

Chapter 1

Modeling Pre-invasive Bronchial Epithelial Lesions

Rita Aggarwala — rita@math.ucalgary.ca
Alma Barranco-Mendoza — abm@cs.sfu.ca
Fiona Beardwood — fiona@apmaths.uwo.ca
Carole Clem — cclem@bccancer.bc.ca
Perry Fizzano — pfizzano@ups.edu
Martial Guillaud — mguillaud@bccancer.bc.ca
Arvind Gupta — arvind@cs.sfu.ca

Compiled by:

Alma Barranco-Mendoza
School of Computing Science
Simon Fraser University

1 Introduction

Pathologists diagnosing lung cancer in a patient must consider the global architecture of the bronchial tissue, as well as the local architecture of cells and the appearance of individual cells. In order to obtain more detailed information on the condition of the bronchial tissue, a bronchoscopy is performed on the patient and a tissue sample of the lesions is obtained. The pathologist obtains a two-dimensional (2-D) photograph of the sample extracted from a region of the tissue containing abnormal cells.

Often an abnormal cell will die naturally without forming a cancer so the pathologists are concerned only with identifying the cells that will eventually develop into a cancer. Since cancer treatment is very aggressive and traumatic for the patient, pathologists want to be fairly certain that the abnormal cells present in the sample will lead to cancer before recommending treatment. When the sample contains large amounts of abnormal cells or none at all the diagnosis is simple; however, in many cases there are just a few abnormal cells in the sample and diagnosis is difficult.

As part of the PIMSIPS Workshop, two members of the Cancer Imaging Lab of the British Columbia Cancer Research Centre, Drs. Carole Clem and Martial Guillaud, presented a model of pre-invasive bronchial epithelial lesions. This model presupposes a large number of parameters designed to accurately reflect the biological process. The two researchers asked the following ques-



tion: Suppose a 2-D cross-section of tissue from the model is presented. What information can be determined about the original lesion? In particular, is it possible to predict the structure of the three-dimensional (3-D) lesion accurately enough to determine whether the lesion will progress towards cancer? (Figure 1).

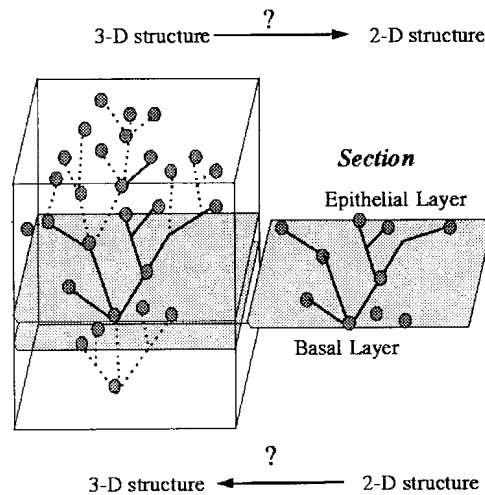


Figure 1: Example of 3-D lesion and 2-D cross-section from the model.

1 Biological Problem

The lung tissue can be seen as divided into three layers: the basal layer, where stem cells divide; an intermediate layer, which thickens as more abnormal cells are present; and the epithelial layer, which is the top layer where cells flatten and die. The tissue is about 10 cells thick. In a normal tissue, a stem cell divides and gives birth to two identical cells. One of the new daughter cells stays in the basal layer and will become a new stem cell, while the other daughter cell differentiates and will slowly move toward the epithelial layer where it will die. A clone is the set of cells that are descendants of the same stem cell. On occasion, abnormalities may occur in a cell. Most of the times, the body has mechanisms that will simply stop the life cycle of such a cell; however, there are a few cases in which the abnormal cell does not die and the cell and its clone multiply out of control.

Flexible fiber-optic bronchoscopy is an important diagnostic technique used for lung cancer. The flexible fiber-optic bronchoscope consists of fiber-optic light bundles for light transmission and auxiliary channels for passing instruments. The flexible video-chip bronchoscope has a small number of glass fibers to deliver the light distally, but the image is recorded directly by a video-chip at the distal end of the bronchoscope [22]. Until recently, the only diagnostic tool available to localize pre-malignant cellular alterations and early bronchial cancer was conventional white light fiber-optic bronchoscopy (FOB). Because only the relatively thick or polyploid lesions are visualized by FOB, only 29% of these lesions were visible to an experienced endoscopist [21], [18]. In an effort to overcome these problems, the Cancer Imaging Laboratory at the British Columbia Cancer Agency (BCCA), Vancouver, has developed the Lung Imaging Fluorescence Endoscopic Device (LIFE) which utilizes differences in tissue autofluorescence to detect precancerous and CIS lesions at a much higher rate



than FOB [15], [16].

The instrument is introduced in the patient's lung orally. To obtain an epithelium biopsy, a 1 cm needle attached to a catheter is placed through the mucosa using the bronchoscope. Using suction, cells are collected for cytologic evaluation. The sample will contain a vertical cross-section of the lung tissue including cells from the three layers. From this sample, 2-D biopsies of the lesion are obtained from which the pathologists must predict whether the lesion will evolve into a malignant tumor or if it will at least not evolve towards cancer.

Recently, new technology has allowed the ready detection of pre-invasive neoplastic bronchial lesions, which are believed to be the possible precursors of malignant tumors. The natural history of lung cancer development from the initial genetic event through other multiple genetic changes, cell-cell, cell kinetics, and cell-host interaction is not completely understood. New techniques (microdissection and PCR amplification) and tools (quantitative cytology and quantitative histology) are elucidating this neoplastic development process. These techniques are generally dealing with snapshots (biopsies, bronchial fragments) of a continuously evolving (changing) epithelium. Current understanding suggests that as pre-invasive neoplastic epithelial tissue becomes more likely to develop into an invasive neoplasia, the amount of genetic changes and genetic heterogeneity in the tissue also increases. While it is still impossible to measure the complete genetic makeup of individual cells in a biopsy or tissue section, it is possible to measure a few selected genetic changes. However, it is still impossible to determine the genetic relationship of all the cells in a pre-invasive neoplastic lesion during the development into invasive cancer. Unfortunately, this is what would be required to completely understand the evolution of normal epithelium into invasive neoplasia. The only feasible alternative is to develop models, which try to simulate the initial stages of the neoplastic process and most importantly to try to simulate the development pathway from normal tissue to abnormal lesion. The simulated development of an abnormal lesion (clonal or multiclonal) requires a model, which takes into account not only the individual cell, but also the whole architecture of the tissue.

2 Model Characteristics

A graphical computer model of the 3-D architecture of bronchial epithelial lesions was developed by Dr. Clem in order to refine hypotheses concerning the progressive spatial disorganization of the bronchial epithelium during the pre-invasive neoplastic process.

There are two main parts in this model. First, there is a static model which simulates the physical arrangement of cells in normal and pre-invasive neoplastic tissue of the bronchial epithelium, and, secondly, a dynamic part which allows the simulation of the continuously interacting nature of living tissue using the 3-D representation obtained from the static model as a starting point.

In the static part, the positions, sizes, shapes and orientations of the nuclei are used as a basis for the 3-D modeling of the architecture. The representation also takes into account the spatial arrangement of the nuclei, with several cell layers modeled. The nuclei are modeled by tri-axial spheroids. The sizes of the major and minor axes of each nucleus are deduced from cytomorphometric analysis. A homogeneous three-dimensional Poisson point process is used to simulate the candidate-positions of nuclei. This point process is layered to take into account the different intensities on the different layers (basal, intermediate and superficial). In addition, the model generates a random angle of orientation of each nuclear axis. Each newly generated nucleus is inscribed in a suitable rectangular parallelepiped with faces parallel to the planes defined by the spheroid axes. If this parallelepiped has an intersection with a parallelepiped of any earlier generated nucleus, the newly generated candidate-position with its nucleus is deleted.

In order to determine whether the model's behaviour has an acceptable range of accuracy for its intended purpose, the system computes the values of 2-D parameters from several computer "sec-



tions” through the simulated 3-D image. An iterative process is used, based on statistical comparison between the 2-D parameters computed and those used from real (2-D) histological sections. If the t-test showed a statistically significant difference between the obtained values and the expected ones, the corresponding values are modified and the process is repeated until no statistically significant differences are found.

The dynamic part of the model can be seen as a tissue growth process applied to the 3-D representations obtained from the static model. Before applying this growth process, an initialization procedure is used in order to define the different cell types that can be found in the tissue (stem or differentiated cells). The simulated tissue can be considered as a closed volume where no cell, even if it is submitted to a force which pushes it out of the box, can go out except by passing into the lumen. Each cell is defined by some internal states which include its capacity of division, its position in the tissue, its age, its displacement capacity, its lifetime and its cell type. Under normal conditions, only the stem cells are able to divide and only the differentiated cells can migrate from basal layer to lumen. At each time step, several events may occur: a stem cell can divide and a new cell can appear; the volume of a stem cell can increase; a differentiated cell can move towards the lumen; a collision between two cells can occur; a cell can die; a nucleus can enter into pyknosis. All these events induce local and global modifications of the tissue architecture and require the model to check the structural stability of the tissue at each time step. Furthermore, all these processes, in order to occur, require an analysis of the local environment of the cell which is involved in one of these events.

The analysis of the local environment of a cell requires the detection of any small changes on the position of a cell and its neighbours. The Gabriel Graph is one of the most commonly used distance mathematical methods for cluster detection. It is sensitive enough to noisy conditions while giving most information necessary for the analysis of the local environment. Simulations of different diffusion patterns of abnormal cells within the bronchial epithelium during the pre-invasive neoplastic process have been obtained as well [4], [5], [6], [7].

3 Assumptions

As with any tractable mathematical model, there were a number of assumptions made about the biopsy procedure and cell behaviour:

1. The cross-section consists of a vertical plane of the lung tissue. That is, the cross-section goes from the basal layer to the epithelial layer.
2. Any given lesion starts with only one abnormal cell, since the probability of an abnormal cell forming from a normal cell is extremely small.
3. The cross-section is taken in the lesion detected by bronchoscopy and will therefore (with high probability) be near the site where the lesion began.
4. Normal cells are formed at the basal layer through cell division of stem cells. They then have a tendency to “drift” towards the epithelial layer. As a simplification to the problem, we will assume that the stem cells will always remain static at the basal layer and the new cells will be the ones moving upwards. This assumption is valid since mother and daughter cells are identical.
5. When a cell reaches the epithelial layer it dies.



4 The General Approach

Clearly there are a large number of parameters to consider in solving this problem. In particular, any mathematical system must form a simpler model that accurately describes Clem's model. To this end, we decided to begin by modeling a 2-D process (that is, we effectively assumed lesions are formed in a 2-D lung). We then studied the difficulty of taking a 1-dimensional cross-section and determining the structure of the 2-D process. In this report we assume that cells divide with a fixed probability and that cells cannot move once they are formed (that is, there is no lateral or upward movement).

2 The 2-Dimensional Model

We will be working with the integer lattice with all points having non-negative coordinates. There will be an initial abnormal cell at position $(0, 0)$. Given an abnormal cell at position (i, j) , an abnormal cell will occur at position $(i + 1, j)$ with fixed probability p ($0 < p < 1$) and independently at position $(i, j + 1)$ with the same fixed probability p ($0 < p < 1$). Since the total height of a cross-section in the 3-D case is at most ten, we will only allow cells to occupy lattice positions (i, j) such that $0 \leq (i + j) < 20, 0 \leq i, j < 10$. That is, we restrict to a height of 10 along the diagonal (i.e. the points $(0, 0), (1, 1) \dots (9, 9)$).

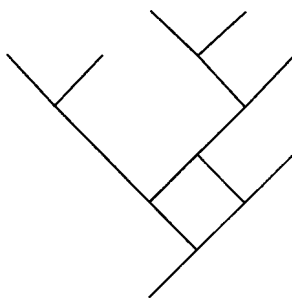


Figure 2: An example growth pattern with $p = 0.6$.

A configuration for $p = 0.6$ is illustrated in Figure 2. Notice that since there are only a finite number of configurations of abnormal cells, it would be possible to enumerate all configurations and assign each a probability (as a function of p). A cross-section of a configuration (the only information we assume is available) will be the line of slope 1 passing through the point $(0, 0)$. As a practitioner, the reproduction rate of abnormal cells, p , will be of interest. Therefore, one problem that we wish to address here is the following:

Given a cross-section of a configuration generated using some probability p , find an interval $[p_L, p_U]$ such that 90% of the time, in repeated experiments, similar intervals will contain the true value p ; that is, find an approximate 90% confidence interval for the parameter p .

Notice that we can view a cross-section as a sequence of 0's and 1's of length 10 where the first element is always 1 and represents the abnormal cell at $(0, 0)$. The sequence for the cross-section in Figure 2 would be 1101000000 since there are abnormal cells at positions $(0, 0)$, $(1, 1)$ and $(3, 3)$.



Mathematically, a cross section can be denoted by (X_0, X_1, \dots, X_9) where $X_i = 0$ if a normal cell is in position i on the cross section and $X_i = 1$ if an abnormal cell is in position i .

1 Experimental results

The model described above was developed and run through computer simulations for various values of p . It is hypothesized (and verified through experimental results) that at $p = .64$ [13] there is a threshold effect – any smaller p results in mainly small configurations whereas any larger p results in most configurations having some cells that reach the last level. Our analysis, described in the following paragraphs, confirms this conjecture.

Using Monte Carlo simulation methods, generating 1000 2-D lattices for each value of p between .1 and .9, incrementing by .1, we were able to determine the (approximate) probability of an abnormal cell being at a particular lattice point in the cross-section, for all points in the diagonal slice under consideration; that is, we were able to find the marginal distributions of the random variables $X_i, i = 0, 1, \dots, 9$. The joint distribution of these indicator random variables would require many more simulations, as there are 2^{10} sample points to consider in the joint distribution. Therefore, a suitable test statistic should be sought in order to make inferences about the parameter p .

Our strategy for estimating p was to choose as a test statistic the last grid point along the cross-section that was occupied (i.e. the position of the right-most 1 in a cross-sectional sequence, or $\max(i : X_i = 1)$). The rationale for choosing this value as a test statistic was that it was evident from running simulations and also intuitively that the last grid point was very sensitive to the true value of p . Therefore, given this information, estimation of the parameter p may be relatively precise. Furthermore, this is a single (univariate) random variable, which can be studied thoroughly using simulation. The choice of a test statistic at this point is highly intuitive, and verification that the test statistic is “good” (that is, a function of a complete and sufficient statistic, unbiased, optimal in terms of variance) has not been formally investigated.

Again using Monte Carlo simulation based on 1000 trials, for values of p between .1 and .9, incrementing by .1, we obtained frequency histograms of the test statistic (comprising the last grid point along the cross-section) (Figure 3). Table 1 shows the results obtained after running the simulation 1000 times.

| p | lvl 0 | lvl 1 | lvl 2 | lvl 3 | lvl 4 | lvl 5 | lvl 6 | lvl 7 | lvl 8 | lvl 9 |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 0.1 | 983 | 16 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0.2 | 915 | 79 | 5 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0.3 | 828 | 126 | 35 | 7 | 3 | 0 | 1 | 0 | 0 | 0 |
| 0.4 | 677 | 190 | 57 | 41 | 15 | 8 | 6 | 3 | 0 | 3 |
| 0.5 | 470 | 189 | 115 | 68 | 37 | 31 | 19 | 27 | 20 | 24 |
| 0.6 | 300 | 102 | 74 | 69 | 43 | 42 | 40 | 47 | 61 | 222 |
| 0.7 | 151 | 47 | 26 | 19 | 8 | 6 | 29 | 36 | 87 | 591 |
| 0.8 | 68 | 1 | 6 | 1 | 2 | 0 | 3 | 10 | 58 | 851 |
| 0.9 | 17 | 0 | 1 | 0 | 0 | 0 | 1 | 2 | 9 | 970 |

Table 1: Table of frequencies of the test statistic (last grid point along the cross-section) after 1000 simulations.

Using these results, we attempted to estimate the value of p used in the sample cross section (Figure 2). A natural method to use with this amount of information is the method of maximum



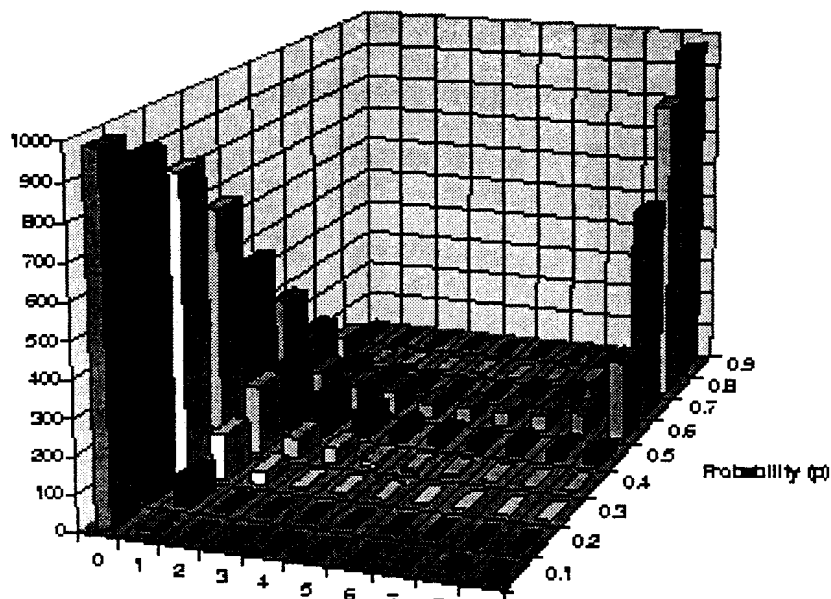


Figure 3: Frequency histograms representing last grid point along the cross-section of 1000 simulations (Table 1). Probability values p between 0.1 and 0.9

likelihood. From the list of frequencies, we found the (approximate) maximum likelihood estimate of p to be 0.6 (that is, the probability function of the test statistic is maximized at $p = 0.6$). This “perfect” estimate is, of course, in light of the fact that we have no histograms for any other values of p between 0.5 and 0.7. We also found an approximate 90% confidence interval for p using these histograms: $[.4, .8]$. Again, the interval would be slightly different (more precise) if simulations for more values of p were performed.

2 Extensions

There are a number of extensions to this model that warrant further investigation. First, we hope to calculate, for a probability p , the probability of a particular configuration occurring, explicitly. Using this information, the joint distribution of (X_0, X_1, \dots, X_9) can be obtained, and the entire sample cross-section may be used to estimate p (we would in this case maximize the joint probability function of (X_0, X_1, \dots, X_9)). This may also allow us to arrive, either explicitly or numerically, at the exact probability distribution of the proposed test statistic (last grid point along the cross-section), and therefore exact estimates and confidence intervals for given confidence levels. The rigorous investigation of the proposed test statistic mentioned earlier would also be possible.

If the above explicit solutions are unfeasible to establish, a next step would be to investigate the distribution of the proposed test statistic for a much finer partition of the interval $(0, 1)$ and to establish much more precise estimates. Next, we will investigate models in which cells are allowed



to move (first vertically and then laterally).

3 Modelling Lesions as Contact Processes

The model proposed above could be seen as a discrete time Markov chain. Hence, we could modify this to be defined over continuous time. Markov chains in continuous time are defined by giving the rates

$$q(x, y) = p(x, y)Q$$

at which jumps occur from state x to state y , where Q is a constant representing the total jump rate and $p(x, y)$ the transition probability at each point of a Poisson process with rate Q .

The finite dimensional distributions of the process at $state_{t[x]}$ at time t is described by the probabilities $P(state_{t[x_1]} = i_1, \dots, state_{t[x_n]} = i_n)$, for each choice of a finite number of sites x_1, \dots, x_n and of possible states i_1, \dots, i_n . The total configuration at time t is described by giving the state of each site x . An initial distribution for the process which does not change in time is called a stationary distribution [17].

If there is a stationary distribution which concentrates on configurations that have infinitely many sites in each possible state then we say that *coexistence* occurs. In most cases in which coexistence occurs there will be a translation invariant stationary distribution where $P(state_t[x] = i)$ is a constant $u[i] > 0$ that we will call the density of type i [17]. *Clustering* occurs if for each x and y the probability of seeing one type of particle at x and a different type of particle at y converges to 0 as t tends to infinity [17].

The *contact process model* was first introduced by Harris in 1974. In this model, each site in the square lattice is either occupied (in state 1) or vacant (in state 0) and follows the conditions:

- (i) An occupied site becomes vacant at a rate δ ; and
- (ii) a vacant site becomes occupied at a rate equal to the fraction of the four nearest neighbours that are occupied. [14].

Much research has been done on these types of models [[17], [8],[9], [10],[1]], but perhaps the most important result on contact process is the *Complete Convergence Theorem*:

When the contact process does not die out then it will converge to the stationary distribution that is the limit starting from all 1's [11].

An immediate consequence of this is that the only stationary distributions for the process are:

- (i) the limit starting from all 1's,
- (ii) the trivial stationary distribution which assigns probability one to the all 0's configuration, and
- (iii) p times (i) plus $(1 - p)$ times (ii) [11].

An interesting modification to this model was presented by Durrett and Levin in 1994 when they proposed that the behaviour of stochastic spatial models could be determined from the properties of the mean field ODE [12].

Going back to our model, we will rephrase it in terms of a modification to the contact process model. We will define our diamond lattice (square lattice rotated 45 degrees) to be at most 10 cells in the diagonal since that is a characteristic of the lung tissue. Each site in the diamond lattice is either occupied by an abnormal cell (in state 1) or vacant (in state 0) and follows the conditions:

- (i) a vacant site becomes occupied at a rate equal to λ times the fraction of the four nearest neighbours that are occupied, and



(ii) an occupied site becomes vacant at a rate equal to δ times the fraction of the four nearest neighbours that are vacant,

where $\delta \leq \lambda < 1$ is the rate at which abnormal cells split and $0 \leq \delta < 1$ is the small probability of an abnormal cell being displaced from the site by a healthy cell. The reason for these constraints is that we are interested in the problem where both normal and abnormal cells coexist (at least in the early stages). It is easy to see that if $\delta \geq 1$ the process would die out, i.e., there would be total recovery; and if $\lambda \geq 1$ then the abnormal cells would take over the entire tissue.

In practice, the vacant sites are not actually vacant but occupied by healthy cells that can be displaced by abnormal ones. However, since we are currently only concerned about the growth of abnormal cells, considering the healthy sites vacant simplifies the problem considerably. Condition (ii) is necessary since there is a very small probability that a healthy cell may displace an abnormal cell. When this occurs, since we are not allowing cells to drift in any direction on the plane, the new cell will be pushed out of the 2-D plane. Hence, for this simple 2-D model, this situation is resolved by setting $\delta = 0$; i.e., once a site is occupied by an abnormal cell, it will never become vacant. However, the condition $\delta > 0$ must be considered when introducing drift and when building the 3-D model.

Let $S_0 = \{\text{finite subsets of } \mathbf{Z}^2\}$. If $A \in S_0$ is the original set of abnormal cells, then let ξ_t^A be the set of sites occupied by abnormal cells at time t . We can rewrite the above as Markov processes $(\xi_t^A)_{t \geq 0}$ and their jump rates are given by

$$A \rightarrow A \cup \{x\} \quad (x \notin A) \text{ at rate } \lambda |\{y \in A : \|y - x\| = 1\}|,$$

$$A \rightarrow A - \{x\} \quad (x \in A) \text{ at rate } \delta |\{y \in \bar{A} : \|y - x\| = 1\}|,$$

where $\|x\|$ is the distance from x to 0; i.e., the rate at which a site becomes occupied by an abnormal cell is dependent on the cardinality of the set of sites occupied by abnormal cells adjacent to the current site.

Note that if $\delta = 0$ and $\lambda = 1$, then our model is a finite version of Richardson's growth model presented in [19]. There Richardson showed that if $B(t)$ is the set of sites occupied at time t , then $B(t)/t$ clusters to a limiting shape, which is roughly but not exactly circular. Since we assumed that the process started with a single abnormal cell at the origin, then we are only interested in the process ξ_t^0 .

Hence, we can represent our model using a partial differential equation to describe the stochastic process. We will consider the following equation:

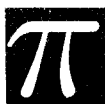
$$\frac{\partial n}{\partial t} = \sigma \Delta n + \lambda n + \mu_y \frac{\partial n}{\partial y}$$

where n represents the number of abnormal cells in the region of interest. We quickly summarize the various parts of the equation.

The term:

$$\frac{\partial n}{\partial t} = \sigma \Delta n$$

represents the diffusion equation, and models the random movement of the abnormal cells in the tissue over time, where σ is the diffusion constant. The birth rate of the abnormal cells is controlled by the parameter λ . Clearly this term allows the number of abnormal cells spawned at any given time to grow linearly with the current number of abnormal cells.



Since there is a natural upwards drift of the cells in the tissue, we use the term:

$$\mu_y \frac{\partial n}{\partial y}$$

to model this phenomenon. Notice that this term depends on the distribution of the cells in the vertical direction, since we can expect more cells to drift upwards if there are more cells clustered near the basal layer of the tissue than elsewhere.

To model other aspects of the biological processes occurring in the diseased tissue, additional terms are required. For example, the term:

$$\mu_x \frac{\partial n}{\partial x}$$

can be added to model the lateral drift of the cells; the rate can be controlled through the parameter μ_x .

Finally, we need to apply suitable boundary conditions. For instance, we could set the Dirichlet conditions:

$$n(0) = 1$$

$$n(10) = 0$$

at the bottom and top respectively of the cell layer. This makes sense physically since in this simple model we only allow one abnormal cell at the basal layer, and assume that once cells reach the top layer they die. The boundary conditions to model the sides of the region are more complicated. A possible solution would be to use moving boundary conditions at these edges, so that as the lesion expands the boundaries would also expand.

4 Conclusions and Future Work

The growth of cancer cells involves many different processes which can only be captured by a complex model. However, simplified models provide a great deal of insight into the fundamental processes involved. In this workshop we proposed two simple models – one discrete stochastic model and one PDE model – to solve a 2-D simplification of the original problem.

It is worth mentioning that, after the PIMSIPS workshop, we came across a model of skin cancer, presented for the first time by Williams and Bjerknes, that follows a similar approach to ours. In their model, each site is either occupied (in state 1) or vacant (in state 0) following the conditions:

- (i) an occupied site becomes occupied at a rate δ times the fraction of the four nearest neighbours that are occupied, and
- (ii) a vacant site becomes occupied at a rate equal to the fraction of the four nearest neighbours that are occupied.

Letting $B(0)$ be a finite set and $B(t)$ be the set of lattice points occupied at time t , if $B(t)$ is ever the empty set then it will remain so for all time, in which case we say the model dies out [20]. In other papers, Bramson and Griffeath showed that if $B(t)$ does not die out, then $B(t)/t$ has a limiting shape [2], [3]. The main difference between this model and ours is that, by the nature of the problem, skin cancer growth was modelled only as sidewise splitting on the basal layer in such a way



that the surface folded onto a torus. On the other hand, our lung cancer model is very constrained since we are modelling all three layers as cross-sections of the bronchial epithelium, which forces the model to be restricted to a finite height along the diagonal. However, the model of Williams and Bjerknes gives veracity to ours since their approach, similar to ours, is plausible and well studied.

Our next step is to add both horizontal and vertical drift to the 2-D discrete model. This will make it more natural to model as a stochastic process as well as make it closer to the experimental observations. There is also a natural extension to using a 3-D lattice for the discrete model. In the partial differential equation model we will require another independent variable to handle movement of cells in the third dimension. Hence, if μ_x and μ_z are the rates of lateral drift at which cells move along the x - and z -axis, respectively, then

$$\frac{\partial n}{\partial t} = \sigma \Delta n + \lambda n + \mu_y \frac{\partial n}{\partial y} + \mu_x \frac{\partial n}{\partial x} + \mu_z \frac{\partial n}{\partial z}$$

By combining analytic techniques with computer simulations we hope to produce a model that is useful in modelling the growth of cancer cells and predicting the existence of lung cancer at an early stage.

5 Acknowledgments

We want to thank Dr. Rick Durrett from the Mathematics Dept. at Cornell University for his valuable ideas and suggestions in the postulation of the model as a contact process.



Bibliography

- [1] Bezuidenhout, C. and Grimmett, G., “The critical contact process dies out,” *Ann. Prob.*, 19: 984-1009, 1990.
- [2] Bramson, M. and Griffeath, D., “On the Williams-Bjerknes tumor growth model, II,” *Proc. Camb. Phil. Soc.*, 88:339-357, 1980.
- [3] Bramson, M. and Griffeath, D., “On the Williams-Bjerknes tumor growth model, I,” *Ann. Prob.* 9:173-185, 1981.
- [4] Clem, C. J., Guillaud, and MacAulay, C., “Computer simulation of the 3-D tissue organizational changes associated with development of pre-invasive neoplastic bronchial epithelial lesions,” *unpublished research*, 1997.
- [5] Clem, C. J., König, D. and Rigaut, J. P., “A three-dimensional dynamic simulation model of epithelium tissue renewal,” *Analytical and Quantitative Cytology and Histology* 19:174-184, 1997.
- [6] Clem, C. J. and Rigaut, J. P., “Computer simulation modelling and Visualization of 3-D architecture of biological tissues,” *Acta Biotheoretica*, 43:425-442, 1995.
- [7] Clem, C. J., Boysen, M. and Rigaut, J. P., “Towards 3-D modelling of epithelia by computer simulation,” *Analytical Cellular Pathology*, 4:287-302, 1992.
- [8] Durrett, R., “On the growth of one dimensional contact processes,” *Ann. Prob.*, 8:890-907, 1980.
- [9] Durrett, R. and Griffeath, D., “Contact processes in several dimensions,” *Z. fur Wahr*, 59:535-552, 1982.
- [10] Durrett, R. and Schonmann, R.H., *Stochastic growth models. In Percolation Theory and the Ergodic Theory of Interacting Particle Systems*, edited by H. Kesten. Springer Verlag, New York, 1987.
- [11] Durrett, R., “The contact process: 1974–1989,” pp. 1-18 in *Mathematics of Random Media*, edited by W.E. Kohler and B.S. White, American Math. Society, 1992.
- [12] Durrett, R. and Levin, S., “The importance of being discrete (and spatial),” *Theoret. Pop. Biol.*, 46:363-394, 1994.
- [13] Durrett, R., personal communication.
- [14] Harris, T.E., “Contact interactions on a lattice,” *Ann. Prob.* 2: 969-988, 1974.



- [15] Hung J, Lam S, LeRiche JC, Palcic B., "Autofluorescence of normal and malignant bronchial tissue," *Lasers Surg. Med.*, 11:99-105, 1991.
- [16] Lam S, MacAulay C, Hung J, Palcic B., "Detection of dysplasia and carcinoma in situ with a lung imaging fluorescence endoscope device," *J. Thorac. Cardiovasc. Surg.*, 105:1035-1040, 1993.
- [17] Liggett, T.M., *Interacting Particle Systems*, Springer, New York, 1985.
- [18] Nagamoto N, Saito Y, Sato M, Sagawa M, Kanma K, Takahashi S, Usuda K, Endo C, Fujimura S, Nakada T., "Lesions preceding squamous cell carcinoma of the bronchus and multicentricity of canceration—serial slicing of minute lung cancers smaller than 1 mm.," *Tohoku J. Exp. Med.*, 170:11-23, 1993.
- [19] Richardson, D., "Random growth in a tessellation," *Proc. Camb. Phil. Soc.*, 74:515-528, 1973.
- [20] Williams, T. and Bjerknes, R., "Stochastic model for abnormal clone spread through epithelial basal layer," *Nature*, 236:19-21, 1972.
- [21] Woolner LB, Fontana RS, Cortese DA et al., "Roentgenographically occult lung cancer: Pathologic findings and frequency of multicentricity during a ten year period," *Mayo Clin. Proc.* 59:453-466, 1984.
- [22] The Virtual Hospital, (internet web page)
<http://indy.radiology.uiowa.edu/Providers/Textbooks/LungTumors/Diagnosis/Bronchoscopy/Text/Bronchoscopy.html>, University of Iowa, 1997.



Chapter 2

Fingerprint Identification — New Directions

Group Members:

Petr Lisonek, *Simon Fraser University*
Akbar Rhemtulla, *University of Alberta*
Rolf Luchsinger, Michael Saliba, Alexandre Zagoskin, *University of British Columbia*
Dan Calistrate, Marc Paulhus, *University of Calgary*

Compiled by:

Michael Saliba
Dept. of Physics and Astronomy
University of British Columbia

25 September, 1997

1 Problem Description

A brief description of this problem as presented by Kinetic Sciences Inc. (KSI) can be found at the beginning of these proceedings. In summary, KSI have developed a small live-scan sensor (imager) that can be mounted onto a small silicon chip. KSI are developing this technology with the objective of producing a fingerprint imager that can be used to limit access to sensitive information or locations. Examples of possible applications of this technology include ATM machines, computer keyboards, and door-unlocking mechanisms.

The KSI sensor is capable of producing an image of the fingerprint that has a resolution of 850 dpi, and that provides high contrast between the ridges and valleys of the fingerprint. A sample



image of part of a fingerprint is shown in Figure 1. The finished product will be able to produce an image of a large portion of the fingerprint (refer to Figure 2). In order for this technology to be successful, KSI requires the use of a computer image-matching algorithm that can make fast and reliable comparisons, in real time, between a scanned fingerprint image and images that are contained in a digital database.

The problem that was submitted to the workshop by KSI involved finding the best model on which to base such an algorithm. The basic questions were:

1. What general features of a fingerprint image should the algorithm search for?
2. What methods should be used to locate these features?

Approach

Conventional fingerprint matching algorithms search for anomalies in the ridge flow (ridge endings and ridge bifurcations) and then, after appropriate post-processing, compare the relative locations of these anomalies to those of the reference fingerprint(s). Our approach during this workshop was to identify new features of the fingerprint that could be used for matching, and to find efficient ways to locate the various features and to store the information.

We identified the following main problems that would need to be addressed by a robust algorithm:

1. Day-to-day changes in the fingerprint (e.g. due to the weather, or due to physical condition).
2. Variations in sensor response (e.g. due to variations in finger pressure applied during scanning).
3. Variations in image location for processing (due to variations in position and/or angular orientation of the finger during scanning).

Furthermore, the following constraints also needed to be addressed:

1. Limits on database memory.
2. Limits on computational time required for matching.

A number of models for suitable algorithms were suggested and discussed at length during the workshop, and the more promising of these are described in the following pages. The various suggestions are presented in the form of a set of individual papers, and a general conclusion is given at the end of this report.



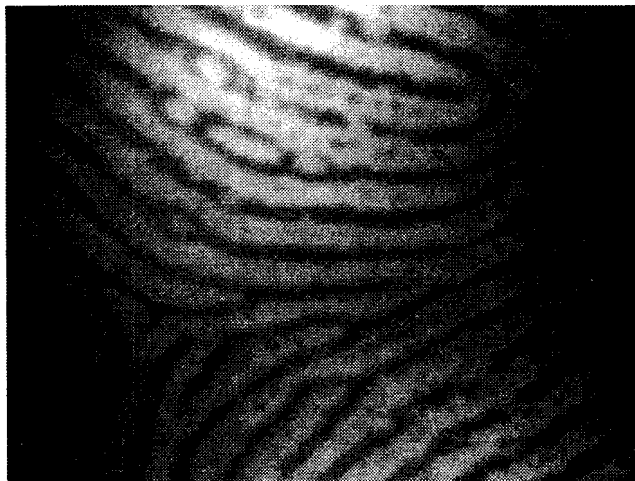


Figure 1: A sample fingerprint image, taken by the KSI live-scan sensor.



Figure 2: Ridge pattern of an inked fingerprint.





Figure 3: Locating the reference point on the image.

2 Finding a Reference Point on the Image

Akbar Rhemtulla¹ and Michael Saliba²

Some of the algorithms that are described on the following pages may be sensitive to changes in the positioning of the finger during scanning. Our first suggestion therefore involves an algorithm that could be used to locate a reference point on the image.

Observation of the fingerprint samples that were available during the workshop suggest that many prints have a clearly identifiable centre, or “heart”, that is either the centre of a spiral ridge or the tip of an inverted U-shaped ridge. An example of the first case is shown in Figure 3. An algorithm to locate the centre would involve drawing a series of horizontal lines down the image as shown in the figure, and locating the position (end-points) of the innermost curve that crosses the line in each case. This should be done until three or more successive lines have the same innermost curve, or until this curve intersects the line at one point only. In Figure 3 this would occur at line 4. In the case of a spiral ridge, this method could also search for the position where the innermost ridge curves *below* the line instead of above it.

¹akbar@malindi.math.ualberta.ca

²saliba@physics.ubc.ca



When the general location of the heart has been found, a higher resolution search can be carried out between the last three or four lines in order to obtain a more exact location of the reference point.

3 An Algorithm for Comparing Fingerprints

Dan Calistrate³ and Marc Paulhus⁴

1 Introduction

In this report we shall describe an algorithm which assigns a sequence of integers to a fingerprint image. This sequence can then be used to decide if two fingerprints are identical. The sequence will be topologically invariant and we will describe how scratches on the finger might be ignored by the algorithm.

Section 2 will present the basic idea of the algorithm. Section 3 will address some of the obvious objections or problems that there might be. Potential solutions to these problems will also be given. Section 4 will highlight some of the advantages of using this technique.

2 Basic Idea

Our basic idea is to find a topologically invariant sequence of integers which uniquely describe a fingerprint (or at least uniquely enough). This reduces the problem of comparing two fingerprint images to the problem of comparing two integer sequences. We will call the sequence of integers which come from a fingerprint image the *image sequence*.

We now describe how to produce the image sequence. Given a fingerprint, start at the lefthand-bottom corner of the image. Choose the first valley in the print and follow along this valley using the right-hand rule of navigating⁵. As you walk down this valley, increment a counter every time you come to a fork in the road. This assigns an integer to that particular valley. Do the same for the next valley on your way up the left side of the fingerprint. You will generate an integer sequence, one integer for every valley you walk down. We will assume that we stop walking when we come to a cul-du-sac although this is not necessary.

If a longer sequence is needed you can follow the same procedure using the lefthand-rule or generate the sequence corresponding to the valleys which start on the right side of the image.

You will find an example for an actual fingerprint in Figure 4. In the example there are 28 valleys starting on the left side of the fingerprint. Below we list the integers corresponding to these valleys.

| | | | | | | | | | | | | | | |
|---------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Valley | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| Integer | 1 | 2 | 2 | 3 | 2 | 3 | 1 | 4 | 1 | 1 | 2 | 0 | 4 | 2 |
| | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 |
| | 3 | 8 | 1 | 1 | 0 | 1 | 2 | 3 | 0 | 0 | 1 | 1 | 0 | 2 |

³calistra@math.ucalgary.ca

⁴paulhusm@math.ucalgary.ca

⁵The right-hand rule for navigating is a well defined concept which amounts to always following the wall to your right side as you navigate a maze. One can think of putting one's right hand on the wall and never removing it as one walks.



The method described will produce a well defined integer sequence for any particular fingerprint image. Two images can then be compared by comparing their integer sequences. There are some obvious problems which we hope to address in the next section.

3 Potential Problems

- The first objection is that the starting line on the bottom could be different in two scanned images of the same fingerprint. Hence we should not ask for an exact duplication of the image sequence but instead look for one sequence as a subsequence in the other.
- Another objection is that the starting points of the valleys on the left side of the fingerprint could be different for two different images. This could affect the length of the fingerprint sequence as well as the value of the digits in the sequence. We suggest that, as long as the fingerprints are truly identical, the choice of a lefthand side starting point will only change a small number of the digits in the image sequence. Also, a comparison program which is intelligent enough to try a backtracking sequence matching algorithm with deletion, could apply a probability of match even when the sequences differ in length or when the sequences are essentially the same except for extra digits and small deviations. A cumulative sum of differences might also be useful.

Some work has been done in the field of sequence comparison. For example, biologists who study DNA sequences commonly compute the probability that two integer sequences are the same, modulo errors in observations.

- Being topologically invariant, the sequence would be very sensitive to scratches on the fingerprint.

If the scratch cuts across the valley you are currently following, then, while walking down this valley, the scratch will appear as an intersection of degree four. In this case ignore the intersection altogether (i.e. do not increment the counter) and continue along the middle (main) valley. Hence scratches of this type will not affect that integer in the sequence. The authors have observed that intersections of degree four rarely occur naturally in a fingerprint (in fact we do not know if they occur naturally at all). In any case, as long as degree four intersections are consistently ignored, the assigned integer will be the same.

If the scratch runs along the valley you are currently following, there is potential for a large deviation in the assigned integer. This large deviation should only occur in one (or at most a few) digits in the sequence. An intelligent enough sequence-comparison program should be able to deal with this deviation.

- The image sequence should be fairly robust under reasonable rotations of the image. No matter which identification algorithm is used, if the user is instructed to place his finger flush against a wall beside the sensor, many problems may be avoided.
- The computational complexity of “Walking down the valley” is unknown to the authors but it does not seem like an unreasonable task.
- Depending on image quality and resolution, some form of image massage might be required to make the procedure robust. The nature of this massage can only be determined by experimentation.



- If the proper rules of navigating the valley are used, then no infinite looping can occur. Even if a navigation rule was chosen which had the potential for infinite looping, an artificial stopping criteria could be imposed (for example stop walking down a valley when your counter reaches some value).
- Is the image sequence sensitive enough to tell fingerprints apart? Only experimenting with a large database of fingerprints can answer this question. Since the sequence can be made quite long we suspect it is sensitive enough.
- It is possible that, for some fingerprints, there is a nearly vertical valley running up the sides of your image. In this case the image sequence is forced to be quite short. A number of possible solutions can be proposed. Which solution works best would have to be determined experimentally.

One way is to translate the starting line along the fingerprint until the starting line intersects the maximum number of valleys. Then assign integers to the valleys (using the same set of rules as before) which radiate from the starting line in the left direction (and then the right direction, if a longer sequence is desired).

Another solution would be to assign integers to valleys starting from the top and the bottom edges as well.

4 Advantages

- Since the image sequence is all that is used to compare two fingerprints, very little memory is required.
- The matching procedure could be made very fast. Also, it has many parameters that could be set by the user for desired levels of security.
- It is a topologically invariant sequence which has the potential to ignore scratches. We think this is quite powerful.

4 Fingerprint Matching - A Topological Approach

(Abstract)

Petr Lisonek
CECM, Simon Fraser University
e-mail: lisonek@cecm.sfu.ca
September 10, 1997

The geometrical approach to fingerprint matching has its drawbacks. In distance-based methods (measurement of distances between minutiae) these are, for example, effects of the ambient temperature and humidity and distortion caused by the softness of the tissue. In grid methods (considering the number of intersections, slopes, etc. relative to a given grid) we additionally encounter the problem of grid positioning whose uncertainty is much higher than the required precision (which is a fraction of the ridge/valley width) and may grossly override the fine resolution offered by the scanning technology.



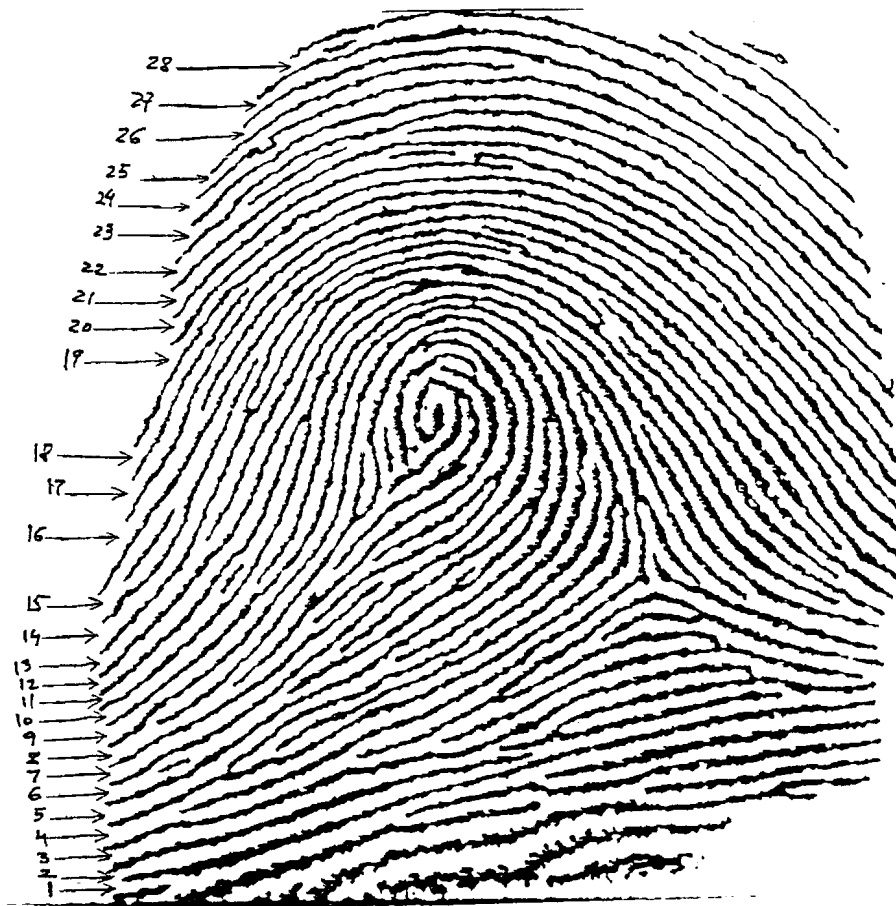


Figure 4: Fingerprint image showing the valley numbers.



Our approach proposes to encode the fingerprint using topological features found in it (spirals, bifurcations, closed ovals, etc.). These structures are (by their mathematical nature) much less sensitive to all kinds of metric distortions described above. The fingerprint is represented as a tree, whose root corresponds to the entire image, and whose branches represent the nesting/branching relations between the corresponding structures in the image. On the low levels of the tree we explicitly encode the distinguished topological features (examples above). Our method may look computationally intensive but certain patterns (spirals, ovals etc.) can be recognized in a single scanning sweep, and thus in real time, given the assumed technology of scanning that was proposed in the project description.

In the initial approximation our method is equivalent to representing the fingerprint as a planar graph (as known from Graph Theory) along with a given planar embedding, where vertices correspond to minutiae and edges to ridges between them. (This can be slightly enriched by adding an encoding for isolated closed ovals, which have no minutiae but still contain information, in the event that such features appear in fingerprints at all.) The fingerprint matching is then reduced to planar graph isomorphism (consistent with given planar embeddings) for which fast algorithms are known.

The set of distinguished topological features, some of which were suggested above, has to be determined by studying a large variety of prints and considering the complexity of their recognition in the given scanning technology (sweep).

5 Fingerprints: Global methods

Rolf Luchsinger

Dept. of Physics and Astronomy, University of British Columbia

e-mail: `luchsi@theory.physics.ubc.ca`

1 Introduction

Instead of looking for local features in the structure of the fingerprint picture, we propose here to map the entire picture of the fingerprint into a space where the specific features of the fingerprint might be better revealed. This is what we mean by a global method. Two possible mappings now commonly used in signal processing are the Fourier Transformation and the Wavelet Transformation. Due to personal experience with Fourier methods and a lack of experience with Wavelets, we will focus in the following on Fourier methods. However Wavelet methods might give better results for the problem at hand.

2 Method

The idea of the Fourier method is to decompose the picture (fingerprint) into a set of periodic structures. The wavelength of the periodic structures varies from the largest possible that fits into the image frame to the smallest wavelength that is comparable to the finest structures that need to be represented. The size of a Fourier coefficient at a given wavelength is a measure of the importance of the periodic structure of this wavelength in the picture. The collection of these Fourier coefficients is called the Fourier spectrum. Thus having a spectrum where only long wavelength coefficients are important, we can say that the picture has a smooth structure. If the short wavelength coefficients dominate, the picture has a rough structure. In that sense the different structure scales of a picture can be separated by the Fourier transformation.

Looking at fingerprints, we see in the first approximation a somehow regular structure of black and white lines. This regular structure will be revealed in the Fourier spectrum by a prominent peak



Fourier-Spectrum of a Fingerprint

(shown in one Dimension)

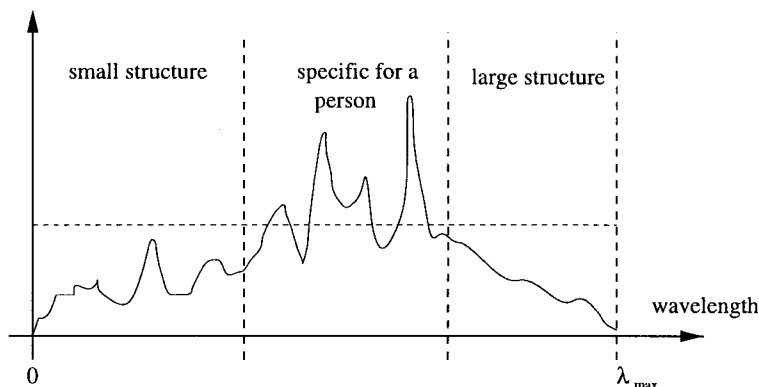


Figure 5:

and is a characteristic feature of the fingerprint. Other peaks will show up in the spectrum which will characterize the fingerprint. The short wavelength scale of the spectrum will be dominated by all kinds of noise and is not specific to a fingerprint. Since the Fourier transformation is a one to one mapping, all the information of the picture is contained in the spectrum. The person specific fingerprint information will presumably be contained in the most prominent Fourier coefficients. A possible Fourier spectrum of a fingerprint is shown in Fig. 5.

The basic idea of the Fourier method is to have a tool which can filter out the person specific information. Since Fourier methods are generally used for data compression, this approach seems to be promising. However the efficiency of the Fourier method depends on whether the characteristic information of a given fingerprint is revealed in a reasonably small number of Fourier coefficients. If this is the case, only this set of important coefficients has to be matched to the true person's set in order to classify a fingerprint. Looking at the variation of the spectrum using different fingerprints or the same fingerprint scanned several times is the only way to determine the efficiency of the method. Unfortunately this was not possible during the workshop.

3 Advantages

1. The method is fast and simple. You feed the picture of the fingerprint to a Fast Fourier Transformation algorithm, cut out the interesting region in the spectrum and compare it with your key data using a least square method with some tolerance criteria. (You don't need to follow lines, find bifurcation points etc. which is a lot of work for a computer.)
2. Fourier methods are in general well understood and used in many applications of signal processing.
3. The location of the finger on the sensor is not important. Shifting the coordinate system will just multiply a phase factor to the Fourier coefficients. This doesn't affect the absolute value of the Fourier coefficients.
4. If the high resolution of the image is useful, (that means it is not mainly noise you see on the small scale), you can make use of this information by including those small wavelength peaks.



5. The method is not sensitive to local damage to the fingerprint (cuts, dirt, etc.) .

4 Disadvantages

1. The method is a very general approach to signal processing. Whether it is reliable and optimal for the identification of fingerprints is an open question. However, fine tuning or a Wavelet approach might be possible. – > Check it out.
2. If it is the method of choice, it might have already been used in this context (patents).

6 Fingerprint Identification by the Measurement of the Angular Orientation of Successive Ridges within Narrow Vertical Bands

Michael Saliba

Dept. of Physics and Astronomy, University of British Columbia

e-mail: saliba@physics.ubc.ca

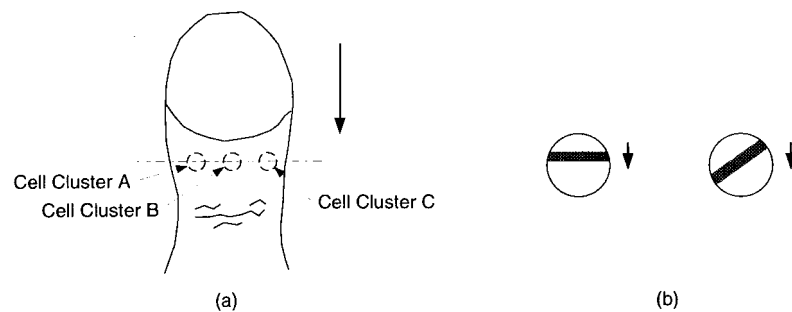


Figure 6: Location of the sensors for fingerprint scanning.

1 Introduction

This suggestion is intended to provide an alternative approach to the more conventional method of applying a pattern recognition algorithm to a static 2-D image. It attempts to take advantage of the specific characteristics of the KSI sensor, and the already selected method of scanning. It does depend rather heavily, however, on the actual capabilities of the sensor.

A more conventional variation of this approach is also suggested at the end of the next section.

2 Concept

Consider a sensor with only three active clusters of cells A, B, and C, each of about 1 mm diameter, as shown in Figure 6(a). As the finger moves down over the sensor, each cluster will be subjected to the passage of successive bright and dark fringes, corresponding to the ridges and valleys of the



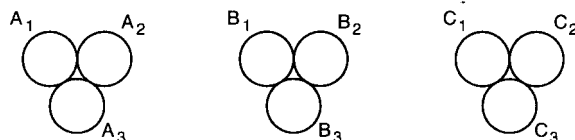


Figure 7: Utilization of sets of cell clusters for increased robustness.

fingertip topography. A horizontal fringe will cross the field of view of the cluster in a shorter time than will a fringe that has an oblique angular orientation (Figure 6(b)).

If the sensor can be made capable of registering the time that it takes for each fringe to cross over the field of view of each cluster, then each cluster can be made to output a string of numbers⁶ corresponding to the times taken by successive fringes to pass across the cluster. For fingerprint identification the three sets of numbers from clusters A, B, and C can then be compared to the reference sets in the database. A suitable algorithm, with the appropriate tolerances, can be utilized to make the comparisons.

For example, it may be sufficient to search for a matching string of only six digits within the much larger string in each of the three sets; or, alternatively, the algorithm would take only (say) every third fringe into consideration. Each digit would also be allowed a certain tolerance (e.g. ± 2 bins) as determined during the testing of the prototype. As a further alternative, the algorithm could also be made to compare the *ratios* of the times between successive ridges, and/or the thickness of the fringes. If the independent matching of the three strings does not provide sufficient security, they can be tested for a correlation on the vertical (i.e. time) scale.

In the final design, in order to allow for possible variations in the lateral positioning of the finger in successive passes, each of the three clusters A, B, and C should actually be replaced by, say, three clusters that are slightly displaced from each other laterally, as shown in Figure 7. The three sets of numbers from clusters A₁, A₂, and A₃ can be compared to the “A” set in the database, and the best set can be selected (the same method is applied for sets B and C). This procedure may also correct for lateral expansion/contraction of the finger from day to day, as well as for the occasional ridge irregularity or junction that may happen to pass through the field of view of the cluster. Since the output strings of numbers will depend primarily on the *macroscopic angular orientation* of successive ridges, they are not expected to be very sensitive to ordinary day to day changes in fingertip topography, or to minor variations in the sensor response. Slight variations in the angular positioning of the finger can be made to fall within the tolerance of the comparing algorithm.

Apart from the possible constraints due to the sensor capability, a possible drawback to this concept may be due to its use of only a *part* of the fingerprint in order to make the identification (specifically, three⁷ rather narrow vertical segments taken from the total 2-D image). It is however quite probable that there is in fact sufficient information within such a portion of the map. A method would also need to be devised to deal with or discard regions where the ridges may be oriented vertically. A major advantage of this overall concept is the small amount of memory needed for the database, as well as the use of relatively simple algorithms that would be computationally very fast.

A more conventional variation of this method would involve using an algorithm that simply measures the angular orientations, and possibly also the absolute and relative frequencies and thicknesses of the fringes in selected bands of the static 2-D image. An example with three bands is shown in Figure 8. The signature of a fingerprint is extracted exclusively from within these narrow regions.

⁶The measured times can be digitized into appropriately sized bins.

⁷Testing of the prototype would help to establish the optimum number of segments that should be utilized.





Figure 8: Selected vertical bands of the static image.

7 Fingerprint recognition algorithms

A. Zagoskin

Dept. of Physics and Astronomy, University of British Columbia

e-mail: zagoskin@physics.ubc.ca

- **Objective:** recognition of a sample fingerprint from a set of 100-1000 for the purposes of user access.
- **Conditions:** small computing time/memory requirements; robustness with respect to natural variations of the user's fingerprint
- **Algorithm A**

1. **Description:** The fingerprint is coded as a square matrix of integer numbers by placing it on a rectangular $N \times N$ grid and assigning each edge of the grid a number n of its intersections by the fingerprint lines. (The actual size of the grid is variable, being determined by the size of the fingerprint, Fig 9.) Then each cell is assigned either a number $M = (n_{top} - n_{bottom}) + (n_{right} - n_{left})$ (algorithm A1), or its parity ($P(M)=0$ if M is even, $P(M)=1$ if M is odd) (algorithm A2) (refer to Fig 10). The resulting matrix is compared to the sample one stored in the system; if the number of matches is below the recognition margin, the access is denied, and vice versa.

2. **An example:** The algorithm A2 was applied to two different fingerprints. (For the sake of simplicity we chose the initial grid 8×8 ; of it the central subgrid 6×6 was used in order to limit the effects of near boundary distortions.) The code matrices (6×6) are presented in Figs 11a,c. In order to model the natural variations of shape and positioning of the same fingerprint, the fingerprint (a) was also coded using a grid that was slightly distorted and displaced (by approx. 5%); the resulting matrix is presented in Fig 11b.

The matching of code matrices is shown in Fig 12. The result suggests that there is a sufficient recognition margin that makes it possible to distinguish between natural variations of the same print (matching 66%) and different fingerprints (matching $\sim 40\%$).



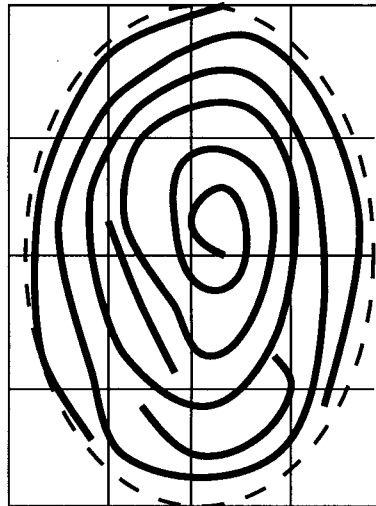


Figure 9: Grid positioning

Of course, in order to make a decisive conclusion and determine statistically significant recognition margins, a large pool of samples must be investigated.

3. **Advantages:** Simplicity; little computation time and memory required, so the check can work in real time; robustness with respect to imperfections of the print (e.g. it is not very likely that the grid line will pass exactly through an accidental gap in the ridge, thus changing the code numbers; in case of algorithm A2 accidental finger cuts will generally change the code numbers of at most 2 cells).
4. **Difficulties:** Sensitivity to the positioning of the grid; e.g., if a ridge goes through an intersection of the grid lines, or ends on the line, small variation of grid position will change the code numbers of several cells at once.
5. **Possible solutions:** The system can be programmed to perform several scans with a varied grid size, and throw away the “rogue” readings. The fringe of the grid can be discarded as in the above example; one expects the distortions to be most prominent there. The optimal grid dimension N should be determined from a large pool of fingerprints, but I estimate that the optimum will be reached at about 5 intersections per grid edge.

- **Algorithm B**

1. **Description:** This is a simpler version of the algorithm A, where instead of the $N \times N$ matrix, the fingerprint is characterized by two sequences of length $N + 1$ (Fig 13). Each term gives the number of interstections of a given *line* of the grid with the ridges. Thus a fingerprint is presented by two “scans” in perpendicular directions.
2. **Advantage:** This method is evidently simpler, faster, and more robust than the algorithm A, since the relative variation in the total number of ridges scanned across the whole grid is less than on a single edge. The shifts in the finger position are also less important, since matching of two one-dimensional sets can easily allow for such shifts.



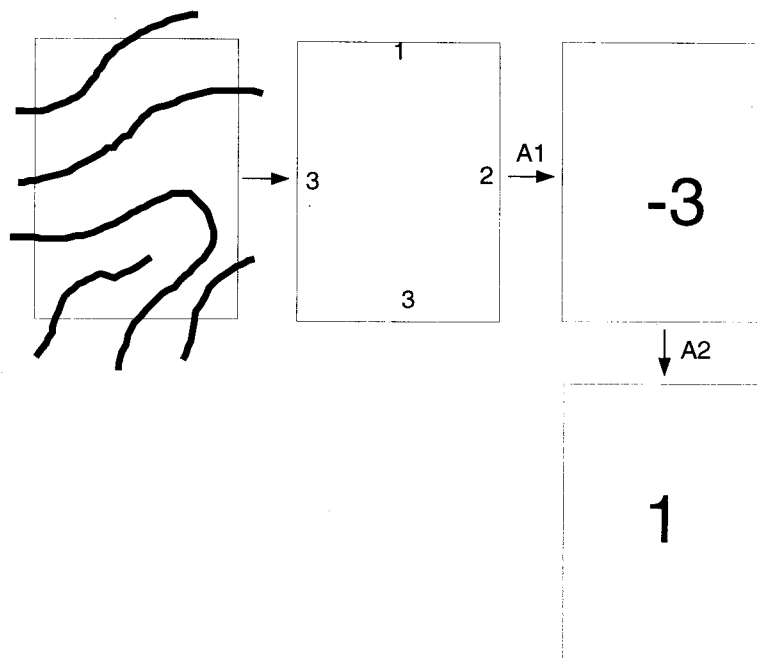


Figure 10: Algorithm A; $M=(1-3)+(2-3)=-3$; $P(-3)=1$

3. **Open question:** It is unclear *a priori* whether the codes created by this algorithm for different fingerprints can be reliably distinguished. This question can be answered only by investigating a large sample pool.

8 General Conclusion

In most of the algorithms that have been suggested in this report, the fingerprint image is reduced to a relatively short sequence of integers. This reduces the memory size required by the database. Each algorithm is intended to exploit specific properties and features of the fingerprint that vary from finger to finger, and that can be localized relatively fast using digital techniques, thus also reducing the computational time requirements to a minimum. In each case, the sensitivity of the algorithm to small variations in the image was also discussed, with the aim of reducing the False Rejection Rate, and of increasing the general robustness of the algorithm.

It is important to point out that the suggestions that have been put forward are conceptual in nature. Each model would require extensive development and testing in order to determine its feasibility, and in order to prepare for final implementation. In particular each model would have to be tested on a large number of fingerprint images, taken from a representative sample of the population. These images should preferably be taken using the KSI sensor. The final solution would possibly involve a *combination* of some of the algorithms that have been presented above.

It is hoped that many of the ideas that are contained in this report will prove useful to KSI in its development of a commercial live-scan fingerprint imager.



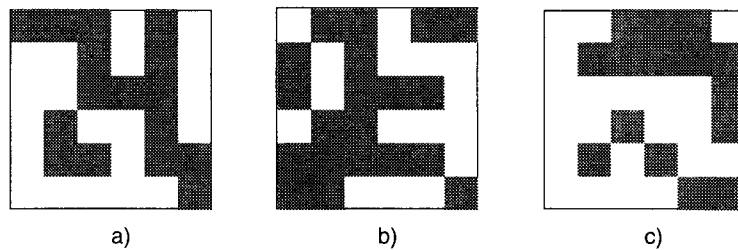


Figure 11: Code matrices (algorithm A2). Odd cells (parity 1) marked. a) Sample 1; b) Sample 1, distorted grid; c) Sample 2.

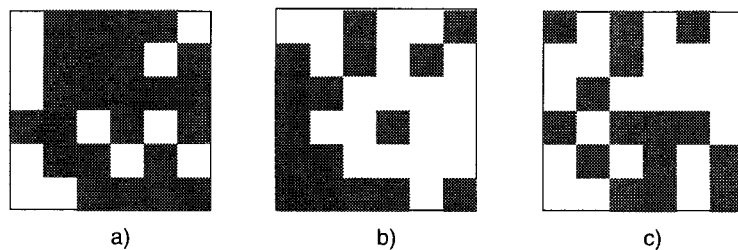


Figure 12: Matching of code matrices. Matching cells marked. a) Fig 11a vs. Fig 11b (same fingerprint): 66% match; b) Fig 11a vs. Fig 11c (different fingerprints): 44% match; c) Fig 11b vs. Fig 11c (different fingerprints); 42% match.



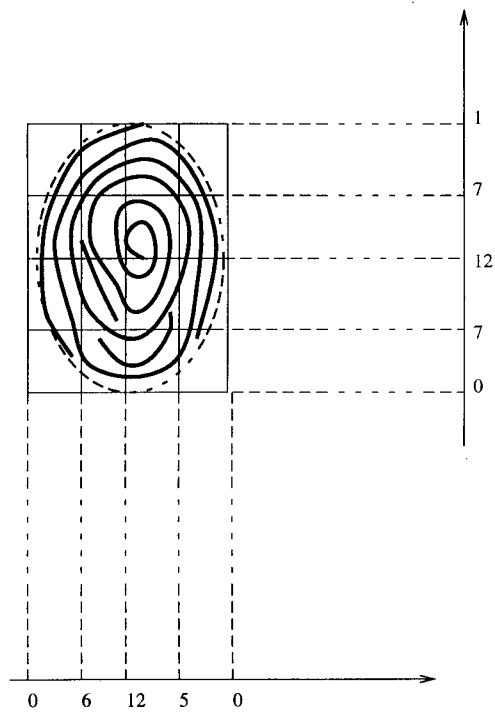


Figure 13: Code sequences (algorithm B).



Chapter 3

Optimal Lumber Production from Softwood Sawlogs

David Baar, Advanced Numerical Methods, davebaar@numericalmethods.com

Lou Hafer, SFU/Comp Sci, lou@cs.sfu.ca

David Kirkpatrick, UBC/Comp Sci, kirk@cs.ubc.ca

Ramesh Krishnamurti, SFU/Comp Sci, ramesh@cs.sfu.ca

Philip D. Loewen, UBC/Math, loew@math.ubc.ca

Michael Monagan, SFU/CECM, monagan@cecm.sfu.ca

Marc Paulhus, Calgary/Math, paulhusm@math.ucalgary.ca

Mihaela Radulescu, UBC/Math, mihaela@math.ubc.ca

written by

Philip Loewen

30 September 1997

1 Problem Statement

This problem deals with the maximization of the “value recovery number” for the processing of raw logs at sawmills in British Columbia. This quantity has the units of dollars per cubic metre, and is defined as the dollar value of usable lumber produced divided by the volume of raw material processed. Sawmills in the interior of British Columbia are of particular interest because the logs processed there are typically of small diameter and the current volume recovered is low, typically 45% to 55%. Each log sawn in a given mill can be given its own value recovery number, and clearly maximizing the aggregated criterion is equivalent to finding out how each individual log can be broken down into boards of the greatest total dollar value.

The first cut running the length of a log is of central importance: it defines a plane one can visualize as being rigidly attached to the log; all subsequent cuts must be either parallel or perpendicular to this plane. Determining the best first cut on each incoming log, which might have nonconvex cross-sections (“cats”), be bent (“swept”), or even have a nonplanar central axis (“corkscrew”), is a complicated optimization problem that must be solved in 3–30 seconds. This is the length of time that elapses as the log moves along a conveyor between a laser-operated surface-measurement station and the first rank of saws the log will encounter during the manufacturing process.

Equipment in the sawmill can shift the front and back ends of the log laterally, and then spin the



log about an axis parallel to the direction in which the conveyor is moving, to provide a prescribed orientation relative to the vertical bandsaw blades that will make the first cut. The lateral positions of these blades can also be varied to control the widths of the boards this first cut produces. Typically there are two or four vertical blades; some mills also have two chipper heads on the outside of the blade assembly. For example, in a mill with two blades and two chippers, a log can be split into three planks with planar vertical sides. Figure 1 shows the end view of a possible first cut in such a mill: the two blades cut the lines separating regions (a) and (b), while the chipper heads reduce the material in region (c) to wood chips. As shown in Figure 1, the “centre cant” (region (a)), is typically wider than the “flitches” (region (b)). One reason for this is that the chain used in the log transport unit limits the separation of the two innermost blades to at most two or three inches, depending on the equipment.

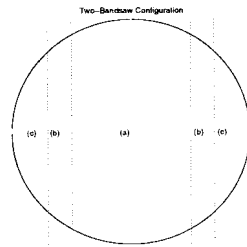


Figure 1: Two-bandsaw configuration.

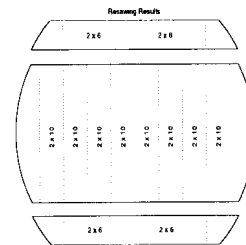


Figure 2: Sample resawing results.

After the first cuts are made, the three timbers in regions (a) and (b) are resawn to produce lumber of standard dimensions. All three are laid down on one of their cut surfaces and passed through ranks of either bandsaws or circular saws. For the flitches labelled (b), dimension lumber lying on its wide face results; for the centre cant (a), resawing produces dimension lumber standing on end. Figure 2 shows a possible breakdown of the boards cut from the log in Figure 1. (Note that the flitches shown at the top and bottom of Figure 2 and the centre cant between them are resawn on quite different pieces of machinery in an actual mill; also, there is no assertion here that the breakdown in Figure 2 is optimal.)

The dollar values of lumber vary according to market conditions, grade, and dimension. For the purposes of this discussion, we considered only the most important of these factors, namely, the correlation between the dimensions of a board and its wholesale value. J.E. Aune of MacMillan Bloedel supplied the typical values in Table 1 below. This table gives the price per cubic metre, calculated as the product of the base price with the length factor. Thus, a load of $1 \times 10''$ boards all 16 feet long is worth \$170 per cubic metre, whereas a load of 1×10 's all 10 feet long is worth only $\$170 \times 0.70 = \119 per cubic metre. Notice that the length factors in this table are not simply increasing (multiples of 8 being especially valuable), and that differences in the length factors can reverse the trend suggested by the base prices. For example, 16-foot lengths of 2×6 are worth only $\$190 \times 0.80 = \152 per cubic metre, but ripping those boards in half lengthwise to produce 16-foot 2×3 's yields a value of \$160 per cubic metre.

The problem first stated at the Workshop was to suggest efficient ways to maximize the value recovery number defined above for each given log. We were given very few constraints within which to work, the most important being that every cut had to run the length of the log, and that all cuts



| Nominal Size | Base Price (\$/m ³) | Length Factors | | | | | | |
|--------------|---------------------------------|----------------|------|------|------|------|------|------|
| | | 8' | 10' | 12' | 14' | 16' | 18' | 20' |
| 1 × 3" | \$100 | 0.60 | 0.70 | 0.80 | 0.90 | 1.00 | 0.90 | 0.85 |
| 1 × 4" | \$120 | 0.60 | 0.70 | 0.80 | 0.90 | 1.00 | 0.90 | 0.85 |
| 1 × 6" | \$150 | 0.60 | 0.70 | 0.80 | 0.90 | 1.00 | 0.90 | 0.85 |
| 1 × 8" | \$150 | 0.60 | 0.70 | 0.80 | 0.90 | 1.00 | 0.90 | 0.85 |
| 1 × 10" | \$170 | 0.60 | 0.70 | 0.80 | 0.90 | 1.00 | 0.90 | 0.85 |
| 2 × 3" | \$160 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| 2 × 4" | \$180 | 1.00 | 0.70 | 0.80 | 0.90 | 1.00 | 0.90 | 1.00 |
| 2 × 6" | \$190 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 1.00 | 1.00 |
| 2 × 8" | \$210 | 1.00 | 0.90 | 0.80 | 0.90 | 1.00 | 0.90 | 1.00 |
| 2 × 10" | \$240 | 0.90 | 0.80 | 1.00 | 1.00 | 0.90 | 0.90 | 0.80 |

Table 1: Typical Lumber Value = Base Price × Length Factor

after the first had to run either parallel to or perpendicular to the plane of the first cut. (The written specifications of the problem do not place restrictions on the number of parallel cuts that can be made in a single pass, although practical limits were discussed in meetings between team members and MacMillan-Bloedel's representative J. E. Aune.) We were allowed to move the log laterally at either end, and rotate it about an axis parallel to the conveyer. An improvement of 1% in value recovered would be valuable to MacMillan Bloedel. Later, we learned that MacMillan Bloedel were primarily interested in finding efficient ways to determine the best first cut. The sequence here is immaterial: *both problems amount to the same thing*. It is clear that regardless of what the first cut may be, once it has been made, the products that result should be resawn in the best possible way in order to maximize profit. So if the goal is to make an intelligent choice of the first cut, the preference of one choice over another can only be determined by optimizing the products resulting from both choices, and associating their payoff values with the corresponding initial decisions. This is why there is no sure way to determine the best first cut without going through some procedure that determines the entire bundle of products that subsequent steps will ultimately associate with the first cut.

If we relax the problem from "find the best first cut" to "find a very good first cut", a second interpretation becomes possible. Experience may well lead to effective "rules of thumb" that often result in first cuts that are very close to optimal, despite being comparatively cheap to implement. In order to gain this experience, of course, one must completely solve the log breakdown problem in many individual cases. But, if we assume that this has been done and that the results are known, there may be some way to make intelligent use of this accumulated wisdom to save computation time, produce better results than the necessarily approximate optimization methods now employed, or even do both.

Here are succinct statements of our two problems.

Problem 1. *Given a log of known dimensions, determine the best location for the opening cut. As a byproduct, completely describe all saleable products that can be produced following this first cut.*

Problem 2. *Given extensive log-optimizing experience, propose new methods to determine a nearly-optimal first cut, without necessarily generating any more information about subsequent stages of processing.*

Problem 1 is a mathematical optimization problem, in which there are three continuous variables (lateral shift of the butt end, lateral shift of the top end, and rotation angle) and a number of discrete variables (number, relative positions, and dimensions of each board in the bundle to be manufactured). Various constraints also come in, due to the number of sawing stations available to process incoming timber, the number of blades at each station, and the physical limits on their relative positions. These features typically vary from one mill to the next, and one expects the optimal breakdown to vary along with them.¹ What is more, any deterministic representation will fail to capture the random errors introduced by the equipment in an actual mill. (For example, Jan Brdiccko says that ordering the rotation of a log by some angle θ about its long axis will trigger a sequence of mechanical events that rotate the log by an actual angle $\hat{\theta}$ that may differ from θ by 10 or 15 degrees.) Thus an implementable solution should take into account the need for robustness with respect to perturbations of the specifications—the truly practical problem might be to maximize the *average* value recovery under plant uncertainty, or even to maximize the *worst-case* value recovery over parameter regions near to the values of top shift, translation, and rotation that our methods propose.

For our first foray into this field, we ignored all these complicating factors, and dealt only with the single constraint explicitly stated in the written problem statement provided to all workshop participants—namely, that all cuts must be parallel or perpendicular to the first one made².

1 Input Data

The information on which the optimization described above is to be based comes from a laser scanning station positioned along the conveyor that moves the log toward the saws. If we consider the z -axis to lie along the direction in which the conveyor is moving, the scanner records the (x, y) -coordinates of 36 points around the perimeter of the log every time its z -coordinate increases by some fixed amount. MacMillan Bloedel provided 15 data sets of such coordinates, which we reformatted into a collection of 45 files—3 for each log. In these files, all measurements are in inches. File `x07.dat` contains a $36 \times N$ matrix of x -values in ascii format; files `y07.dat` and `z07.dat` contain $36 \times N$ matrices with the corresponding y and z -values. Taking the k -th column of these three matrices gives the 36 data values for the log in position k : the number of columns (N) varies, depending on the length of the log. The naming conventions are simple: log numbers from 01 to 15 can take the place of 07 in the description above. For each of these logs, the fixed spacing between z -coordinates of successive measurements is 4 inches, so the z -data is particularly simple. (Each column in matrix `znn.dat` is constant, and the values are simply 0, 4, 8, 12, ...) These data sets are available for anonymous ftp, from `ftp.math.ubc.ca`, in directory `/pub/loew`.

The test data provided by MacMillan Bloedel can be visualized in several ways. A three-dimensional mesh plot joining the measured data points helps one to visualize the log as a whole (see Figure 3 below). In each of the data sets provided, this picture looks very much like a truncated right circular cone. In many cases the lateral sides of the cone are very nearly parallel to its axis, so that the truncated cone is virtually a cylinder. In other cases, the taper is pronounced. (J. E.

¹Comparing optimal breakdowns among various mill configurations could thus provide valuable guidance in choosing the configuration for a new or reconditioned sawmill.

²Some of our formulations may be adaptable to accommodate the curved cuts now being proposed in the lumber industry.



Aune writes, “Log scalers talk of taper as the length of log corresponding to a one inch reduction in diameter, i.e., one inch in twelve feet meaning little taper, one inch in six being a lot of taper.”)

Another way to plot the data is to make a two-dimensional plot showing the rings of measurement points superimposed on one another. This view, which amounts to looking straight down the log along the conveyor axis, reveals the effect of taper most clearly. It can also reveal situations where the natural axis of the log does not line up with the conveyor axis. (See Figure 5 below.)

To see the taper, one can plot the distance from each of the measured points on the log to either the conveyor axis or to (some estimate of) the log’s natural axis as a function of distance along the log. (See Figure 4 below.) When the conveyor axis and the natural axis coincide, this provides a series of vertical clusters of 36 radius values: the taper can be estimated by fitting a straight line to this data, but (as J. E. Aune has noted) a parabolic fit is usually better.

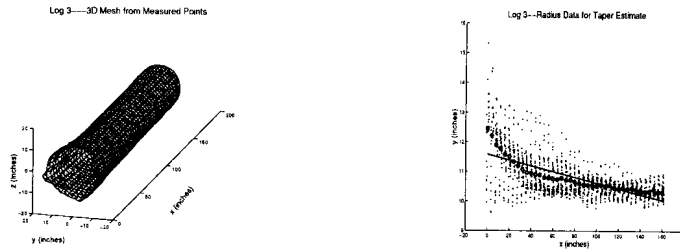


Figure 3: 3D Image of Test Log No. 3. Figure 4: Radius data for taper estimate.

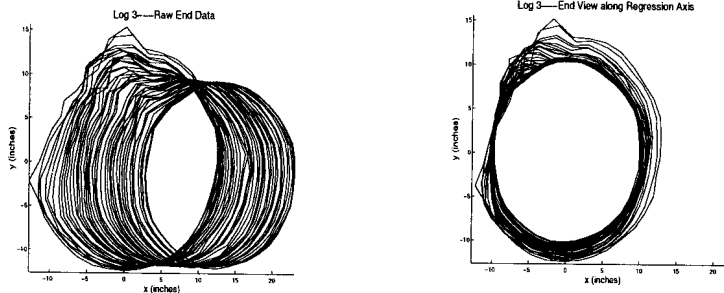


Figure 5: Raw end data.

Figure 6: End view along regression axis.

A summary of physical characteristics of the test logs appears in Table 2. Here, the natural axis of the log was computed using linear regression through the set of all input data points, the taper was estimated as above, and the column labelled “top shift” gives the minimum distance the log’s top needs to be moved to make the log’s natural axis parallel to the conveyor axis. The table shows that log number 12 tapers very little, while log number 13 tapers a lot. Most of the logs landed on the conveyor within a fraction of a degree of the conveyor axis, but perfect alignment could still entail shifting one end by a noticeable amount; the alignment is worst for log number 3, where a shift of over 13 inches seems to be indicated. We emphasize that this is actually the relative shift between the top and bottom ends of the log: in a working mill, it may also be necessary to shift the whole



| Log Number | Length L (inches) | Log-to-Conveyor angle θ (degrees) | Top shift (inches) | $\Delta \text{diam}/\Delta z = 1/x$, where $x = \dots$ |
|------------|---------------------|--|--------------------|---|
| 1 | 168 | 0.96 | 2.8 | 42 |
| 2 | 248 | 0.47 | 2.0 | 229 |
| 3 | 160 | 4.67 | 13.1 | 52 |
| 4 | 160 | 2.77 | 7.7 | 98 |
| 5 | 160 | 1.59 | 4.4 | 108 |
| 6 | 160 | 1.11 | 3.1 | 78 |
| 7 | 160 | 0.40 | 1.1 | 312 |
| 8 | 160 | 0.44 | 1.2 | 198 |
| 9 | 160 | 0.96 | 2.7 | 37 |
| 10 | 160 | 1.96 | 5.5 | 56 |
| 11 | 160 | 0.82 | 2.3 | 45 |
| 12 | 156 | 1.86 | 5.1 | 645 |
| 13 | 160 | 1.26 | 3.5 | 36 |
| 14 | 124 | 0.81 | 1.8 | 145 |
| 15 | 1100 | 0.12 | 2.3 | 104 |

Table 2: Physical Characteristics of Test Logs

log parallel to the cutting axis so that it is roughly centred as it approaches the headrig. For the purposes of optimal breakdown, however, we can ignore this rigid translation simply by assuming that there are no limits on the lateral positions of the saws.

(The top shift amount equals $L \tan \theta$, where L is the length of the log's projection onto the conveyor axis and θ is the angle between the log's natural axis and the conveyor axis, as shown in Table 2. It is included mainly to convey a sense of scale, since the required motion may not lie in either a horizontal or a vertical plane, and when it has been completed, the log's natural axis may be some (parallel) distance away from the axis of the conveyor.)

Figures 3–5 illustrate the various properties of test log number 3. This log has the highest value we found for the angle between the natural axis and the conveyor axis (and consequently for the top shift amount), and tapers more than one inch in five feet.

2 Mathematical Optimization

With regard to the full optimization problem of most efficient log breakdown, we have progress to report on several fronts: selecting the cutting axis, breakdown for cylindrical logs, and cutting planks from flitches. Although a week was not enough time to produce a full three-dimensional optimization package, we are confident that the analysis and ultimate solution of these subproblems will provide insight and perhaps even auxiliary code that will be useful in the general value-recovery problem.

1 Axis Selection by Standard Means

The regression axis used to investigate the given logs in the previous section is optimal in one sense: it is the line for which the sum of all squared distances to the data points (measured in planes perpendicular to the conveyor axis) is the least possible. It clearly gives a better approximation to



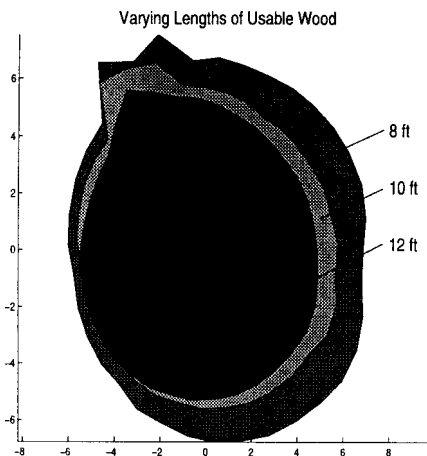


Figure 7: Computing the volume of usable wood.

a reasonable cutting axis than the conveyor axis does, and has the advantage of being very easy to compute. An alternative is to choose a line that minimizes the sum of squared distances to the data points (measured not in planes, but in \mathbb{R}^3): this can be determined using the Singular Value Decomposition at somewhat greater computational expense. Our experience on the test data described above was that the SVD approach gave axes that were indistinguishable from the regression lines for practical purposes, so we will not discuss SVD in detail here. However, for logs with substantial curvature, the SVD may be an appropriate tool to identify the natural plane in which the log lies. (The results for logs in our data set were inconclusive, since they exhibit very little “sweep”.)

2 Axis Selection based on Maximum Usable Volume

The standard mathematical techniques just described are optimization-based, but they involve criteria that are geometric, and not easily related to the real objective of maximum value (or volume) recovery. Another way to choose a cutting axis without going through a complete breakdown optimization is to choose the axis along which you find the maximum volume of “usable wood”. By usable wood, we mean wood of allowable lengths from Table 1, i.e., 8', 10', ..., 20'.

To find this, postulate a cutting axis and have the computer construct a contour map of the butt end of the log by shading all areas that go 20' down the log, 18' down the log, 16' down the log, and so on, down to 8'. See Figure 7. Multiply the shaded areas by their associated lengths to find the volume of usable wood associated with the given cutting axis. Use a continuous optimization routine to choose the cutting axis for which this volume is maximized.

A randomized optimization approach would be easy to implement. Given a proposed cutting axis, Marc Paulhus proposes that one choose at random a pivot point and a direction to move the end of the log, and test if this perturbation improves the usable volume. If so, then accept it; otherwise, keep the original orientation. Repeat this process until random choices of a given size scale no longer help. At this point reduce the size of random perturbations and continue.

In Paulhus's experience, this “tweaking process” will quickly find the maximum weighted contour



map that corresponds to the cutting axis which provides the greatest potentially usable volume. Even if this is not the axis you wish to cut, it might be a good starting point for a search.

3 Vertical Slices for Cylindrical Logs

When we looked at the wire frame model of the first log from the test data MacMillan Bloedel supplied (compare Figure 3), we were struck by the log’s almost perfect cylindrical shape. We had been discussing an algorithm for determining the axis to cut along by finding the largest cylinder that fits entirely inside a given log, expecting that we would later have to do better even than this. When we realized that our suite of test data included a number of logs that are very nearly cylindrical, we put this discussion aside in order to look directly at the problem of breaking down a cylindrical log. We wanted to determine whether the “centre cant plus side fitches” approach reviewed by MacMillan Bloedel was the best sawing strategy after all. The “cant” algorithm is described below.

1. Find the largest width w in the list of usable dimensions (see Table 1) that fits horizontally into the log. If the log radius is r and the prescribed width is w , one can make a rectangular plank of any thickness up to $t^* = \sqrt{4r^2 - w^2}$ down the length of the log.
2. Stack as many planks of the chosen width w as possible vertically into the log. Since planks come in thicknesses of 1 and 2 inches, we can pack any integer height. Thus the thickness actually to be used should be

$$t = \left\lfloor \sqrt{4r^2 - w^2} \right\rfloor.$$

This thickness will be cut into two-inch thick planks in preference to one-inch thick ones as the former are (always?) more valuable. So there will be $t/2$ two-inch planks if t is even, and $(t - 1)/2$ two-inch planks plus one one-inch plank if t is odd.

These two steps determine the width and thickness of the centre cant. MacMillan Bloedel argued that this should be a good approach because it tries to extract as many of the most valuable planks from the log as possible. Jan Aune also pointed out two optimization-based refinements to this method:

1. The number of planks in the cant should be optimized over different rotations of the log. This is irrelevant for a cylindrical log (or, more generally, for a truncated cone), but will often be important for a log of irregular shape.
2. Since the width w and thickness t of the centre cant must be chosen from a short list of marketable values, there will often be freedom to shift the cant to the left, to the right, up, or down inside the log. For example, moving the cant slightly to the left may permit a thicker plank to fit on the right—or vice versa.

Macmillan Bloedel stated that they did not know whether this centre cant approach was the best in general or not. We believe that it is not, and describe a better general approach below.

An obvious problem with the cant approach is that the four pieces which are left over are often not efficiently utilized. We thought that it might be better to split the log vertically into slices, preferably two inch thick slices, and then optimally pack each vertical slice. The intuition is that this approach will utilize the top and bottom of the log better than the cant approach because more often it will be filled with the thicker planks. Moreover, this approach still yields a simple cutting



solution—there is one vertical cutting pass required for the log followed by one horizontal cutting pass for each vertical slice.

To test our hypothesis that the vertical slice approach is better than the cant approach, we constructed the following experiment. We implemented a program which computes a packing for the cant approach as described above where we allow it to add any number of planks to the left of the cant (and symmetrically to the right), and above the cant (and symmetrically below it) to fill up the remaining space. Then we implemented the vertical slice method with three variations:

- choose one 2" wide plank centered in the middle of the log,
- choose one 1" wide plank centered in the middle of the log, and
- choose two 2" wide planks separated by the centre line of the log.

We computed the best packing for each and take the maximum of the three.

We used the sample market price data provided by MacMillan Bloedel to determine which was the optimum decomposition of the vertical planks.

Using only cylindrical logs, we applied both algorithms to a series of 101 logs of fixed length whose diameters ran from 10 inches up to 20 inches in 0.1-inch increments and took the average improvement in the wholesale value of the products introduced by our approach over the cant method. We repeated this experiment seven times, once for a log of each of the standard lengths. We did not make any cuts perpendicular to the long axis of the log (such as cutting a 20-foot board into a pair of 10-foot boards).

Some of the cutting schemes resulting from our implementation appear in Figures 8 and 9. The experimental results are summarized in Table 3. This table shows seven standard lengths, and the average improvement in value recovery for the slice algorithm over the cant approach. In most of the individual trials contributing to this average statistic, the slice algorithm gave a packing with the better dollar value. The cant approach did come out ahead in a reasonable number of trials, however, and occasionally the two methods tied. Obviously, one could try both methods and use the one with the greatest value.

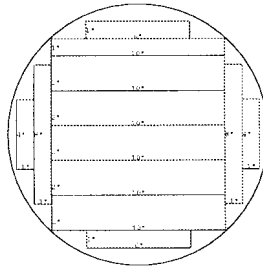
| Log/Lumber Length (ft) | Average Gain in Value using Vertical Slices (%) | Vertical Slicing over Cant Method | | |
|------------------------|---|-----------------------------------|----------|--------|
| | | # Wins | # Losses | # Ties |
| 8 | 10.8 | 97 | 0 | 4 |
| 10 | 2.8 | 69 | 29 | 3 |
| 12 | 0.6 | 52 | 46 | 3 |
| 14 | 5.1 | 77 | 20 | 4 |
| 16 | 2.1 | 67 | 23 | 11 |
| 18 | 2.0 | 64 | 34 | 3 |
| 20 | 11.8 | 97 | 1 | 3 |

Table 3: Average Improvement using Vertical Slices

The reason for the high variance is due in part to the prices. The prices in Table 1 show that for 8' and 20' planks, the value of the 2" thick timber compared to the 1" thick timber is higher than that for the other lengths. If one compares the packings visually one can see clearly what is going on, namely that the vertical slice algorithm packs in more of the thicker lumber. Hence the higher prices for 8' and 20' lengths yield the highest improvements in Table 3.

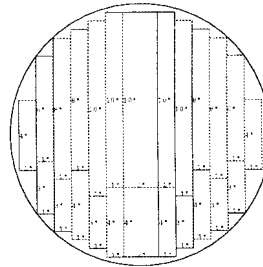


Diameter 15" gives 82.6% yield of 20.28 cubic feet for \$98.44



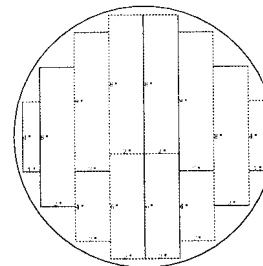
(a) Value-Optimized Cant Breakdown

Diameter 15" gives 87.1% yield of 21.39 cubic feet for \$81.22



(b) Volume-Optimized Slice Breakdown

Diameter 15" gives 81.5% yield of 20 cubic feet for \$111.8

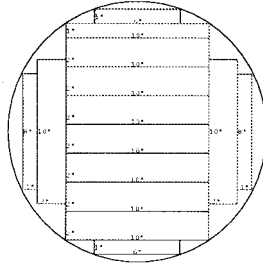


(c) Value-Optimized Slice Breakdown

Figure 8: Possible Breakdowns of a Cylinder with Diameter 15", Length 20'

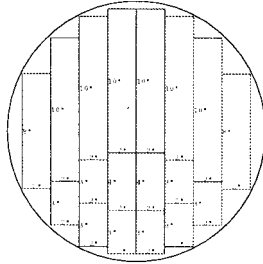


Diameter 18.1" gives 84.7% yield of 15.14 cubic feet for \$76.10



(a) Value-Optimized Cant Breakdown

Diameter 18.1" gives 83.9% yield of 15 cubic feet for \$73.63



(b) Value-Optimized Slice Breakdown

Figure 9: Possible Breakdowns of a Cylinder with Diameter 18.1", Length 10'

We have not yet implemented the cant-placement optimization suggested by MacMillan Bloedel. Of course, this same idea should be applied to both algorithms. But, because the vertical slice algorithm packs the top and bottom better already, it will not benefit from this refinement as much as the cant algorithm. Thus the percentage improvements reported here are somewhat higher than we would expect in practice. Other optimizations may further narrow the gap.

Before we attempted to implement this optimization we found out from Jan Brdicko of MPM Engineering that the vertical slice approach, called “live sawing” in the industry, is already known, and indeed known to be more efficient in general. “So why isn’t it used?” we asked. We were informed that in some (very few?) mills, it is in fact used, but because most(?) saw mills use only 4 vertical blades (some have 6) it cannot be done in a single pass. However, if the vertical slice algorithm is significantly better, then one could imagine making two passes. Given this very positive data, MacMillan Bloedel might consider a study to implement and compare a vertical slice algorithm with the cant algorithm (with optimizations included) on real data to determine how much better it would be in practice, and whether it would be worthwhile to reconfigure any sawmills to accommodate the vertical slice approach.

4 Optimized Fitch Sawing

To make efficient use of the sideboards labelled (b) in Figure 1, it is necessary to know how to get the maximum possible amount of lumber out of a one- or two-inch thick plank with irregular edges. This is a stand-alone sawmill optimization problem of independent interest, but it also plays a vital role in the full problem, since a complete analysis of the full-log problem requires that the optimal value of each cutting product be worked into the calculations. Also, the pieces of timber produced by the live-sawing method just discussed are mathematically equivalent to fitches. Finally, the fitch-sawing problem is a natural place to start for mathematical reasons: it is (approximately) two-dimensional instead of three, so it should be easier to solve than the full problem; furthermore, ideas arising in the 2D case could possibly turn out to remain useful in the general case.

We imagine the sideboard lying on one of its cut faces, and assume for simplicity that all of its edges are vertical. Then the measurement points of the edges of the sideboard provide the vertices of a polygon F in the (x, y) -plane, from which we desire to cut rectangular subsets whose total dollar value is a maximum. The long sides of all these rectangles must be parallel. We have partial results for two approaches, both of which apply to an arbitrary plane polygon F (although we expect F to be long and narrow): first, a method for maximizing the length of a board with given width lying inside F ; second, a simple way to detect if a board of given width fits inside F .

(a) Maximum Length for Given Width

Given the vertices of closed polygon (“fitch”) F in \mathbb{R}^2 whose sides do not intersect, one can easily compute the intersection between F and a horizontal strip $S(a, b) = \{(x, y) : a \leq y \leq b\}$, where a and b are prescribed saw positions and $w = b - a$ is one of the widths of standard lumber. The idea is simply to work around the perimeter of the fitch F , taking note of the points where line segments between successive vertices (measured data points) cross the top or bottom of the strip. The polygon of intersection, corresponding to the shape of the board that would remain after sawing along the lines $y = a$ and $y = b$, has for vertices the collection of all the crossing points together with all the vertices lying inside the strip. With this description for the set $C = F \cap S(a, b)$ in hand, finding the maximum number of rectangles (“planks”) in the set C and their lengths is also straightforward. To illustrate, Figure 10(a) shows a 12-foot long fitch and two parallel lines separated by 4 inches. In Figure 10(b), the computation of yield for this scheme has been determined by the method above.



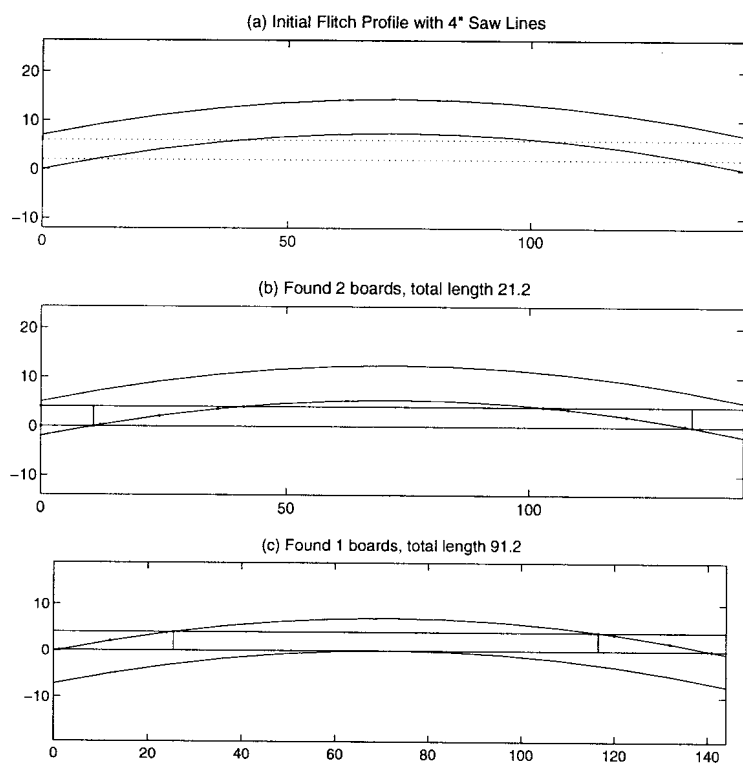


Figure 10: Steps in maximizing length for given width.



When these capabilities are combined with a method for shifting and rotating the flitch F into an arbitrary location relative to the saw lines $y = a$ and $y = b$, one has a method for evaluating the possible number and length of boards of a specific cross-section associated with a particular cutting scheme. Optimizing the value of such a solution over the possible shifts and rotations of the given flitch can be done by standard subroutines. Figure 10(c) illustrates the outcome for this example.

For wide flitches, it may be possible to extract more than one usable board. Our methods for finding the length of boards compatible with a given pair of cut lines extend easily to finding the lengths of a bundle of boards lying side-by-side, in various positions down the flitch. We would plan to use some discrete optimization engine to generate sequences of board widths for processing by this length-optimization algorithm, and do a rapid search for the best possible collection of widths (with its corresponding orientation).

One aspect of this method requiring further study is the objective function in the optimization just demonstrated. So far, we just maximize the total length of all boards produced. Thus, our implementation assigns the same value to one twelve-foot board as it does to four three-foot boards. This is obviously unrealistic, but in order to get reasonable results from any continuous optimization package, one needs an objective function whose gradient (exists and) is nonzero at the vast majority of points. The true objective function—valued in dollars—is constant in the whole range of lengths from 8 to 10, then constant again from 10 to 12, and so on. Further study would be required to find an objective function that captures the essence of the problem’s discreteness while having no flat spots.

(b) Discrete Search Plus Oracle

Reversing the design in the previous subsection holds considerable potential. Under this scheme, the discrete optimization routine would propose not just a list of widths, but the full shape of a “template” of lumber, including lengths and relative positions of the finished boards. Then a continuous subprogram (“oracle”) would be used to determine if the current template can actually be cut, and if so, to give the flitch orientation relative to the blades that makes it happen. The discrete driver would take the usual branch-and-bound or knapsack approach, although the details of how to structure the search space will take some careful thinking to work out. We can, however, describe an oracle that detects the viability of a given template.

We outline the special case in which the template to be realized is a single plank, 6 inches wide. Deciding whether or not this fits into a given flitch is more complicated than simply asking if the flitch is at least six inches wide everywhere along its length—the flitch in Figure 10 is seven inches across everywhere, and we did not even get eight feet of 4-inch wide lumber! The key concept here is convexity: a simple plank can only fail to be realizable if one of the vertices of the given polygon lies inside it. So if we can lay the bottom edge of the plank along a line through the flitch that lies above all the lower vertices and at least 6 inches away from all the upper vertices, the plank will fit inside the flitch. The line we want must therefore separate the set of lower vertices from a system of disks with radius 6 centred on the upper vertices. And a line has this separation property if and only if it separates the upper convex envelope of the flitch’s bottom vertices from the lower convex envelope of the disk-augmented upper vertices. Figures 11 and 12 illustrate this: both involve computer-generated flitches with considerable taper and exaggerated random variations in upper and lower surfaces. Both figures show (i) the flitch profile itself; (ii) the upper vertices enlarged into disks of radius 6”; and (iii) the convex envelopes just mentioned. In Figure 11 the convex envelopes cross, so it is impossible to cut a 6” plank the length of the flitch; in Figure 12 the convex envelopes are separated, and a possible pair of cut lines to produce a 6” board is shown. (In the interests of error rejection, we position the cut lines to divide any excess width equally between the top and bottom



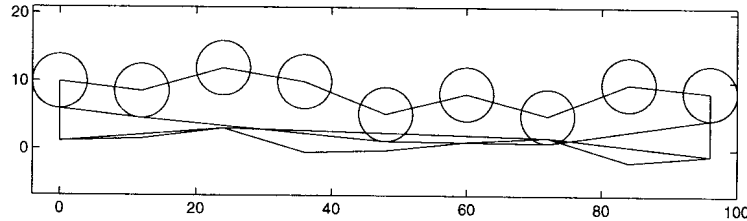


Figure 11: Convex chains overlap, so plank won't fit.

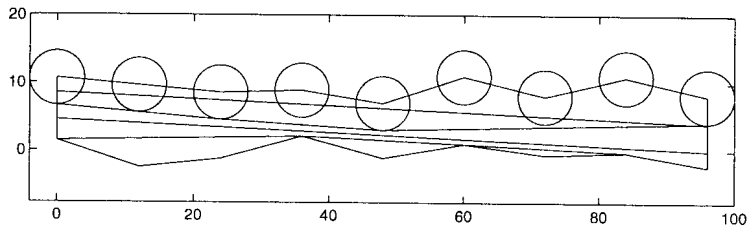


Figure 12: Successful separation and plank cutting.

edges, and to choose the midpoint of any nondegenerate interval in which the side slope of the cut board is allowed to lie.)

This method can be extended to determine the feasibility of templates more complicated than just a single plank. To accommodate a shorter board of width 4" on top of the left end of the plank, it suffices to know the length of this board, and to increase by 4" the radii of the circles in the corresponding portion of the upper flitch boundary. For this more general template, we cannot guarantee feasibility of any template that avoids the flitch vertices—one can imagine a straight-line segment of the boundary that cuts off the corner of the shorter board just mentioned—but we expect the slope of the separating line we find to be so small in practice that if we agree always to position both ends of any shorter boards directly above or below a vertex of the flitch, this problem will never become a practical issue. Indeed, an even coarser approximation may be adequate in practice: given enough data points, we may be able to enlarge the upper vertices not into disks, but simply into vertical segments, or to use segments at vertices near the flat parts of the template and disks near its nonconvex corners. (This will save some effort in the computation of the upper convex envelope for the augmented top edge, since each disk can in principle contribute a circular arc, whereas each segment will contribute just its lowest point.)

(c) Other Directions

The problem of finding how far along a given line segment one can travel before hitting the boundary of a given polygon in the plane or polyhedron in three-space is one that computer graphics experts need to solve in order to predict the shading of computer-generated solids. This is exactly the same problem we face, either in measuring distances inside the log from the butt end toward the top, or in seeking dimension lumber inside a given flitch. In the computer graphics community, the desired



procedure is called “ray tracing.” It would be instructive and likely very useful to explore the ray tracing literature and assess its possible impact on log optimization technology.

3 Rules of Thumb

So far we have dealt with quantitative approaches that could be used as components of a full cutting optimization system. If, instead, the goal is to make an educated (hopefully near-optimal) guess at a reasonable first cut without solving the full breakdown problem in advance, several ad-hoc procedures come to mind.

1 Different Algorithms in Parallel

In the absence of a perfect algorithm a selection of good and essentially different algorithms can help. If two algorithms exist, called A and B , and both are applied to every incoming log, one can choose the one producing the higher value for each instance. Even if algorithm A produces the better recovery 99 times for every 100 logs, the presence of algorithm B has made a contribution.

2 Hybridized Lookup and Optimization

As computers become cheaper and more powerful, there has been a move away from lookup tables and towards real-time optimization for calculating log breakdown strategies. A hybrid approach may be worth considering, however. Here the idea would be to single out a small list of essential log parameters that can be computed using the measurement information, and to use this list as a key into some large database of successful cutting patterns used in the past. In a pure lookup scheme, the database would provide the cutting axis directly; in a hybrid approach, the database would give intervals likely to contain the optimizing values of the continuous variables, and conventional optimization software would be instructed to restrict its search to these regions. This would either reduce execution time, or use the time available more efficiently, to make a more thorough exploration of the set of most interesting parameter values.

3 A Learning Approach

Another idea is to build a database that will become more intelligent the more it is used. Again, use essential log parameters to determine the “type” of each incoming log. Consult the database to find all of the cutting patterns which have applied to this type before. Apply them all to the log (under different orientations) and choose the best result, then cut the log according to this scheme. If the value recovery number that results falls below some preassigned efficiency level (which may depend on the log’s type), send the image of the log to an offline computer that spends as much time as necessary to find what the best way to cut this log would have been. Then add this cutting scheme to the database.

The latter two approaches incorporate the notion of “essential parameters” that capture features of the log that are important when selecting a good cutting pattern and orientation. There are obvious suspects: the length, taper, and sweep of the whole log come immediately to mind, along with the sizes, eccentricities and principal axes orientations of the cross sections; other quantitative measures can also be imagined. (Characterizing the cross sections using the first few Fourier coefficients of their (2π -periodic) radius functions, for example, might provide an easy way to detect “cat face”.) To identify correlations between the value recovery number and the parameters thought to be



essential, we propose a statistical search³ through idealized cutting reports. These would be based on the raw log data gathered by a particular mill in real operation, but with the actual breakdown replaced by a truly optimal cutting pattern generated by exhaustive offline search. If the observed correlations can capture the variation in optimal yield, then we can have some confidence that the parameters singled out as “essential” are sufficient. If not, and unexplained variation remains, it would indicate that some important property of the log has been overlooked; a natural response would be to look for it and repeat the statistical analysis described above. (Offline computation of the “idealized” cutting reports mentioned above may represent a significant computational challenge, but the payoff would be a database containing many patterns that improve on the best current practice. Running the statistical search on actual cutting schemes used in the real mill could also be done, but it would produce recommendations that are no better than what is done already . . . although they may be produced faster than they are by full-scale optimization.)

Until the statistical analysis described above is done, we cannot be sure that a correlation between yield and a small number of log shape parameters will exist. If there turns out to be no correlation, we will be forced to conclude that there is no way to consistently produce good cutting patterns based on a small number of log parameters, and encouraged to pursue further the idea of efficient full-scale optimization in real time.

4 Acknowledgements

The authors thank Jan Aune of MacMillan Bloedel for proposing this problem, and for continued email correspondence during the weeks following the Industrial Problem Solving Workshop. Credit is also due to Jan Brdicko of MPM Engineering for coming to visit the group during the Workshop and providing useful information about the state of the log optimization industry.

5 Internet References

- Log data by anonymous ftp: [ftp.math.ubc.ca/pub/loew](ftp://ftp.math.ubc.ca/pub/loew)
- Email address for Jan Aune: je.aune@mb ltd.com
- Web page for MPM Engineering: <http://www.mpm-eng.com/>
- Web pages on Lumber Recovery from *Logging and Sawmilling Journal*:
<http://www.forestnet.com/log&saw/lrf/lrf.htm>
- Web page for Nanoose Systems of Parksville (another maker of log optimization software), showing nice graphics of the optimization process:
<http://www.nanoose.com/>
- Web page for Inovec Optimization and Control Systems of Eugene, Oregon (another maker of log optimization software):
<http://www.inovec.com/index.shtml>

³The method of “Kriging”, known to geophysicists, statisticians, and others who need to interpolate large data sets, may be of use here.



Chapter 4

Stress intensity in a thermoroll

D. Beder¹, T. Myers², J. Ockendon³, A. Peirce⁴, B. Van Fliert⁵, M. Ward⁶

written by Michael J. Ward

1 Introduction

This report describes the mathematical results obtained by a team of researchers working at the 1997 PIMSIPS Workshop investigating the stress buildup and temperature profiles in a thermoroll. The problem under consideration was brought to the workshop by Dr. Roman Popil of MacMillan Bloedel Research representing MacMillan Bloedel Ltd. The problem description provided, including the background, questions, and data are given below in §1, §2 and Appendix A.

The outline of the report is as follows. In §3 we give some physical estimates. In §4 and §5 we model the temperature field in the thermoroll in two different regions of the roll and we calculate the corresponding temperature gradient. In §6 we estimate the stress induced by the temperature gradients in order to determine the location of the maximum stress and to determine any possible singular behavior of the stress field. Such a singular behavior could lead to the formation of cracks. Finally, in §7 we state our conclusions.

¹Dept. of Physics, UBC, Vancouver

²Dept. of Mathematics, Cranfield Institute, U.K.

³Dept. of Mathematics, Oxford University, U.K.

⁴Dept. of Mathematics, UBC, Vancouver

⁵Dept. of Mathematics, Leiden University, Netherlands

⁶Dept. of Mathematics, UBC, Vancouver



2 Background

During the manufacture of coated paper products, a paper-making stock consisting of water and 1% or less wood fibers is prepared by chemically or mechanically separating the fibers from wood. A screening process removes most of the water; the remainder is removed through pressing against felts and contact drying. The web is further densified by passing it through high pressure calender rolls, resulting in about a two-fold decrease in caliper of the pressed and dried paper. The web may then pass through a number of calender nips. The geometry for one such nip is shown in Fig. 1. This last stage of densification involves high temperatures and pressures that lead to high stresses in the roll material.

A stack consists of two rolls: one has a polymeric elastomer covering, the other is a solid iron alloy (the thermoroll). It is our task to estimate the stresses in the thermoroll under standard operating conditions, and determine whether it is possible, under certain conditions, for cracking or roll failure to occur.

3 A Detailed Description

The thermoroll consists of a hollow cylinder, rigidly attached at either end to a rotating bearing. It is hydraulically loaded to 350kN/m, leading to a deflection at its center of 1.6mm. The roll has inner and outer radii of 560 and 750mm, respectively. It contains 45 bore holes with radii 16mm. Oil, heated to a temperature of 253°C, flows through these boreholes at a rate of 33.7 L/s with an energy influx of approximately 800kW. As noted by Dr. Roman Popil, the temperature on the inner bores may be a little different from the oil temperature and is typically an unknown quantity. We label the bore temperature as T_{bore} in the analysis below. The typical configuration is shown in Fig. 1 and 2.

The thermoroll is made by a casting process such that the outer layer, the so-called ‘chill’, having cooled first upon contact with the walls of the mold, imparts a residual radial stress into the roll. The thermal and elastic properties of the chill are different from that of the bulk material.

The web enters the nip at a temperature of 66°C, typically traveling at 18m/sec. Its temperature, measured some distance from the nip, increases by approximately 26°C.

Periodically the polymer covered roll is washed, providing a source for dripping water. This may fall onto the iron roll and subsequently evaporate or ‘jump’ off the hot surface (and so have little effect on heat transfer).

A number of sources for stress build-up in this system are easily identified, for example:

- Large temperature gradients (and consequently thermal stresses) will occur in the vicinity of the nip.



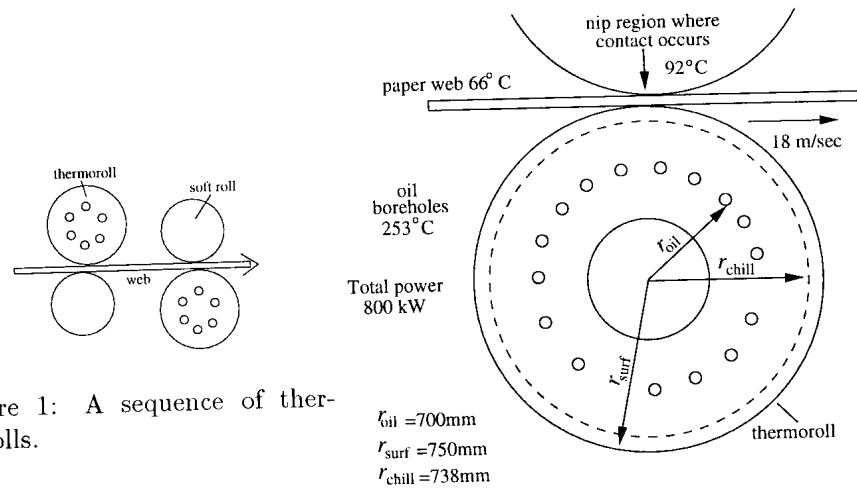


Figure 1: A sequence of thermorolls.

Figure 2: Schematic plot of geometry simplified.

- Cooling by dripping water may have a significant local effect.
- The iron roll manufacturing process will invariably lead to built-in stresses (estimated by the manufacturers to have compressive radial stress components of 150 MPa).
- Another compressive stress field, transmitted through the web, due to the roll's weight and loading.

The main questions for the study group to address are:

- What will be the total stress intensity (residual, thermal, loading ...) induced during this process?
- Are there conceivable operating conditions under which the stresses could become high enough to cause cracking, or even worse for catastrophic failure to occur?

The main focus of our group was to study the second question.

4 Physical Estimates

We first estimate whether the energy input from the oil boreholes and the quoted temperature difference across the roller are consistent. Specifically, we would like to estimate the surface temperature of the roller.



As will be shown in §4, the perturbing effect of the nip on the rapidly rotating drum is small except near the nip. Thus, we can approximate the temperature in the annular iron drum between the oil boreholes and the outer surface as the radial function

$$T(r) = A + B \ln r, \quad r_{oil} < r < r_{surf}. \quad (1)$$

The values for r_{oil} and r_{surf} are given in Appendix A, and are $r_{oil} = 700\text{mm}$ and $r_{surf} = 750\text{mm}$, respectively. The boundary condition $T = T_{bore}$ at $r = r_{oil}$ determines one relation between A and B while the other relation is found from the given total heat flux into the roller from the oil, which was quoted in §2. We estimate

$$\frac{dQ_{in}}{dt} = 2\pi r_{oil} \kappa T_r L_{roll} = P \quad \text{on} \quad r = r_{surf}. \quad (2)$$

Here P is the total heat flux and L_{roll} is the length of the roller. Using the data provided, and taking κ as that for iron, we estimate $B = 398^\circ\text{C}$. This indicates a temperature difference between the borehole radius and the outer surface of

$$\Delta T = B \ln(r_{surf}/r_{oil}) = 28^\circ\text{C}. \quad (3)$$

Using a more refined calculation, taking into account the different values of the thermal diffusivity κ for the chill and the core, we estimate a temperature difference of 32°C . In §5 below, we use the surface temperature of 192°C as measured by MacMillan Bloedel.

Next, we estimate the heat flux and temperature gradient in the roller surface just below the nip. With a specified nip width of 1.1cm and the heat flow to the web, the estimated radial heat flux ϕ is

$$\phi = 50 \frac{\text{kW}}{\text{m}} / 0.011\text{m} = 4.5 \times 10^6 \text{ W/m}^2. \quad (4)$$

Thus, since $\phi = \kappa T_r$ on the roller surface where κ is the thermal diffusivity of the chill, we get the estimated surface temperature gradient

$$T_r = O(10^5) \text{ }^\circ\text{C/m} = O(10^3) \text{ }^\circ\text{C/cm}. \quad (5)$$

This is much larger than the overall, or average, radial temperature gradient of 5°C/cm through the roll radius. Therefore, it is clear that we must do a careful analysis of the temperature gradients and the stress field near the nip. This is done below in §4 – §6.

Finally, we estimate the effect of water droplets. Water may be dripped onto the thermoroll when the paper web is interrupted in order to clean the top roller. The effect of these water droplets on the formation of micro-cracks seems very difficult to determine analytically. The goal would be to first calculate the temperature gradient on the surface of the roller just under the droplet. To do so we would need better observations of droplet size and lifetime. Apparently the manufacturer claims



a droplet lifetime of the order of 1 second whereas MacMillan Bloedel has estimated the “Leidenfrost thermal flux lower limit” as input information.

We now consider a simple scenario. Suppose that we have a hemispherical droplet of radius a , where $a = 6.2\text{mm}$. The required latent heat to evaporate the droplet is $C_{lat} = 2.25610^6\text{J/kg}$. Thus, the estimate of the heat flux ϕ into the droplet across a small patch of roller surface under the droplet is

$$\phi = \frac{2\pi}{3t_{drop}} \left(\frac{a^3 \rho C_{lat}}{\pi a^2} \right) = \frac{2a\rho C_{lat}}{3t_{drop}}. \quad (6)$$

Using the estimates $a = 6.2\text{mm}$ and the droplet lifetime $t_{drop} = 1\text{sec}$, we get an estimate $\phi = 10^7\text{W/m}^2$, which is four times larger than the heat flux estimated from the nip region. It therefore, appears crucial to do a careful analysis of temperature gradients and the resulting stress field near the droplet. This analysis is very difficult and was not done by our group.

5 The Global Temperature Field

In this section we calculate the temperature field in the region away from the nip to determine the effect of the oil bores on the heating of the drum. This is referred to as the **global problem**. In this problem, the nip region is replaced by a point sink of strength Q . The strength of this sink can be determined by the estimate obtained in §3.

It is convenient in the analysis to fix ourselves in a frame in which the thermoroll is stationary and the nip region on the edge of the thermoroll is rotating at an angular velocity $\omega = v/r_{surf}$. The mathematical model for the temperature field T , where T is measured in $^\circ\text{C}$, is

$$T_t = \kappa (T_{rr} + r^{-1}T_r + r^{-2}T_{\theta\theta}), \quad r_{oil} < r < r_{surf}, 0 \leq \theta < 2\pi, \quad (7a)$$

$$T(r_{oil}, \theta, t) = T_{bore}, \quad (7b)$$

$$\kappa_{ch} T_r(r_{surf}, \theta, t) = Q\delta(\theta - \omega t). \quad (7c)$$

Here κ_{ch} is the thermal diffusivity of the chill and δ is the Dirac delta function. The parameter κ in (7a) is piecewise constant, with $\kappa = \kappa_{ch}$ in the chill region $r_{chill} < r < r_{surf}$, and $\kappa = \kappa_{co}$ in the core region $r_{oil} < r < r_{chill}$. The values for these constants are given in Appendix A. As a simplifying approximation to the geometry, in this model we have replaced the individual boreholes by a line of boreholes along $r = r_{oil}$. Although such an approximation should warrant further study, it greatly simplifies the analysis. In this model we have also assumed a negligible heat transfer between the air and the thermoroll. In a more refined analysis than is presented below, a Newtonian cooling boundary condition should be imposed on the edge of the thermoroll.

There are two goals to the analysis below. Firstly, we would like to justify why the temperature can be approximated by a radially symmetric function away from the nip region. Secondly, we would like to estimate the temperature gradient as we approach the nip region.

We introduce non-dimensional variables by $\rho = r/r_{surf}$ and $\tau = t/\omega$. The non-dimensional rotation rate ω_n is defined by

$$\omega_n = \omega r_{surf}^2 / \kappa_{ch}. \quad (8)$$

We estimate that $\omega_n = 10^6$, which is very large. This value measures the importance of rotation as compared to thermal diffusion, and can be thought of as a Peclet number. The non-dimensional model is

$$\omega_n k^{-1} T_\tau = T_{\rho\rho} + \rho^{-1} T_\rho + \rho^{-2} T_{\theta\theta}, \quad \rho_0 < \rho < 1, \quad 0 \leq \theta < 2\pi, \quad (9a)$$

$$T(\rho_0, \theta, t) = T_{bore}, \quad (9b)$$

$$T_\rho(1, \theta, t) = \bar{Q} \delta(\theta - \tau). \quad (9c)$$

Here $\rho_0 \equiv r_{oil}/r_{surf}$, $\bar{Q} = Q r_{surf} / \kappa_{ch}$, while $k = 1$ in the chill region and $k = \kappa_{co} / \kappa_{ch}$ in the core. A schematic plot of the geometry is shown in Fig. 3.

We look for a solution to (9) in the form

$$T(\rho, \theta, t) = F(\rho, \phi) \quad \text{where} \quad \phi = \theta - \tau. \quad (10)$$

The problem for $F(\rho, \phi)$ becomes

$$-\omega_n k^{-1} F_\phi = F_{\rho\rho} + \rho^{-1} F_\rho + \rho^{-2} F_{\phi\phi}, \quad \rho_0 < \rho < 1, \quad 0 \leq \phi < 2\pi \quad (11a)$$

$$F(\rho_0, \phi) = T_{bore} \quad (11b)$$

$$F_\rho(1, \phi) = \bar{Q} \delta(\phi). \quad (11c)$$

Next, we seek a solution to (11) in the rapid rotation limit $\omega_n \gg 1$. In the outer region, defined away from the nip near $\phi = 0$, we substitute the expansion

$$F(\rho, \phi) = F_0(\rho, \phi) + \frac{1}{\omega_n} F_1(\rho, \phi) + \dots, \quad (12)$$

into (11) and collect powers of ω_n to obtain that

$$F_{0\phi} = 0, \quad (13a)$$

$$F_{0\rho\rho} + \rho^{-1} F_{0\rho} = -k^{-1} F_{1\phi}, \quad \rho_0 < \rho < 1. \quad (13b)$$

Since F_1 is periodic in ϕ , it follows that $\int_{-\pi}^{\pi} F_{1\phi} d\phi = 0$. Therefore, upon integrating (13b), we obtain that F_0 satisfies

$$F_0 = A + B \ln \rho. \quad (14)$$



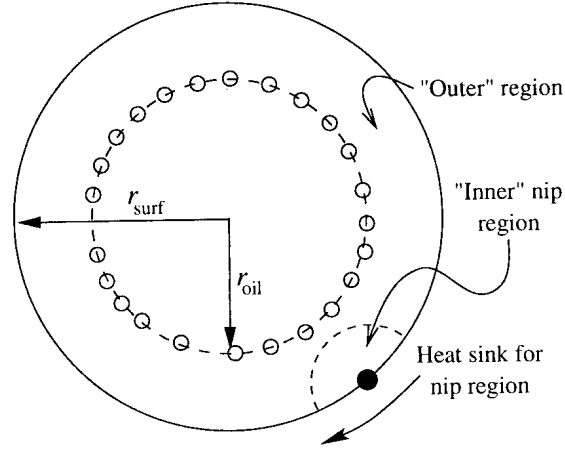


Figure 3: A heat sink is rotating on the boundary of the thermoroll.

One boundary condition for F_0 is $F_0 = T_{bore}$ at $\rho = \rho_0$. There are two possible conditions that can be imposed for the second relation. One choice would be to specify the flux out of $\rho = \rho_0$ as in §3, which would determine B . The second choice would be to satisfy the boundary condition $F_{0,\rho} = 0$. Note that a better approximation to the boundary condition would be to impose a Newtonian cooling condition on the surface of the roller characterized by some Biot number, representing a heat transfer coefficient between the air and the chill. Specifying the flux out of the boreholes would enable us to calculate this coefficient. The temperature field, away from the nip region, is obtained by substituting (14) in (10).

Our first conclusion is that in the limit of rapid rotation the temperature field away from a thin zone near the nip region is radially symmetric.

The solution (14) is not valid in the vicinity of the nip region where $\phi = 0$. Thus, as is usual in singular perturbation problems, we must construct an inner solution near $\phi = 0$, $\rho = 1$. The extent of this region is $O(\omega_n^{-1})$. We introduce the local variables $\hat{\rho}$ and $\hat{\phi}$ by

$$\hat{\rho} = (1 - \rho)\omega_n, \quad \hat{\phi} = \phi\omega_n^{-1}. \quad (15)$$

The corresponding problem for F becomes

$$-F_{\hat{\phi}} = F_{\hat{\rho}\hat{\rho}} + F_{\hat{\phi}\hat{\phi}}, \quad 0 < \hat{\rho} < \infty, \quad -\infty \leq \hat{\phi} < \infty, \quad (16a)$$

$$F_{\hat{\rho}}(0, \hat{\phi}) = \bar{Q}\delta(\hat{\phi}), \quad (16b)$$

$$F \rightarrow 192 \quad \text{as} \quad \hat{\rho} \rightarrow \infty. \quad (16c)$$

We have used the surface temperature of 192°C quoted by MacMillan Bloedel Research as the far field condition. A plot of the geometry is shown in Fig. 4.



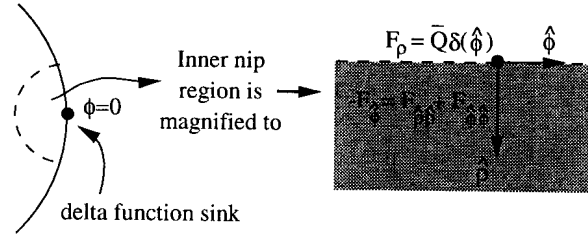


Figure 4:

The solution to this problem is readily found to be

$$F(\hat{\rho}, \hat{\phi}) = -\bar{Q}e^{-\hat{\phi}/2}K_0\left(\frac{1}{2}(\hat{\phi}^2 + \hat{\rho}^2)\right) + 192. \quad (17)$$

Here $K_0(z)$ is the modified Bessel function of the first kind of order zero. In terms of the original variables, we have

$$T(r, \theta, t) = 192 + \bar{Q}e^{-\omega_n(\theta - \omega t)}K_0\left(\frac{\omega_n}{2}\left[(1 - r/r_{surf})^2 + (\theta - \omega t)^2\right]^{1/2}\right). \quad (18)$$

Therefore, using the asymptotic behavior of $K_0(z)$, we observe that as we approach the nip region the change in temperature $\Delta T = T - 192$ behaves logarithmically like

$$\Delta T \sim \frac{\bar{Q}}{2} \ln \left[(1 - r/r_{surf})^2 + (\theta - \omega t)^2 \right], \quad \text{as } r \rightarrow r_{surf}, \quad \theta \rightarrow \omega t. \quad (19)$$

Here \bar{Q} is the strength of the heat sink. Thus, for the global problem the temperature decrease as we approach the nip region is only logarithmic with the distance from the nip. The effect of the resulting temperature gradient on the stress field is estimated in §6.

6 The Local Temperature Field

We now calculate the temperature field in the region near the nip. This is referred to as the **local problem**. In the near-nip region, the geometry is approximately planar as shown in Fig. 5. The contact region between the paper and the roller is approximately 2cm as shown in this figure. The goal is to estimate the temperature gradients near the **leading edge** where the paper first comes into contact with the roller (see the figure). It is in this region that we expect the temperature gradients to be largest. Note that the calculations below are done in a frame for which the leading edge is located at $X = 0$ (see Fig. 5).

From the data given in Appendix A, the values for the thermal diffusivity of the water κ_w , the chill κ_{ch} and the core κ_{co} are

$$\kappa_w = 0018 \text{ cm}^2/\text{sec}, \quad \kappa_{ch} = .071 \text{ cm}^2/\text{sec}, \quad \kappa_{co} = .143 \text{ cm}^2/\text{sec}. \quad (20)$$



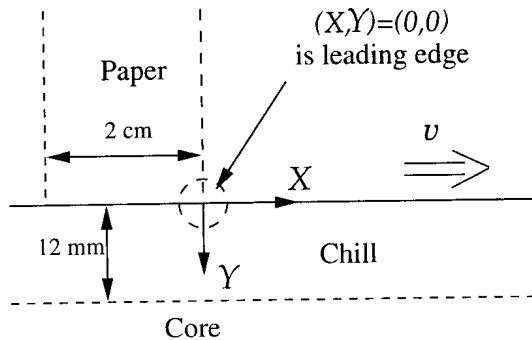


Figure 5: Schematic plot of contact region.

From these values, we can estimate a Peclet number, which is a dimensionless parameter giving a measure of the relative strengths of convection compared to the thermal diffusion. Since $Pe = vL/\kappa$, we can estimate a Peclet number using $v = 18\text{m/sec}$, $L = 2\text{cm}$, and $\kappa = 0.1\text{cm}^2/\text{sec}$ to get $Pe = 18000$, which is very large. Hence the temperature field near the nip is dominated by convection.

We now formulate the model. We approximate the temperature field in the near nip region using a steady-state convection-diffusion equation

$$-vT_X = \kappa(T_{XX} + T_{YY}), \quad (21)$$

where κ is piecewise constant. We then introduce non-dimensional variables by

$$x = XPe/L, \quad y = YPe/L, \quad (22)$$

where we take $L = 2\text{cm}$ and $Pe = vL/\kappa_{ch}$. In terms of these variables, the length of the contact region between the paper and the roller is extremely large and thus, as a good approximation, we take the contact region to be of semi-infinite extent occupying the region $x < 0$, $y = 0$. In terms of these new variables, the chill region extends very deeply below the contact region and hence we will take the chill to occupy the region below the x -axis.

We now formulate the boundary conditions. The line $x > 0$, $y = 0$ is where the roller is exposed to the air and hence we assume that there is negligible heat transfer between these two media (i.e. $T_y = 0$ for $x > 0$, $y = 0$). This assumption should be examined more carefully in a more refined analysis. In addition, we assume that the temperature field and the heat flux are continuous across the contact region.

Next, we give far field conditions for the temperature field. The surface temperature of the roller obtained from the global problem is estimated by MacMillan Bloedel Research to be $T = 192^\circ\text{C}$. This yields the asymptotic matching condition $T \rightarrow 192^\circ\text{C}$ as $y \rightarrow -\infty$. Finally, the web temperature into the nip is 66°C and thus we set $T \rightarrow 66^\circ\text{C}$ as $y \rightarrow +\infty$.



To summarize, the model for the temperature field in the near nip region (see Fig. 6), where T is measured in $^{\circ}\text{C}$ is

$$-T_{+x} = \delta(T_{+xx} + T_{+yy}) \quad y > 0, \quad -\infty < x < \infty, \quad (23a)$$

$$-T_{-x} = T_{-xx} + T_{-yy} \quad y < 0, \quad -\infty < x < \infty, \quad (23b)$$

$$T_{+y}(x, 0) = T_{-y}(x, 0) = 0 \quad \text{for } x > 0 \quad (23c)$$

$$T_{+}(x, 0) = T_{-}(x, 0); \quad \delta T_{+y}(x, 0) = T_{-y}(x, 0) \quad \text{for } x < 0, \quad (23d)$$

$$T_{+} \rightarrow 66 \quad \text{as } y \rightarrow +\infty; \quad T_{-} \rightarrow 192 \quad \text{as } y \rightarrow -\infty. \quad (23e)$$

Here T_{+} and T_{-} are the temperature fields in the web and the chill regions, respectively. The small parameter δ is the ratio of the thermal diffusivities of the web and the chill,

$$\delta = \kappa_w / \kappa_{ch} \approx 1/40. \quad (24)$$

The goal is to calculate the temperature gradient near the leading edge $(x, y) = (0, 0)$ and to determine its singular behavior. In particular, we calculate

$$I(x) \equiv T_{-y}(x, 0^-) \quad \text{as } x \rightarrow 0^-. \quad (25)$$

The problem (23) is difficult to solve analytically and hence we will seek an asymptotic solution valid for small δ (i.e. $\delta \ll 1$). When δ is small, the extent of diffusion of heat from the roller into the web is limited, and hence the temperature in the chill region deviates only slightly from its asymptotic value as $y \rightarrow -\infty$. In addition, a thermal boundary layer occurs in the web region near $y = 0$ and it has a width $O(\delta^{1/2})$. The relevant asymptotic expansion for T_{+} is

$$T_{+}(x, \tilde{y}\delta^{1/2}) \equiv u(x, \tilde{y}) = u_0(x, \tilde{y}) + O(\delta), \quad (26)$$

while for T_{-} it is

$$T_{-}(x, y) = 192 + \delta^{1/2}v(x, y) + \dots \quad (27)$$

Substituting (26) and (27) into (23), and collecting powers of δ , we find that u_0 satisfies

$$-u_{0x} = u_{0\tilde{y}\tilde{y}} \quad y > 0, \quad (28a)$$

$$u_{0\tilde{y}}(x, 0) = 0 \quad \text{for } x > 0; \quad u_0(x, 0) = 192 \quad \text{for } x < 0, \quad (28b)$$

$$u_0 \rightarrow 66 \quad \text{as } y \rightarrow +\infty. \quad (28c)$$

In terms of this solution, the correction v to the temperature in the thermoroll satisfies

$$-v_x = v_{xx} + v_{yy} \quad y < 0, \quad -\infty < x < \infty, \quad (29a)$$

$$v_y(x, 0) = 0 \quad \text{for } x > 0; \quad v_y(x, 0) = u_{0\tilde{y}}(x, 0) \quad \text{for } x < 0. \quad (29b)$$



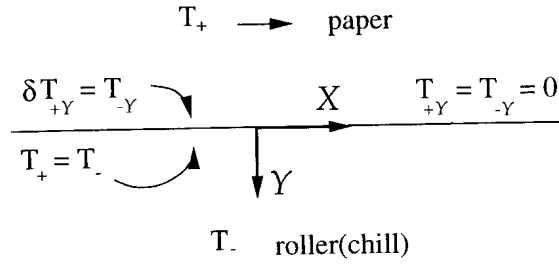


Figure 6:

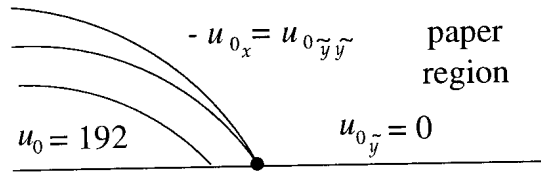


Figure 7: Sketch of u_0 isotherms.

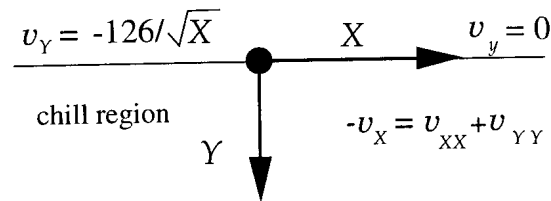


Figure 8:

To get a well-posed problem, we solve (28) in the region $x < 0$, with the “initial” condition $u_0(0, \tilde{y}) = 192$. The solution is readily found to be

$$u_0(x, \tilde{y}) = 66 + 126 \operatorname{Erfc} \left(\frac{\tilde{y}}{2(-x)^{1/2}} \right), \quad (30)$$

where $\operatorname{Erfc}(z)$ is the complementary error function. Thus, the isotherms occur along the curves $\tilde{y} = c(-x)^{1/2}$ for $c > 0$ and $x < 0$ (see Fig. 7). A simple calculation then yields

$$u_{0\tilde{y}}(x, 0) = -\frac{126}{(-\pi x)^{1/2}}, \quad \text{for } x < 0, \quad (31)$$

which is used in (29) (see Fig. 8).

The solution to (29) can be found in terms of the Green’s function, $G(\mathbf{x}'; \mathbf{x})$, which solves the adjoint problem corresponding to the convection-diffusion operator in (29)

$$G_{x'x'} + G_{y'y'} - G_{x'} = \delta(\mathbf{x}' - \mathbf{x}), \quad y' < 0, \quad -\infty < x' < \infty, \quad (32a)$$

$$G_{y'} = 0 \quad \text{on } y' = 0. \quad (32b)$$



The solution to (32), found from the method of images, is

$$G(\mathbf{x}'; \mathbf{x}) = -\frac{1}{2\pi} e^{(x'-x)/2} [K_0(|\mathbf{x}' - \mathbf{x}|) + K_0(|\mathbf{x}' - \mathbf{x}^*|)]. \quad (33)$$

where $\mathbf{x} = (x, y)$, $\mathbf{x}' = (x', y')$, and $\mathbf{x}^* = (x, -y)$. This yields the integral representation

$$v(x, y) = -\frac{132}{\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{-x'}} G(\mathbf{x}'; \mathbf{x}) \Big|_{y'=0} dx'. \quad (34)$$

Substituting (30) and (34) into (26) and (27) yields the temperature profile in the web and the chill regions, respectively. It is easy to show that the limiting behavior of v as we approach the leading edge is

$$v \sim Cr^{1/2} \sin \theta/2, \quad \text{as } r = |\mathbf{x}| \rightarrow 0, \quad (35)$$

where $C = -252/\sqrt{\pi}$.

The solution for T_- in the paper becomes invalid in an $O(\delta)$ neighborhood of leading edge. This follows from the fact that we have neglected the diffusion term T_{-xx} in obtaining (28). Hence we need an inner-inner region where $x = O(\delta)$ and $y = O(\delta)$. Introduce the new variables

$$\hat{x} = x/\delta, \quad \hat{y} = y/\delta, \quad (36)$$

and expand

$$T_+ = \hat{u}(\hat{x}, \hat{y}) + \dots, \quad T_- = 192 + \delta \hat{v}(\hat{x}, \hat{y}) + \dots \quad (37)$$

Substituting (36) and (37) into (23), we obtain that \hat{u} satisfies

$$-\hat{u}_{\hat{x}} = \hat{u}_{\hat{x}\hat{x}} + \hat{v}_{\hat{y}\hat{y}} \quad \hat{y} > 0, \quad (38a)$$

$$\hat{u}_{\hat{y}}(\hat{x}, 0) = 0 \quad \text{for } \hat{x} > 0; \quad \hat{u}(\hat{x}, 0) = 192 \quad \text{for } \hat{x} < 0, \quad (38b)$$

$$\hat{u} \rightarrow 66 \quad \text{as } \hat{y} \rightarrow +\infty. \quad (38c)$$

In terms of the solution to this problem, \hat{v} satisfies

$$\hat{v}_{\hat{x}\hat{x}} + \hat{v}_{\hat{y}\hat{y}} = 0 \quad \hat{y} < 0, \quad -\infty < \hat{x} < \infty, \quad (39a)$$

$$\hat{v}_{\hat{y}}(\hat{x}, 0) = 0 \quad \text{for } \hat{x} > 0; \quad \hat{v}_{\hat{y}}(\hat{x}, 0) = \hat{u}_{\hat{y}}(\hat{x}, 0) \quad \text{for } \hat{x} < 0 \quad (39b)$$

The solution \hat{v} must match with the behavior (35) in the far field.

The solution to (38) can be found explicitly in terms of the parabolic coordinates ξ and η defined by

$$\hat{x} = \frac{1}{2}(\xi^2 - \eta^2), \quad \hat{y} = \xi\eta. \quad (40)$$



In terms of these coordinates the problem (38) reduces to the following ordinary differential equation problem for $\hat{u}(\eta)$:

$$\hat{u}_{\eta\eta} + \eta\hat{u}_{\eta} = 0, \quad 0 < \eta < \infty; \quad \hat{u}(0) = 192, \quad \hat{u}(\infty) = 66. \quad (41)$$

The solution to this problem is

$$\hat{u}(\eta) = 66 + 126 \operatorname{Erfc}(\eta/2). \quad (42)$$

Using the definition of the change of coordinates we can calculate the derivative needed in the problem (39) for \hat{v} . We get

$$\hat{u}(\hat{x}, 0) = -\frac{126}{(-\pi x)^{1/2}}, \quad \text{for } x < 0, \quad (43)$$

Finally, we can substitute (43) into (39) and solve the resulting problem for \hat{v} to get

$$\hat{v} = C\hat{r}^{1/2} \sin(\theta/2), \quad \text{where } C = -252/\sqrt{\pi} \quad \text{and } \hat{r} = (\hat{x}^2 + \hat{y}^2)^{1/2}. \quad (44)$$

Substituting (44) into (37) determines the temperature field in the chill region of the thermoroll in the immediate vicinity of the leading edge where the paper first makes contact with the roller.

The main conclusion from this analysis is that this temperature field near the leading edge has the behavior

$$T_- \sim Br^{1/2} \sin(\theta/2) \quad \text{as } r \rightarrow 0, \quad (45)$$

for some constant B . Thus, it has an infinite gradient of square-root type at the leading edge. In §6 we estimate whether this singular form leads to a singular stress field at the leading edge.

7 The Stress Field

We now estimate the stress field induced by the temperature gradients calculated in §4 and §5. We will consider both the local and the global problems.

The equilibrium equation from elasticity theory for the displacement vector \mathbf{u} is

$$\frac{3(1-\nu)}{(1+\nu)} \nabla(\nabla \cdot \mathbf{u}) - \frac{3(1-2\nu)}{2(1+\nu)} \nabla \times \nabla \times \mathbf{u} = \alpha \nabla T. \quad (46)$$

Here ν is Poisson's ratio, $\alpha \nabla T$ is the body force induced by the thermal gradient, and α is a constant.

The stress tensor σ_{ij} is determined in terms of \mathbf{u} by

$$\sigma_{ij} = \lambda \delta_{ij} \frac{\partial u_k}{\partial x_k} + G \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right). \quad (47)$$



Here λ and G are the Lamé constants. Thus, the stress is (essentially) proportional to the first derivatives of \mathbf{u} .

Consider the **local problem** for which the temperature field behaves like (45) at the leading edge. Then decomposing $\nabla\mathbf{u}$ in terms of a radial component u_r and an angular component u_θ , we get the following equations from (46):

$$\frac{3(1-\nu)}{(1+\nu)} (\partial_{rr} u_r + \dots) - \frac{3(1-2\nu)}{2(1+\nu)} \left(\frac{\partial_{r\theta} u_\theta}{r} + \dots \right) = \alpha \left(\frac{C}{2r^{\frac{1}{2}}} \right) \sin \left(\frac{\theta}{2} \right), \quad (48a)$$

$$\frac{3(1-\nu)}{(1+\nu)} \left(\frac{\partial_{r\theta} u_r}{r} + \dots \right) + \frac{3(1-2\nu)}{2(1+\nu)} (\partial_{rr} u_\theta + \dots) = -\alpha \left(\frac{C}{2r^{\frac{1}{2}}} \right) \cos \left(\frac{\theta}{2} \right). \quad (48b)$$

From these equations, a simple dominant balance argument shows that the two components satisfy $u_r = O(r^{3/2})$ and $u_\theta = O(r^{3/2})$ as $r \rightarrow 0$. Thus the stress field has the form

$$\sigma_{ij} = O \left(r^{\frac{1}{2}} \right) \quad \text{as } r \rightarrow 0. \quad (49)$$

Hence the stress field is not singular at the leading edge, and there is no significant stress intensification, such as that determined by a stress intensity factor. In fact, the stress field behaves like that for a contact problem (i.e. a Barenblatt “crack”).

Now consider the **global problem** where the nip region is approximated by a delta function. The geometry is shown in Fig. 4 and the temperature field has the behavior given in (19). Substituting $T \sim \ln r$ into (46) we can obtain the behavior of u_r using a dominant balance argument. We estimate $\partial_r u_r = O(\ln r)$ and hence the radial component of the stress is $\sigma_r = O(\ln r)$. Thus, the stress field grows logarithmically in the outer region but is then cut off as we approach the inner-inner region. Once again, we conclude that there is no significant intensification of the stress, such as that determined by a square root singularity for which a stress intensity factor can be defined.

8 Conclusions

The main focus of our group was to calculate the temperature gradient in the thermoroll and to determine whether this gradient can lead to an intensification of stress in the nip region.

The temperature gradient was calculated for both a global temperature model in which the nip is represented by a point source and a local temperature problem, defined in the vicinity of the nip region, where we studied in detail the region where the paper first comes into contact with the roller. For both the local and global temperature problems we calculated the singular behavior of the thermal field. In §6, we used the singular behavior of the thermal field in a model to estimate the stress field for both the local and global problems. The goal was to ascertain whether such a temperature gradient can lead to a singular stress field. Such a stress field is known in other circumstances to induce crack formation. Our conclusion in §6 is that the stress field is not singular



for either the local or global problems, and hence, from our model of the thermoroll, the temperature gradients are not likely to be the cause of roll fracture.

We also showed that in the limit of large drum rotation, which is the usual operating regime of the thermoroll, the temperature distribution inside the thermoroll is well approximated by a radially symmetric function. This can allow for an accurate calculation of the surface temperature on the roller once a more precise boundary condition can be applied to the roller surface.

Finally, a crude physical estimate in §3 suggested that there can be extremely large temperature gradients as a result of dripping water onto the thermoroll. This problem certainly warrants an intensive investigation.

1 Acknowledgements:

We are grateful to Dr. Roman Popil of MacMillan Bloedel Research Ltd. for his explanation of this problem to us and for his detailed comments on this report. We would also like to thank Dhavide Aruliah for preparing the figures for this paper.

2 Appendix A: Physical Constants

We summarize some relevant physical constants given by MacMillan Bloedel Research:

| | |
|----------------------------|-------------|
| Web temperature out of nip | = 92°C |
| Web temperature into nip | = 66°C |
| Web velocity | = 18m/sec |
| Web width | = 7770 mm |
| Inner core radius | = 560mm |
| Outer core radius | = 750mm |
| Thickness of chill | = 12mm |
| Thickness of shell | = 49mm |
| Thermal conduct. chill | = 24W/mK |
| Thermal conduct. core | = 48W/mK |
| Bore hole radius | = 16mm |
| Number of bores | = 45 |
| Oil flow rate | = 33.7L/sec |
| Oil temperature | = 253°C |

The thermal diffusivities of the water, chill and core, which we label by κ_w , κ_{ch} and κ_{co} , respectively, are now calculated in terms of more conventional units. For water, $\rho = 10^3 \text{ kg/m}^3$, $C_\rho = 4186 \text{ J/kg-K}$, $k = 0.6 \text{ W/mK}$, and thus

$$\kappa_w = k/\rho C_\rho = 0.142 \times 10^{-6} \text{ m}^2/\text{sec} = 0.142 \times 10^{-2} \text{ cm}^2/\text{sec}. \quad (50)$$



For the chill, $\rho = 8.0 \times 10^3 \text{ kg/m}^3$, $C_\rho = 420 \text{ J/kg-K}$, $k = 24 \text{ W/mK}$, and thus

$$\kappa_{ch} = k/\rho C_\rho = 0.71 \times 10^{-5} \text{ m}^2/\text{sec} = 0.071 \text{ cm}^2/\text{sec}. \quad (51)$$

Lastly, $\kappa_{co} = 2\kappa_{ch}$.

Finally, we quote the estimate of the heat flow into the web given by Macmillan Bloedel Research. For each unit length along the roll we have $dm/dt = 1.1 \text{ kg/s}$ and the heat flux into the web is

$$\frac{dQ}{dt} = \sum C_\rho \frac{dm}{dt} \Delta T = 50 \text{ kW}. \quad (52)$$

Here $\Delta T = 26^\circ\text{C}$ is the temperature difference along the web. This input estimate might be based on too much water content, and if so it would turn out that the heat input into the web is smaller by a factor of ten.



Chapter 5

Inversion for Anisotropic Velocity Parameter

September 1997

*Ben Aggarwala¹, Ellis Cumberbatch², Jeff Grossman³,
Michael Lamoureux⁴, Vlad Shapiro⁵, Mark Solomonovitch⁶, Paul Webster⁷*

written by Michael Lamoureux

1 Introduction

This report describes the mathematical results of a team of seven researchers working under the auspices of the 1997 PIMSIPS Workshop, organized in Vancouver by the Pacific Institute for the Mathematical Sciences. The problem under study was raised by PetroCanada, and concerns the robust computation of a certain parameter of anisotropy from observed traveltimes of a seismic shear wave propagating through a geological medium.

2 Problem Description

A routine simplifying assumption in geophysical studies is that the velocity of a seismic wave in a given layer of material is independent of the direction of propagation; such a material is said to be isotropic. While this assumption greatly simplifies the mathematics describing certain geophysical phenomena, it ignores the physical reality that many materials which occur in geological layers are

¹Dept. Math. & Stat., University of Calgary, Calgary AB

²Dept. Math., Claremont Graduate University, Claremont CA

³Dept. Math. & Stat., University of Calgary, Calgary AB

⁴Dept. Math. & Stat., University of Calgary, Calgary AB

⁵TRIUMF, UBC, Vancouver BC

⁶Dept. Math., University of Alberta, Edmonton AB

⁷Dept. Physics, University of Calgary, Calgary AB



not isotropic. Understanding and measuring this so-called velocity anisotropy plays a key role in the development of a more accurate geophysical model.

Researchers at PetroCanada have developed an algorithm for computing a measure of anisotropy from recordings of traveltimes of seismic signals traversing various paths through horizontally layered media. The algorithm is based on an approximation to an exact physical model, and assumes an elliptical velocity profile in the anisotropy. The main problem addressed in this project is that the resulting approximate formula for anisotropy is extremely sensitive to all input parameters, in particular to the traveltimes measurements which form the standard set of geophysical data.

PetroCanada is seeking to understand why this method is so sensitive to input data. Some possibilities are that the numerical methods used are not robust, or that the approximation is a source of error. More generally, they wish to find alternative, more robust methods of computing the anisotropy parameter.

3 Background Material

PetroCanada provided us with two references as background material: the 1996 Ph.D. thesis entitled “On elastic-wave propagation in anisotropic media: . . . ” by Michael Slawinski; and an abstract of a presentation to the recent SEG International Exposition entitled “Analytic inversion for Thomsen’s γ parameter in weakly anisotropic media” by Michael Slawinski and Raphael Slawinski.

Within these works, the mathematics describing ray propagation through anisotropic media is presented without derivation. A first step is a reduction of the problem to two dimensions, which is typical in certain seismic imaging situations where all measurements are carried out in one plane. In two dimensions, the direction of propagation may in principle be described by a single angle. However, there are in fact two angles, the phase and group angles θ_p and θ_g , which are of interest in the mathematical formulation, along with the related phase and group velocities c_p and c_g . The anisotropy of a medium is indicated by Thomsen’s parameter γ , from which one may describe the elliptical profile of the velocity, as a function of the phase angle, by the formula

$$c_p(\theta_p) = \beta \sqrt{1 + 2\gamma \sin^2 \theta_p}. \quad (1)$$

Here, β is the vertical velocity of propagation. The group velocity, as a function of phase angle, is given by

$$c_g(\theta_p) = \beta \sqrt{\frac{1 + 2\gamma(1 + \gamma)(1 - \cos 2\theta_p)}{1 + \gamma(1 - \cos 2\theta_p)}}, \quad (2)$$

and the phase and group angles are related by

$$\tan \theta_g = (1 + 2\gamma) \tan \theta_p. \quad (3)$$



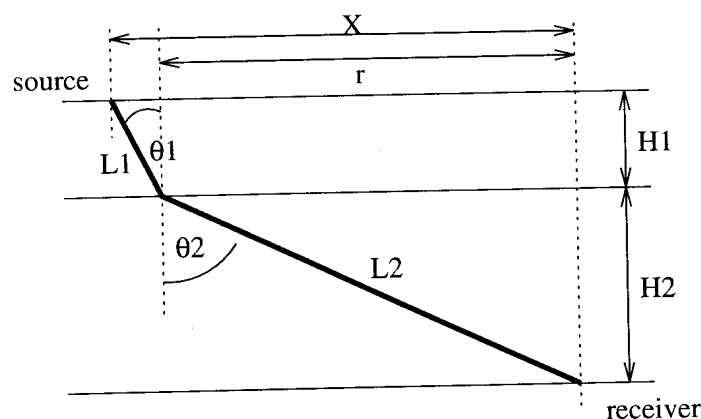


Figure 1: Two-layer model

Figure 1: Propagation of a seismic wave through a two layer medium.

An approximation is then made, assuming γ is small, so that

$$\theta_p \approx \theta_g \quad (4)$$

and

$$c_g \approx c_p \approx \beta(1 + \gamma \sin^2 \theta_p). \quad (5)$$

The inversion formula for γ is based on this approximation.

The physical model considered in the reference work is propagation of a seismic wave through a two layer medium as illustrated in Figure 1. The upper layer is isotropic, with velocity parameter c_1 , while the lower layer is anisotropic with vertical velocity parameter β and anisotropic velocity parameter γ . A source on the surface, at some distance X from a vertical bore hole, generates a seismic signal which travels to a receiver at the bottom of the bore hole along a bent ray that refracts at the interface between the two layers some distance r from the bore hole. Noting that $r/L_2 = \sin \theta_2$ is the sine of the angle of propagation, one finds the traveltime for the signal is

$$\begin{aligned} t &= \frac{L_1}{c_1} + \frac{L_2}{c_2} \\ &\approx \frac{\sqrt{(X-r)^2 + H_1^2}}{c_1} + \frac{\sqrt{r^2 + H_2^2}}{\beta(1 + \gamma r^2/(r^2 + H_2^2))}. \end{aligned} \quad (6)$$

Inverting this approximate formula gives an equation for γ as a function of the measured traveltime t and the refraction point r .



To find r , one applies Fermat's principle of stationary time and solves

$$\frac{dt}{dr} = 0. \quad (7)$$

For the inverse problem, one solves

$$\frac{d\gamma}{dr} = 0 \quad (8)$$

to obtain the refraction point r and then evaluates the inverse formula for γ at the given r, t values. This forms the basis for an approximation algorithm to compute γ .

The background material supplies a concrete computational example. A two layer medium is given, with the upper isotropic layer of thickness 355m and velocity parameter 1030m/s, and the lower anisotropic layer of thickness 1045m, vertical wave speed of 1609m/s and anisotropic parameter $\gamma = .096$. A chart is given of various traveltimes, for X in the range 0 to 990m, and the computed gamma obtained from the inversion formula. This chart is reproduced in Table 3, including the erroneous first four traveltimes.

| Offset X | Traveltime t | Inverted γ |
|------------|----------------|-------------------|
| 0 | 1.25 | N/A |
| 90 | 1.25104 | .0959772 |
| 190 | 1.27558 | .0961738 |
| 290 | 1.3484 | .0958288 |
| 390 | 1.02475 | .0959165 |
| 490 | 1.04204 | .0959752 |
| 590 | 1.06286 | .0959918 |
| 690 | 1.08699 | .0959842 |
| 790 | 1.11419 | .0959669 |
| 890 | 1.1442 | .0959865 |
| 990 | 1.17678 | .0959947 |

Table 1: Approx. traveltime & inversion

4 Development of the solution

Our team began with an investigation aimed at understanding the elliptical velocity profile for an anisotropic material, with the intent of deriving from first principles the mathematical equations presented in the background material. We modeled the wave propagation with a constant coefficient wave equation and found the exact relationship between two orthogonal components of velocity and the anisotropic parameter γ . From this model we successfully derived all the formulas stated in the background material.

We then investigated the accuracy of the approximation used in the background work. From our model above, we produced exact formulas for Snell's law, traveltime through the two layer medium,



and the inversion formula for γ . We noted that the exact formulas are no more difficult to work with than the approximations of the background material.

The next step was to redo the numerical work done using Slawinski's approximate formula, and to compare these results with the numerical results obtained using our exact formulation. We uncovered some numerical errors in the data in Table 3, and we made an observation from the numerical work that finding critical points via $dt/dr = 0$, $d\gamma/dr = 0$ is equivalent to minimizing t and γ over a permissible range of r values. We then proved this result mathematically, and used the minimization routine in MATLAB as a more robust method of solving the propagation model.

From this numerical work, we were then able to observe the sensitivity to input parameters for the γ inversion, even in the exact formulation. We noted the cause of the sensitivity, which is intrinsic to γ , and stabilized the problem by recasting the inversion in terms of two velocity parameters α and β . Plots of surfaces of intersection were created in Maple to demonstrate the stabilization.

We then produced asymptotic formulas for time of travel and γ which show more directly the sensitivity of γ , and finally proposed what we believe is a promising method of collecting detailed seismic data which gives a more stable measure of anisotropy.

5 Model, solution, and results

1 Wave equation formulation

We begin with a constant coefficient wave equation to describe the propagation of a wave in an anisotropic, two dimensional medium as

$$\frac{\partial^2 u}{\partial t^2} = \alpha^2 \frac{\partial^2 u}{\partial x^2} + \beta^2 \frac{\partial^2 u}{\partial y^2}, \quad (9)$$

where α, β are the velocity parameters in the horizontal and vertical directions, respectively. The mixed term $\frac{\partial^2 u}{\partial x \partial y}$ is absent in the equation, due to the horizontal/vertical orientation of the velocity ellipse.

For a plane wave with wave number $k > 0$ traveling at some phase angle θ_p relative to the vertical, we choose a normal vector $\mathbf{k} = (l, m) = (k \sin \theta_p, k \cos \theta_p)$ to obtain the plane wave solution

$$u = e^{i(xk \sin \theta_p + yk \cos \theta_p - \omega t)}. \quad (10)$$

Inserting u into the wave equation yields

$$\omega^2 = k^2 (\alpha^2 \sin^2 \theta_p + \beta^2 \cos^2 \theta_p), \quad (11)$$



from which we derive the phase velocity of the wave as

$$\begin{aligned}
c_p = \frac{\omega}{k} &= \sqrt{\alpha^2 \sin^2 \theta_p + \beta^2 \cos^2 \theta_p} \\
&= \beta \sqrt{1 + \frac{\alpha^2 - \beta^2}{\beta^2} \sin^2 \theta_p} \\
&= \beta \sqrt{1 + 2\gamma \sin^2 \theta_p},
\end{aligned} \tag{12}$$

where we have identified $\gamma = (\alpha^2 - \beta^2)/2\beta^2$ as Thomsen's anisotropy parameter. Thus equation (12) captures exactly the elliptical velocity profile of the phase velocity as described in equation (1) of the background material.

Group velocity is a vector, obtained by the vector derivative

$$\mathbf{c}_g = \frac{\partial \omega}{\partial \mathbf{k}} = \left(\frac{\partial \omega}{\partial l}, \frac{\partial \omega}{\partial m} \right). \tag{13}$$

As $\omega^2 = \alpha^2 l^2 + \beta^2 m^2$ we have $2\omega \frac{\partial \omega}{\partial l} = 2l\alpha^2$ and $2\omega \frac{\partial \omega}{\partial m} = 2m\beta^2$, thus

$$\mathbf{c}_g = \frac{1}{\omega} (\alpha^2 l, \beta^2 m) = \frac{k}{\omega} (\alpha^2 \sin \theta_p, \beta^2 \cos \theta_p). \tag{14}$$

The magnitude of the group velocity vector is thus

$$\begin{aligned}
c_g &= \frac{k}{\omega} \sqrt{\alpha^4 \sin^2 \theta_p + \beta^4 \cos^2 \theta_p} \\
&= \frac{\sqrt{\alpha^4 \sin^2 \theta_p + \beta^4 \cos^2 \theta_p}}{\sqrt{\alpha^2 \sin^2 \theta_p + \beta^2 \cos^2 \theta_p}},
\end{aligned} \tag{15}$$

while the ratio of the components of \mathbf{c}_g gives the tangent of the group angle θ_g as

$$\tan \theta_g = \frac{\alpha^2 \sin \theta_p}{\beta^2 \cos \theta_p} = \frac{\alpha^2}{\beta^2} \tan \theta_p. \tag{16}$$

Rewriting equations (15) and (16) in terms of the parameters β and γ yields immediately equations (2) and (3) of the background material, showing that our wave equation model is consistent with the description in the background material.

We now derive an exact formula for time of travel of a ray propagating through a two layer medium as in Figure 1. It is important to note that the angle θ_2 in the diagram is in fact the group angle, and the relevant velocity in the anisotropic layer is the group velocity, thus the time of travel is given by

$$\begin{aligned}
t &= \frac{L_1}{c_1} + \frac{L_2}{c_g} \\
&= \frac{\sqrt{(X-r)^2 + H_1^2}}{c_1} + \sqrt{r^2 + H_2^2} \sqrt{\frac{\alpha^2 \sin^2 \theta_p + \beta^2 \cos^2 \theta_p}{\alpha^4 \sin^2 \theta_p + \beta^4 \cos^2 \theta_p}} \\
&= \frac{\sqrt{(X-r)^2 + H_1^2}}{c_1} + \sqrt{r^2 + H_2^2} \sqrt{\frac{\alpha^2 \tan^2 \theta_p + \beta^2}{\alpha^4 \tan^2 \theta_p + \beta^4}},
\end{aligned} \tag{17}$$



which reduces, by the relation $r/H_2 = \tan \theta_g = \frac{\alpha^2}{\beta^2} \tan \theta_p$ to the simple form

$$t = \frac{\sqrt{(X-r)^2 + H_1^2}}{c_1} + \sqrt{\frac{r^2}{\alpha^2} + \frac{H_2^2}{\beta^2}}. \quad (18)$$

Fermat's principle of stationary time $\frac{dt}{dr} = 0$ yields the equation

$$\frac{X-r}{c_1 \sqrt{(X-r)^2 + H_1^2}} = \frac{r}{\alpha^2 \sqrt{\frac{r^2}{\alpha^2} + \frac{H_2^2}{\beta^2}}} \quad (19)$$

or equivalently

$$\frac{\sin \theta_1}{c_1} = \frac{\sin \theta_p}{c_p} \quad (20)$$

which is Snell's law at the interface of the isotropic/anisotropic layers. Alternatively, Snell's law may be obtained by equating plane wave solutions at the two layer boundary.

Inverting equation (18) for α^2 yields

$$\alpha^2 = \frac{r^2}{(t - \sqrt{(X-r)^2 + H_1^2}/c_1)^2 - (H_2/\beta)^2} \quad (21)$$

and thus we obtain an exact formula for γ as

$$\begin{aligned} \gamma &= \frac{1}{2} \left(\frac{\alpha^2}{\beta^2} - 1 \right) \\ &= \frac{1}{2} \left(\frac{r^2}{(t - \sqrt{(X-r)^2 + H_1^2}/c_1)^2 - (H_2/\beta)^2} - 1 \right) \end{aligned} \quad (22)$$

We note that the exact formula for γ has a singularity in it, which is relevant when attempting to minimize γ as a function of r .

The background material mentions that the refraction point r could be obtained by solving either $dt/dr = 0$ (Fermat's principle) or $d\gamma/dt = 0$. We note the following equivalence for minima.

Proposition 4.1. *If the critical point for $dt/dr = 0$ is a local minimum, then so is the corresponding critical point for $d\gamma/dr = 0$.*

This proposition is a result of the chain rule. To summarize, observe that we have a formula $t = F(r, \gamma)$ describing travelttime in terms of the parameters r and γ ; from the inversion formula, we may write γ as a function of r and t , so inserting into F gives

$$t = F(r, \gamma(r, t)). \quad (23)$$

Treating r, t as independent variables, we may differentiate (23) twice with respect to r to obtain

$$0 = \frac{\partial^2 F}{\partial r^2} + \frac{\partial^2 F}{\partial r \partial \gamma} \frac{\partial \gamma}{\partial r} + \frac{\partial F}{\partial \gamma} \frac{\partial^2 \gamma}{\partial r^2}. \quad (24)$$



But $\frac{\partial \gamma}{\partial r} = 0$ at the critical point, while $\frac{\partial^2 F}{\partial r^2}$ is positive there (as this is just the second derivative of t with respect to r , evaluated at the minimum), and $\frac{\partial F}{\partial \gamma}$ is negative for the physical reason that traveltime decreases as the anisotropy parameter increases, due to the increasing horizontal velocities. Thus from (24), the second derivative $\frac{\partial^2 \gamma}{\partial r^2}$ is positive, indicating a minimum for γ .

Since the seismic apparatus actually records the minimum time of arrival of a signal, we may assume the critical point for t , and thus for γ , is in fact minimizer. Thus we obtain the following algorithms for computing traveltime and inverting γ based on minimization.

Algorithm for traveltime

Given parameters $c_1, \alpha, \beta, H_1, H_2$ and X , minimize t with respect to r as per formula (18). Resulting t value is the traveltime.

Algorithm for γ inversion

Given parameters c_1, t, β, H_1, H_2 and X , minimize γ with respect to r as per formula (22). Resulting γ value is the anisotropic parameter.

Note that we may also minimize α^2 directly, as it is a linearly increasing function of γ .

2 Asymptotics

The numerical results show substantial errors and sensitivity to input data, in particular at the smaller values of X . We develop some asymptotic approximations using the exact equations for γ in order to obtain explicit formulas for the inverse problem. The first approximation we considered is for $X - r \ll H_1$. Expanding the exact relations (18), (19) in this approximation yields

$$t = \frac{H_1}{c_1} \left(1 + \frac{(X - r)^2}{2H_1^2} \right) + \frac{1}{\alpha\beta} (\alpha^2 H_2^2 + \beta^2 r^2)^{\frac{1}{2}}, \quad (25)$$

$$\frac{\alpha^4 H_2^2}{\beta^2 r^2} + \alpha^2 = c_1^2 \frac{H_1^2}{(X - r)^2}. \quad (26)$$

At first sight keeping the term $(X - r)^2/H_1^2$ in (25) may seem inconsistent, since a term of the same relative order has been dropped from the right-hand side of (26). Neglecting this term in (25) facilitates a simple explicit solution to the inverse problem: r/α is obtained from (25), and its substitution into (26) yields α , which can then be eliminated. The result is

$$r(X - r) = c_1 H_1 [(t - H_1/c_1)^2 - H_2^2/\beta^2] / (t - H_1/c_1) \quad (27)$$

However, inserting traveltimes into this expression for typical parameter values results in complex r values. This is further evidence of the sensitive nature of the inversion algorithm to small errors.



Our second asymptotic formula is for almost vertical rays. Here we consider both $X, r \ll H_1$ so that the previous approximations (25), (26) hold. In addition, for almost vertical rays, the traveltime t almost equals t_0 , where

$$t_0 = H_1/c_1 + H_2/\beta$$

is the traveltime in the vertical case, $X = 0$. Equations (25), (26) imply that $t - t_0, r^2$ and X^2 are of like order. We put

$$t - t_0 = \Delta^2 H_2/\beta, \quad X = \Delta qh, \quad r = \Delta ph, \quad (28)$$

where $h = (2H_1H_2c_1/\beta)^{1/2}$ is a convenient length scale, and p, q, Δ are scalars, with $\Delta \ll 1$. Inserting these into (25), (26), and retaining the leading order terms, gives

$$p = q - 1/q \quad (29)$$

and

$$\frac{\alpha^2}{\beta^2} = \frac{h^2}{H_2^2} \left(\frac{q^2 - 1}{2} \right). \quad (30)$$

The above asymptotic results for this approximation allow the identification of α/β from almost vertical traveltime measurements via simple formulas. First, Δ and q are obtained from (28) in terms of $t - t_0$ and X . Then α/β is found from (30). In terms of original parameters,

$$\frac{\alpha^2}{\beta^2} = \frac{c_1 H_1}{\beta H_2} \left(\frac{X^2}{2c_1 H_1 (t - t_0)} - 1 \right) \quad (31)$$

The accuracy of this formula is discussed further in the section on numerical results.

We may observe directly the effect of measurement errors by this vertical ray approximation of equation (30), the horizontal velocity α in the anisotropic medium is presented in terms of the measured traveltime and other parameters. Since it is available as an explicit formula, it provides an excellent resource to examine parametric dependence of the identification problem. Of particular interest is the error in α resulting from measurement error in β . Let δ be the fractional change in β . That is β is replaced by $\beta(1 + \delta)$ in (31). The fractional change in α is calculated from

$$\frac{1}{\alpha^2} \Delta(\alpha^2) = \frac{1}{\alpha^2} \frac{\partial \alpha^2}{\partial \beta} \beta \delta = \left[1 - \frac{X^2}{2\alpha^2} \frac{1}{(t - t_0)^2} \right] \delta \quad (32)$$

For typical values ($\gamma = 0.06$, $X = 190\text{m}$) the right-hand side has the value -95δ . That is, a 1% error in β gives rise to a 95% error in α^2 . This large amplification is, of course, due to the smallness of $t - t_0$, its occurrence in the denominator of (31), and the dependence of t_0 on β .



6 Numerical results

We implemented the minimization algorithms for finding traveltime t and anisotropic parameter γ in MATLAB, using both our exact formulas and Slawinski's approximations. The parameter values for $c_1, \gamma, \beta, H_1, H_2$, are exactly as in the computational example in the background material. We explicitly avoid the singularities in the minimization routines to guarantee accurate results.

A comparison was made of the results of the exact formula with the approximate formula, with results summarized in Table 6. The approximate values of Table 3 's are recovered (note the first four values in Table 3 are in error; Slawinski indicates an editing problem), the exact answers are in good agreement with the approximation, and our γ inversions are exact to machine accuracy. Note that in the background work, the γ inversion was not exact even for this synthetic data; we do not know if this indicates an inaccuracy in the MATHEMATICA software used in the background material, or simply the use of inexact input data in the background work.

| Offset X | Traveltime | | | Inverted γ | |
|---------------|------------|---------|-------|-------------------|----------|
| | Approx | Exact | Error | Approx | Exact |
| 0 | 0.99413 | 0.99413 | 0 % | .0960000 | .0960000 |
| 90 | 0.99579 | 0.99584 | .01 % | .0960000 | .0960000 |
| 190 | 1.00148 | 1.00171 | .02 % | .0960000 | .0960000 |
| 290 | 1.01117 | 1.01169 | .05 % | .0960000 | .0960000 |
| 390 | 1.02474 | 1.02564 | .09 % | .0960000 | .0960000 |
| 490 | 1.04204 | 1.04336 | .13 % | .0960000 | .0960000 |
| 590 | 1.06286 | 1.06465 | .17 % | .0960000 | .0960000 |
| 690 | 1.08699 | 1.08925 | .21 % | .0960000 | .0960000 |
| 790 | 1.11418 | 1.11690 | .24 % | .0960000 | .0960000 |
| 890 | 1.14420 | 1.14735 | .28 % | .0960000 | .0960000 |
| 990 | 1.17678 | 1.18034 | .30 % | .0960000 | .0960000 |

Table 2: Approximate vs. exact comparisons

We then investigated the sensitivity of the γ inversion to input data, the results of which are summarized in Table 6. Boosting t by 1 percent gave errors in γ as large as 600 percent, as did boosting β by 1 percent, with the larger errors occurring for small values of X . Even for midrange values of X , the errors are in the range of 50 to 200 percent.

However, further numerical work showed that if we compute the α parameter from the time data, the result is much less sensitive to input data, as summarized in Table 3. In this case, boosting t by 1 percent gave errors in the inverted α only as large as 43 percent (disregarding the physically impossible result of zero velocity), while boosting β by 1 percent gave errors as large as 33 percent for similar offset. Again, the larger errors appeared for small values of X , with midrange values for X giving errors on the order of 5 to 10 percent. The overall error for α is much smaller than for γ .



| Offset X | Time error = 1% | | β error = 1% | |
|-------------|-----------------|-------|--------------------|-------|
| | γ | Error | γ | Error |
| 90 | -.5000 | 621 % | -.4621 | 581 % |
| 190 | -.3064 | 419 % | -.2311 | 341 % |
| 290 | -.1635 | 270 % | -.0984 | 202 % |
| 390 | -.0794 | 183 % | -.0299 | 131 % |
| 490 | -.0298 | 131 % | .0076 | 92 % |
| 590 | .0008 | 99 % | .0298 | 69 % |
| 690 | .0207 | 78 % | .0438 | 54 % |
| 790 | .0343 | 64 % | .0532 | 45 % |
| 890 | .0438 | 54 % | .0597 | 38 % |
| 990 | .0509 | 47 % | .0644 | 33 % |

Table 3: Parameter γ sensitivity to input error

| Offset X | Time error = 1% | | β error = 1% | |
|-------------|-----------------|-------|--------------------|-------|
| | α | Error | α | Error |
| 90 | 0 | 100 % | 443 | 75 % |
| 190 | 1001 | 43 % | 1180 | 33 % |
| 290 | 1320 | 25 % | 1442 | 18 % |
| 390 | 1476 | 16 % | 1560 | 11 % |
| 490 | 1560 | 11 % | 1621 | 8 % |
| 590 | 1610 | 8 % | 1656 | 6 % |
| 690 | 1642 | 7 % | 1678 | 5 % |
| 790 | 1663 | 5 % | 1692 | 4 % |
| 890 | 1678 | 5 % | 1702 | 3 % |
| 990 | 1689 | 4 % | 1709 | 3% |

Table 4: Velocity α sensitivity to input error

For the same parameter values, we tested our asymptotic formulas for inverting γ , as summarized in Table 6. We see this approximation formula recovers γ from exact data to within a couple of percent.

| Exact $\gamma =$ | .03 | .06 | .09 | .12 | .15 |
|------------------|-------|-------|-------|-------|-------|
| X= 90 | .0307 | .0607 | .0907 | .1207 | .1507 |
| X=190 | .0330 | .0631 | .0931 | .1231 | .1532 |
| X=290 | .0371 | .0672 | .0972 | .1273 | .1574 |

Table 5: Asymptotic results for γ

From these numerical observations, we conclude a major problem with the γ inversion is that the parameter γ is intrinsically sensitive to the input data for the given measurement technique. The problem does not arise just from the approximation originally used, or the particular numerical methods, although our minimization method in MATLAB was an improvement over the zero finding of MATHEMATICA.

A physical explanation can be given for this phenomena. In the measurement technique under study, small X values give nearly vertical rays, which measure accurately the vertical velocity parameter β , while large X values give more horizontal rays, which measure accurately the horizontal velocity parameter α . The general ray measures some weighted average of α and β . However, the anisotropic parameter $\gamma = (\alpha^2 - \beta^2)/(2\beta^2) \approx \alpha - \beta$ is the difference of the two velocity parameters, and the current measurement technique does not directly measure it. An accurate measure of γ requires very high accurate, independent measurements of α and β .

To demonstrate the stability of the α , β parameter formulation, we used MAPLE to plot certain surfaces of the parameters to show steepness of the intersection, which implied robustness of the method. The plots have been omitted from the report.

7 Further directions

In order to accurately measure the anisotropic velocity parameter γ , which is essentially the difference of the two velocity parameters α and β , it is necessary to devise a measurement technique that measures some quantity that depends directly on the difference $\alpha - \beta$. Several methods occur as possibilities, however, they are limited by the physical difficulty of performing the measurements.

A key notion is to record the differential time of travel for a single signal traveling along two different paths. If one path is near vertical, and the other more horizontal, the difference in the time of travel will be related to the difference $\alpha - \beta$. Signal processing methods can be devised that record the received signals simultaneously, and compute the differential in time of travel directly.

Generating two signal paths from one seismic event is more problematic. One possibility is to



consider two different polarizations of the seismic signal, as separated out by a 3-axis geophone. Depending on the nature of the anisotropy of the geological layer, it is possible that different polarizations of a single signal will travel along different paths; the difference in time of travel will depend on $\alpha - \beta$.

Another possibility is to have more than one recording device, at widely separated positions in the bore hole. One seismic signal on the surface would generate a pulse at both receivers, again traveling along different paths. With proper placements of the receivers, it is possible to have one path nearly vertical, another more horizontal.

A third, but probably more difficult method would be to generate the seismic source at the bottom of the bore hole, and receive simultaneous signals on the surface at widely separated X offsets. Again the differential time of arrival will be a useful measure for extracting γ .

It is also quite possible that in certain geophysical applications, the precise value of γ is not directly useful, but instead the velocity parameters α and β are the important measures. In this case, the measurement technique considered in this paper is completely appropriate for a robust recovery of α from measured time of travel. The most accurate recovery of α occurs for the most horizontal ray paths; that is, for large offsets X . The minimization algorithm may be applied directly to the exact formula for α^2 for a robust computation of this parameter.

8 Conclusions

We have obtained an exact mathematical description of a geoseismic signal propagating through an anisotropic medium using a constant coefficient wave equation as the basic model. This model captures exactly the elliptical velocity profile required in the formulation of the geophysical model from which we obtained exact formulas describing the traveltimes through a two layer geological structure, and an exact inversion formula for computing the anisotropic velocity parameter γ . A robust numerical method based on a minimization technique was presented as an accurate method of computing both traveltimes and the inverted γ .

The exact formulas and robust numerical methods are significant improvements over the approximations and root finding methods discussed in the background material, and we note our formulation is no more difficult than these background methods.

We derived asymptotic formulas valid for the near vertical case, which describe accurately the high sensitivity of γ to the input parameters in this case. Our numerical work also confirms this sensitivity, even using exact formulas and robust numerical methods.

We conclude that the computation of the anisotropic velocity parameter γ for the given physical measurements from a series of surface signals and single borehole receiver is intrinsically unstable. By changing to the α, β velocity parameter space, we obtain an inversion method that is much less



sensitive to input errors. For certain geophysical problems, the α, β parameters may suffice for an accurate description of the material.

When the anisotropic velocity parameter γ is needed directly, a different measurement technique is required. This route will require further investigation, and we have proposed a number of promising possibilities involving a differential time measure.

