, f_s less sharply resolved.

In Figures 4.6(c) and 4.7(c), regions of large deviation from the mean f_s are highlighted in blue. Again, in Figure 4.6(c) one sees an anomaly at depth 350. The width of this region appears larger than in 4.6(b); it would be interesting to cross-check this with any known geological features there. Likewise, Figure 4.7(c) picks out a wide band of anomalous attenuation at depth 300, an a narrower, more localized band at depth 600.

4.4 Wavelet-Based Techniques

In this section, we started with the assumption that the Fourier transform of the seismic trace s is proportional to the Fourier transform of the reflectivity r, where the proportionality factor is positive and contains information on the source signal and the attenuation factor. In other words, we can say

$$\hat{s}(\omega) \propto \hat{r}(\omega) e^{-A(\omega,t)}.$$

Here A can be parametrized by Q, and has the form $A(\omega, t) = \int_0^t a(\omega, u)\, du$. By taking the Fourier transform, we turn the convolution in the time domain into a multiplication in the Fourier domain.

The basic idea is to remove the reflectivity r in order to be left only with the information on the attenuation and the source function. The deconvolution being a hard problem, we would like to rewrite the problem in a much easier way so that we could separate the reflectivity from the other information. So after taking the Fourier transform, we take logarithms to transform the multiplication into a summation:

$$\log(|\hat{s}|) = -A(\omega, t) + \log(|\hat{r}|).$$

Then we can consider the data d as a sum of a model $m = -A(\omega, t)$ (including attenuation and the source signature) and coloured noise $n = \log(|\hat{r}|)$ (reflectivity). We have now a denoising problem in the *log-Fourier domain*.

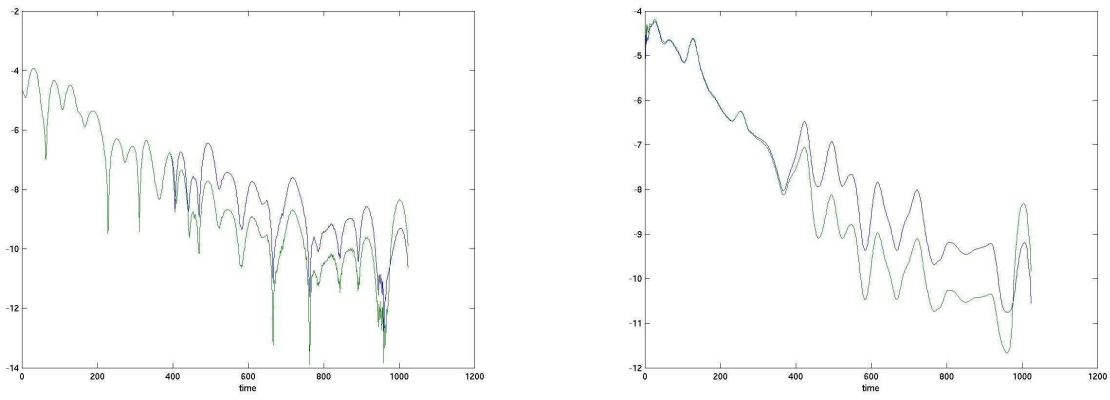
4.4.1 Denoising Using Wavelets

In order to denoise the data, we use wavelet thresholding (we used only hard thresholding). Practically, we took a Windowed Fourier transform of the seismic traces, the result of which depends on both time and frequency. Then we took the logarithm of its absolute value.

The smoothness properties are different in the time and the frequency direction, so we exploited that by taking a different wavelet transform along each direction. Along the time direction, the traces should be piecewise constant as it should be constant where the Q is constant and only change when Q changes. On the other hand, the traces should be relatively smooth along the frequency direction. It is important to choose the right wavelet to represent our signal.

So the denoising method is as follows:

1. We take a 2D wavelet transform
2. We threshold the coefficients according the noise level (prior information on the statistics of the reflectivity)
3. We take a 2D inverse wavelet transform. The result is the denoised data which should scale like the amplitude $-A(\omega, t)$.



(a) Normal (blue) and anomalous (green) traces in the *log-Fourier domain* before denoising.

(b) With thresholding.

Figure 4.8: Normal (blue) and anomalous (green) traces in the *log-Fourier domain*, before and after denoising.

We apply this technique to the synthetic data as generated in Section 4.2.3. We can see from the two figures in Figure 4.8 that the difference between the normal and anomalous traces is enhanced by the thresholding because we took most of the reflectivity contribution out which is mostly the oscillatory part.

4.4.2 Denoising Using a Minimization Technique

Recall that we are seeking to locate $-A(\omega, t)$ from noisy data d, where

$$d = \log(|\hat{s}|) = -A(\omega, t) + \log(\hat{r}) = m + n.$$

We have already discussed a wavelet-based denoising strategy above. Another possible means of removing the noise from the data is to solve a minimization problem

$$\min_m \frac{1}{2}\|C_n^{-1/2}(d - m)\|_2^2$$

where C_n is the covariance matrix of the noise $C_n = E[nn^T]$. We can use a wavelet transform W to convert the minimization problem into

$$\min_{\tilde{m}} \frac{1}{2}\|\Gamma^{-1}(\tilde{d} - \tilde{m})\|_2^2 + \lambda\|\tilde{m}\|_p, \qquad \text{where } \tilde{d} := Wd,\ \ \tilde{m} = Wm,\ \ \tilde{n} = Wn,\ \ E[\tilde{n}\tilde{n}^T] = \Gamma^2.$$

The noise thresholding used is called hard or soft, depending on whether $p = 1$ or 2 respectively. Solving the minimization problem and subsequently inverting allows us to reconstruct an approximation to m.

4.5 Extension of the Wiener Technique

The relation between the measurement $s(t)$ and the initial wave $w(t)$ can be written in form of a modified convolution, i.e.

$$s(t) = w(t) * a \odot r(t) := \int w_\alpha(\tau, t - \tau) r(\tau) d\tau. \tag{4.13}$$

The kernel w_α contains the attenuation function $Q(t)$, which we eventually would like to determine:

$$w_\alpha(u, v) = \int \hat{w}(\eta)\alpha(u, \eta)\mathrm{e}^{i\eta v} d\eta, \tag{4.14}$$

$$\alpha(u, \eta) = \exp\left(-\frac{u\,\mathrm{sgn}(\eta)\eta}{2Q(u)} + i\frac{u\,\mathrm{sgn}(\eta)\mathcal{H}(\eta)}{2Q(u)}\right). \tag{4.15}$$

Here, $\hat{w}$ denotes the Fourier transform of w and $\mathcal{H}$ the Hilbert transform. In the case without attenuation, this expression reduces to $s = w * r$. Therefore, we can write for small attenuation, corresponding to large Q, a Taylor expansion of the kernel α and use the first nontrivial terms in this expansion in order to derive higher order corrections to the relation $s = w * r$. For $\alpha(u, \eta)$ we find

$$\alpha(u, \eta) = \sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{u}{2Q(u)}\right)^k \kappa(\eta)^k \tag{4.16}$$

with the abbreviation

$$\kappa(\eta) = -\mathrm{sgn}(\eta)\eta + i\mathrm{sgn}(\eta)\mathcal{H}(\eta). \tag{4.17}$$

In the above expression, the arguments u and η are decoupled in the sense that we can perform the integration with respect to η in the expression (4.14) for w_α. Thus we obtain for w_α the expansion

$$w_\alpha(u,v) = w(v) + \frac{u}{2Q(u)}J_1(v) + \frac{1}{2}\left(\frac{u}{2Q(u)}\right)^2 J_2(v) + \dots \tag{4.18}$$

where the $J_n(v)$ depend on the initial wave w by

$$J_n(v) = \int \hat{w}(\eta)\kappa^n(\eta)\mathrm{e}^{i\eta v}d\eta. \tag{4.19}$$

Now we are ready to write the higher order correction terms to the relation between the initial wave w and our measurement s:

$$s = w * r + \left(\frac{tr(t)}{2Q(t)}\right) * J_1 + \frac{1}{2}\left[\left(\frac{t}{2Q(t)}\right)^2 r(t)\right] * J_2 + \dots . \tag{4.20}$$

The basic problem of determining Q lies in the fact that we don't have complete knowledge of either r or w. We have, however, information about the statistical properties of r. A very realistic assumption is that the autocorrelation function of r is a delta function. Writing $\tilde{f}(t) = f(-t)$ this means that

$$r * \tilde{r} = \delta(t). \tag{4.21}$$

Therefore, looking at the autocorrelation of s, we can try to eliminate first the terms involving r in order to reduce our problem significantly. Let us assume for sake of simplicity first that Q is a constant. Then, our expansion reduces to

$$s * \tilde{s} = w * r * \tilde{w} * \tilde{r} + \frac{1}{2Q}(tr) * J_1 * \tilde{w} * \tilde{r} - \frac{1}{2Q}(t\tilde{r}) * \tilde{J}_1 * w * r + \dots . \tag{4.22}$$

Assuming now that all convolutions of r can be approximated by delta functions

$$(t^n r) * \tilde{r} = \beta_n \delta(t) \tag{4.23}$$

we see that, in the above expression, all dependency of r vanishes and we obtain an equation that *only* depends on w and Q but not on r

$$s * \tilde{s} = w * \tilde{w} + \frac{\beta_1}{2Q}J_1 * \tilde{w} - \frac{\beta_1}{2Q}w * \tilde{J}_1 + \dots . \tag{4.24}$$

This result reduces for $Q \to \infty$ to the classical Wiener technique, meaning that then the initial wave w can be found by the power spectrum of s and causality assumptions. Equations (4.24) incorporates the influence of Q in form of higher order correction terms.
There are two major directions for future research on the basis of this result:

1. Approximation of w by a Gabor transform using the first part of the measured signal s and finding Q through the above formula with the estimated w.

2. Use the above result in order to reduce the space for possible candidate functions for w and Q in combination with the other developed techniques.

4.6 Conclusions and Future Work

During the course of the workshop, our team did not conclusively solve the severely ill-posed problem of computing seismic attenuation from given surface seismic trace data. However, we learnt about several techniques currently being used, and began studying some other possibilities. Our efforts can be summarized by:

1. an investigation of the utility and robustness of frequency-related attributes as predictors of anomalous attenuation. These efforts are closely related to those of [6]. Under this heading, we investigated two attributes: the centroidal frequency of a windowed seismic trace, and the second moment of frequency;

2. using wavelet-based techniques to remove reflectivity information from the trace (denoising) and subsequently extract attenuation information

3. an extension of the Wiener technique to the case with attenuation.

In addition to pursuing these further (particularly through numerical implementation of the two latter strategies), we would also like to investigate the use of statistical and linear-algebra techniques to retrieve information about attenuation anomalies.

Bibliography

[1] Hedlin, K. Mewhort, L. & Margrave, G. (2001). *Delineation of steam flood using seismic attenuation*. Expanded Abstracts, 71st. Annual International Meeting, Society of Exploration Geophysicists, pp. 1572–1575.

[2] Kumar, G. Batzle, M. & Hofmann, R. (2003). *Effects of fluids on the attenuation of elastic waves*. Expanded Abstracts, 73st. Annual International Meeting, Society of Exploration Geophysicists, pp. 1592–1595.

[3] Pride, S. R. & Berryman, J. G. (2003). *Linear dynamics of double-porosity dual-permeability materials I. Governing equations and acoustic attenuation*. Physics Review E, 68, 36603.

[4] Pride, S. R. & Berryman, J. G. (2003). *Linear dynamics of double-porosity dual-permeability materials II. Fluid transport equations*. Physics Review E, 68, 36604.

[5] Pride, S. R., Harris, J. et al. (2003). *Permeability dependence of seismic amplitudes*. The Leading Edge, 22(6), pp. 518–525.

[6] Quan, Y. & Harris, J. M. (1997). *Seismic attenuation tomography using the frequency shift method*. Geophysics, 62(3), pp. 895–905.

[7] Winkler, K. W. & Nur, A. (1982). *Seismic attenuation: Effects of pore fluids and frictional-sliding*. Geophysics, 47(1), pp. 1–15.

Appendix A

List of Participants

Vilen Abramov	Kent State University	vabramov@math.kent.edu
Benjamin Akers	University of Wisconsin, Madison	akers@math.wisc.edu
Mohammad Al-Khaleel	McGill University	khaleel@math.mcgill.ca
Amirhossein Amiraslani	University of Western Ontario	aamiras2@uwo.ca
Nancy Azer	University of Victoria	nazer@math.uvic.ca
Boyan Bejanov	University of Alberta	bbejanov@math.ualberta.ca
C. Sean Bohun	Pennsylvania State University	csb15@psu.edu
Mariana Carrasco-Teja	University of British Columbia	teja@math.ubc.ca
John Ceko	Husky Injection Molding Systems	jceko@husky.ca
Benjamin Chan	Cornell University	bsc28@cornell.edu
Wan Chen	University of British Columbia	wanchen@math.ubc.ca
Linping Dong	University of Calgary	dongl@ucalgary.ca
Catherine Dupuis	University of British Columbia	dupuisc@math.ubc.ca
Ian Frigaard	University of British Columbia	frigaard@math.ubc.ca
Peter Gibson	York University	pcgibson@mathstat.yorku.ca
Radu Haiduc	Cornell University	haiduc@math.cornell.edu
Kenneth J. Hedlin	Husky Energy	kenneth.j.hedlin@huskyenergy.ca
Gilles Hennenfend	University of British Columbia	g.hennenfend@evc.net
Felix Hermann	University of British Columbia	felix@eos.ubc.ca
Andreas Hofinger	Johannes Kepler University, Linz	hofinger@indmath.uni-linz.ac.at
Huaxiong Huang	York University	hhuang@yorku.ca
Hooman Javidnia	University of Western Ontario	hjavidn2@uwo.ca
Surrey Kim	Random Knowledge Inc.	surrey.kim@randomknowledge.net
Olga Krakovska	University of Western Ontario	okrakovs@uwo.ca
Michael Lamoureux	University of Calgary	mikel@math.ucalgary.ca
Heejong Lee	Seoul National University	snp110@snu.ac.kr
Jinwoo Lee	Seoul National University	person@snu.ac.kr
Joohee Lee	U of North Carolina, Chapel Hill	jooheele@email.unc.edu
Namyong Lee	Minnesota State University	namyong.lee@mnsu.edu
Greg Lewis	University of Ontario	Greg.Lewis@uoit.ca
Song Li	University of Saskatchewan	song.li@usask.ca

Robert Liao	University of Alberta	rjliao@math.ualberta.ca
Jason Lobb	Carleton University	jlobb@math.carleton.ca
Hongwei Long	University of Alberta	long@math.ualberta.ca
Jack Macki	University of Alberta	jmacki@ualberta.ca
Gary Margrave	University of Calgary	margrave@ucalgary.ca
Bill Mawby	Michelin Tire Corporation	Bill.Mawby@us.michelin.com
Andrei Maxim	Cornell University	maxim@math.cornell.edu
Gabriel Mititica	University of British Columbia	mitica@math.ubc.ca
Carlos Montana	University of Calgary	rcamonta@ucalgary.ca
Tim Myers	University of Cape Town	myers@maths.uct.ac.za
Nilima Nigam	McGill University	nigam@math.mcgill.ca
Anthony Peirce	University of British Columbia	peirce@math.ubc.ca
Stanislava Peker	Concordia University	stanislava@sympatico.ca
Yury Petrachenko	University of Alberta	ypetrachenko@math.ualberta.ca
Peyman Poor Moghaddam	University of British Columbia	ppoor@eos.ubc.ca
Randall Pyke	University College of the Fraser Valley	pyke@math.toronto.edu
Malcolm Roberts	Univeristy of Alberta	mroberts@math.ualberta.ca
Yulia Romaniuk	University of Alberta	romaniuk@ualberta.ca
Kes Salkauskas	University of Calgary	ksalkaus@math.ucalgary.ca
Mahin Salmani	University of Victoria	mahin@uvic.ca
Tobias Schaefer	U of North Carolina, Chapel Hill	tobiass@amath.unc.edu
Azar Shakoori	University of Western Ontario	ashakoor@uwo.ca
Weiguang Shi	University of Alberta	wgshi@cs.ualberta.ca
Shijun Song	University of British Columbia	ssong@stat.ubc.ca
John Stockie	Simon Fraser University	stockie@math.sfu.ca
Tzvetalin Vassilev	University of Saskatchewan	tsv552@mail.usask.ca
Mengzhe Wang	University of Alberta	mwang@math.ualberta.ca
Pengpeng Wang	Simon Fraser University	pwangf@cs.sfu.ca
Zhian Wang	University of Alberta	zhian@ualberta.ca
Rex Westbrook	University of Calgary	westbroo@ucalgary.ca
Jinbiao Wu	York University	jbwu@mathstat.yorku.ca
Lang Wu	University of British Columbia	lang@stat.ubc.ca
Yan Wu	University of Manitoba	umwuy4@cc.umanitoba.ca
Mohammad Ali Yassaei	University of Alberta	ali@math.ualberta.ca
Bo Zeng	Purdue University	bzeng@purdue.edu
Xuekui Zhang	University of British Columbia	xzhang@stat.ubc.ca
Jiaping Zhu	University of Victoria	zjiaping@math.uvic.ca

PIMS Contact Information

email: pims@pims.math.ca

http://www.pims.math.ca

Central Office
Pacific Institute for the Mathematical Sciences
Room 200, 1933 West Mall
University of British Columbia
Vancouver BC V6T 1Z2
Canada