Novel Techniques for Acoustic Emission Monitoring of Fatigue Fractures in Landing Gear

James John Hensman

A thesis submitted to the University of Sheffield for the degree of Doctor of Philosophy in the faculty of Engineering.

February 2009

Dynamics Research Group
Department of Mechanical Engineering
The University of Sheffield

ABSTRACT

Acoustic Emission (AE) is a technique for performing Non-Destructive Evaluation (NDE) of structures, whereby ultrasonic transducers are placed upon the surface of the structure in order to detect ultrasound resulting from damage-related events within the structure. Unlike many other NDE techniques, AE is an entirely *passive* process: the transducers operate in receive mode only.

Advances in the sophistication of AE equipment and computer hardware in general mean that it is now possible to perform AE tests involving many sensors over a large structure whilst recording and storing every *waveform* received at every sensor. There is much interest in using this data to perform detailed analysis of structural integrity, particularly because AE testing is entirely non-invasive and can be largely automated.

There are many hurdles to be overcome before AE can be routinely used in such a situation: the quantity of data is huge, and it is of a format which is unusual to most engineers - AE signals comprise short, transient, 'burst' like signals. A significant portion of this thesis is given to considering what to do with all the data, and how it may be understood.

The thesis focusses on the development of a data processing method, and its application to landing gear certification tests. The methodology is designed to be generic, in that it could find application in the on-line monitoring of a variety of engineering structures in the aerospace, automotive and civil-engineering sectors.

Table of Contents

ACKNOWLEDGEMENTS

This work would not have been possible without the help, advice, support and cajoling of a number of people, and I am pleased to be able to thank some of them here.

Firstly I am indebted to my supervisor Professor Keith Worden, for his advice, encouragement, and expertise which he has lent me over the last three years. His supervision has allowed me to learn freely and think clearly, whilst his clear-headed guidance filtered some of the chaos from my trajectory.

Thanks must also go to Dr Gareth Pierce of Strathclyde University, who's advice, encouragement and sheer technical know-how has proved invaluable.

At Cardiff University, I am indebted to Dr Rhys Pullin, Dr Mark Eaton, Professor Karen Holford and Dr Sam Evans. Their work on the practical side of this project has been enormous, and without it my contribution would have been impossible.

This work was funded by EPSRC grant number GR/T26757/01 with the assistance of Messier-Dowty. The project was steered by representatives from Physical Acoustics, the National Physical Laboratories and Messier Dowty.

This document was created in LaTeX on the Gnu/Linux operating system, and the work herein would not have been possible without the use of free software. A thousand thank-yous to everybody involved.

To all of my friends, especially Rob, Dickie, Josephine, Mike; to the folk in the office, Frank, Will, Vaggelis, Giovanni, Maria; to the rest of the Dynamics group including Rob, Chian, Daley and Benny: thank you for three years of friendship and camaraderie, may there be many more.

Finally, to my parents: your love and encouragement has let me wander happily along the meandering path of education, for this I am eternally grateful; and to my wonderful girlfriend Amy, thank you for your love and seemingly endless patience.

PUBLICATIONS

Journal Papers

1. Hensman (J.), Pullin (R.), Eaton (M.), Worden (K.), Holford (K.M.), & Evans (S.L.) 2009. *Measurement Science and Technology* vol. 20 pp 1–10. Detecting and identifying artificial acoustic emission signals in an industrial fatigue environment.

2. Hensman (J.), Mills (R.), Pierce (S.G.), Worden (K.) & Eaton (M.) 2008. *To appear in Mechanical Systems and Signal Processing.* Locating acoustic emission sources in complex structures using Gaussian processes.

3. Leeds (J.S), McAlindon (M.E.), Grant (J.), Robson (H.E.), Morley (S.R.), Donnelly (M.T), Hensman (J.), Worden (K.) and Sanders (D.) 2008. *Submitted to The Lancet.* Predicting outcomes following gastrostomy insertion using the Sheffield Gastrostomy Score. A prospectively devised scoring system with validation and analysis with artificial neural networks.

4. Pullin (R.), Eaton (M.), Hensman (J.), Holford (L.), Worden (K.) and Evans (S.) 2009. *To appear in Strain.* Principal component analysis of acoustic emission signals from landing gear components: an aid to fatigue fracture detection.

5. Holford (K.M.), Pullin (R.), Evans (S.L.), Eaton (M.J.), Hensman (J.) & Worden (K.) 2009. *Journal of Aerospace, in press.* Acoustic emission for monitoring aircraft structures.

Refereed Conference Papers

1. Hensman (J.) & Worden (K.) 2008. *Proceedings of the 5th International Conference on Condition Monitoring and Machinery Prevention Technologies, Edinburgh, 2008.* A framework for detecting fatigue fractures using acoustic emission.

2. Hensman (J.), Pullin (R.), Eaton (M.J.), Worden (K.), Holford (K.M.) & Evans (S.L.) 2008. *Proceedings of 6th BSSM International Conference on Advances in Experimental Mechanics, NPL London* pp.251-260. Detecting and identifying artificial acoustic emission signals in an industrial fatigue environment.

3. Pullin (R.), Eaton (M.J.), Hensman (J.), Holford (K.M.), Worden (K.) & Evans (S.L.) 2008. *Proceedings of 6th BSSM International Conference on Advances in Experimental Mechanics, NPL London* pp.41-47. A principal component analysis of acoustic emission signals from a landing gear component.

4. Hensman (J.) & Worden (K.) 2007. *Proceedings of 13th International Conference on Experimental Mechanics, Alexandroupolis* on CD. Wavelet based feature extraction for acoustic emission.

5. Hensman (J.), Cristodaro (C.V.), Pierce (S.G.) & Worden (K.) 2006. *Proceedings of 27th European Conference on Acoustic Emission Testing, Cardiff* pp.117-124. On the reproducibility of transducer coupling for acoustic emission testing.

6. Pullin (R.), Hensman (J.), Holford (K.M.), Worden (K.) & Evans (S.L.) 2007. *Proceedings of DAMAS 2007, Torino* pp.139-144. A principal component analysis of acoustic emission signals from a landing gear component.

Conference Papers

1. Hensman (J.), Mills (R.), Pierce (S.G.), Worden (K.) & Eaton (M.) 2008. *Proceedings of 23rd International Conference on Noise & Vibration Engineering, Leuven 2008.* Locating acoustic emission sources in complex structures using Gaussian processes.

2. Spencer (A.B.), Worden (K.), Pierce (S.G.), Hensman (J.), Chetwynd (D.) & Staszewski (W.J.) 2008. *Proceedings of 23rd International Conference on Noise & Vibration Engineering, Leuven.* Force characterisation of a laser impulse using differential evolution with a local interaction simulation algorithm.

3. Hensman (J.), Worden (K.) & Mohd-Saleh (S.) 2008. *Proceedings of 4th European Workshop on Structural Health Monitoring, Krakow.* Feature extraction from acoustic emission waveforms using a discrete wavelet transform.

4. Chetwynd (D.), Worden (K.), Spencer (A.) & Hensman (J.) 2008. *Proceedings of the 26th International Modal Analysis Conference, Orlando, Florida* p.12. The effects of uncertainties within acoustic emission modelling.

5. Leeds (J.S.), Morley (S.R), Robson (H.E.), Grant (J.), Hensman (J.), Worden (K.), McAlindon (M.E.), Sanders (D.). 2009. *Poster at GUT Gastroenterology conference.* Comparison of the Sheffield gastrostomy score with an artificial neural network analysis.

Technical Reports

1. Holford (K.M), Evans (S.), Pullin (R.), Eaton (M.), Worden (K.), Hensman (J.), Baker (T.), Bond (D.), Cole (P.), Bradshaw (T.) and Theobald (P.) 2008. *Technical report submitted to Messier-Dowty.* Characterisation of damage in complex structures by acoustic emission during fatigue testing (methodology document).

2. Holford (K.M), Evans (S.), Pullin (R.), Eaton (M.), Worden (K.), Hensman (J.), Baker (T.), Bond (D.), Cole (P.), Bradshaw (T.) and Theobald (P.) 2007. *Technical report submitted to Messier-Dowty.* Characterisation of damage in complex structures by acoustic emission during fatigue testing (progress review Jan. - Dec. 2007).

List of Figures

List of Tables

Chapter 1

INTRODUCTION

The technique of Acoustic Emission (AE) holds much potential as a non-destructive technique. It has been proven to be a useful method in several areas such as the monitoring of pressure vessels [1], but is still not considered a reliable technique capable of competing with established technologies such as ultrasonic scanning, eddy-current inspection or magnetic particle inspection. This is due in part to the huge amount of incomprehensible data acquired during testing, which is infathomable to all but the most experienced AE user. This thesis sets out to provide a series of tools, techniques and recommendations in order to make sense of the data. Whilst the focus lies on detecting fractures in landing gear components, the techniques discussed should be applicable to many engineering structures.

This chapter begins with a short introduction to the field of acoustic emission in order to give the reader some insight into the technological status of the field at the beginning of the project, and to introduce the reader to some key concepts. Next, a project focus section sets out some clear goals for the project, and describes the intended application of the technology developed. The chapter closes with an overview of the coming chapters.

1.1 What is Acoustic Emission?

Various physical events which occur in structures result in sudden redistributions of stress: this stress results in an elastic wave which propagates through the structure,

and can be detected by surface-mounted transducers. These bursts of energy are called *acoustic emissions*, and the technique of detecting and recording them is called Acoustic Emission (AE).

One eloquent AE researcher has informally referred to AE as *"little earthquakes in structures"*: the comparison is quite accurate. In an earthquake, friction and slipping of tectonic plates in the earth's crust give rise to vibrations, which travel as sound waves through the earth and shake structures at the surface. Acoustic emission is very similar, though the frequencies of the vibrations involved is much higher, and the aim is to detect and interpret the vibration of surface mounted objects (usually some form of piezo-electric transducer) in order to assess the cause of the emission, and potentially asses the damage-state of the structure.

The causes of acoustic emissions are numerous. In composite structures, some causes include fibre breakage, matrix cracking and fibre–pullout [2]. Analysis of these acoustic emissions can lead to the identification of the different source types; this is the subject of extensive research (see for example Giordano et al. [3]). In metals, many acoustic emission sources are application specific, for example fragment breakage in wire rope [4, 5], or tribological mechanisms in bearings [6, 7]. This project is concerned with the detection of fatigue fractures using the acoustic emission technique.

1.1.1 Some History

The beginning of acoustic emission research can be attributed to Kaiser [8], whose PhD thesis on acoustic emission in tensile testing gave rise to the first research into the engineering applications of AE. Early research into acoustic emission was limited by the hardware available; it was not possible to record waveform data due to lack of high speed acquisition systems and storage requirements. Much research relied on AE 'count rate', a simple count of the number of times that the voltage crosses some threshold.

With advances in acquisition hardware, computing power and digital storage requirements, it is now possible to capture and analyse waveform data from acoustic emission tests. AE hardware records waveforms at 40MHz, and is capable of storing many waves across many channels in quick succession. This presents the opportunity to analyse the received signals in greater detail.

For many years, acoustic emission researchers lacked sufficient computing power to analyse their data. Now that such computing power is cheaply available, the question remains as to which algorithms and techniques to use in order to make use of it.

1.2 Landing Gear

Landing gear is the structure that supports the aircraft on the ground, allowing it to taxi, as well as playing a critical role in the safe landing of aircraft. This mixed usage of the component places significant demands on the engineering designers – the landing gear lust be able to support bending and torsional loads during taxi and take-off, as well as massive impact load during landing [9]. In addition, the engineer is constrained to minimise weight, leading to structures with little or no structural redundancy.

The arrangement of landing gear are generally twofold [9]: the conventional or *taildragger* gear consist of two wheels at the front of the aircraft and a smaller wheel (or occasionally, for light aircraft, a skid) at the rear. The *tricycle* undercarriage consists of two wheels (or sets of wheels) under the wings and a support wheel at the front. Additionally, some aircraft have a small wheel at the rear to avoid tail strikes during take off. As aircraft increase in size, they require more wheels to cope with the increasing weight: modern passenger aircraft have sets of wheels arranged on a bogie, usually four or six wheels arranged to adjust to the aircraft's orientation.

1.2.1 Landing Gear Monitoring

Landing gear is subject to extensive structural health monitoring, due in part to the "safe-life" design method. In contrast to the rest of an aircraft structure, the landing gear is damage intolerant, and a monitoring system is vital. Modern aircraft fly with systems installed to monitor tire pressure and brake temperature [10].

Tire monitoring systems are used to eliminate the need for periodic manual inspection, and to avoid incidents due to unpressurised tires, whilst brake temperature sensors report the current brake temperature to the pilot in the cockpit. Monitoring of brake temperature can be used to determine remaining brake life [10].

The monitoring of the shock-strut (main tube fitting) of the gear is an area of ongoing research. Of particular interest is the measurement of nitrogen gas-pressure within the strut. Current procedures are either inaccurate due to the complicated nature of the system; or else unweildly, time consuming and expensive, often including taking the aircraft out of service. ultrasonic approaches have been suggested to alleviate these problems [11, 12].

Monitoring the fatigue condition of landing gear is an important yet difficult problem. Some magnetic techniques have been developed for the determining of fatigue state of steels, but these are not yet suitable for commercial use [10]. An alternative approach is known as *Loads Monitoring*, where the loads into the landing gear are measured throughout it's service life (either directly or by inference from the aircrafts state), and compared to either a model of the fatigue life or the measurements from the fatigue testing of a similar component.

1.2.2 Landing Gear Research

Research in the area of landing gear is ongoing, as engineers strive to design safer, lighter, more cost-efficient landing gear. With the modern foci of environmental impact and noise reduction, the optimal solution is a moving target. The modern design phase of landing gear includes extensive modelling and simulation, particularly at the concept stage [13], in order to evaluate new design concepts.

Landing gear is a contributor to aircraft noise - particularly as they are deployed during the take off and landing phase of flight, where aircraft noise has the most impact. Previously, attempts have been made to reduce the aero-acoustical noise by the addition of fairings [13], though current work involves using noise reduction as a design driver, in order to produce landing gear that are low-noise without the addition of extra components.

Materials research is key to the future of landing gear. Traditionally, ultra high tensile strength (UHTS) steel such as 300M [13] has been used, which is extremely strong but lacks ductility (and hence the need to detect even the smallest fatigue fracture). Future directions include the use of composite components where appropriate (the Boeing B787 will feature organic matrix composite braces) as well as titanium components and titanium matrix composites.

1.3 About this Project

This thesis forms part of a larger project, with the aim of developing a system to monitor landing gear structures during fatigue testing. The project involves various researchers at Cardiff university, as well as work at Sheffield. The project is jointly funded by Messier Dowty Ltd. and the Engineering and Physical Sciences Research Council (EPSRC). The project receives input from advisors at the National Physical Laboratory (NPL) and Physical Acoustics Corporation UK. This side of the project aims to develop the signal processing portion of the project, by use of various techniques including wavelet transforms and supervised and unsupervised forms of machine learning.

1.3.1 Project Focus

The objective of the project is to detect crack growth during fatigue testing of landing gear. This testing is performed as part of the certification process for the structure: the gear is tested in a large facility in a loading cycle designed to represent its service life. A series of hydraulic actuators apply loads to the structure to simulate taxiing, take-off and landing; various cycles are used to simulate different working conditions.

A typical test lasts two years. During this time, the structure will be periodically removed from the rig and inspected for damage using dye-penetrant and magnetic particle inspection. This is a time consuming and costly procedure. If parts of the structure have been found to fail, they must be replaced and the test re-started.

This project aims to develop an early warning system using AE. A series of AE sensors are to be attached to the structure during testing. Upon detection of a fatigue fracture, the system can be dismantled and the diagnosis confirmed using current methods. Further, location of the detected fracture will allow easier analysis.

Figure 1.1 shows a schematic diagram of the main section of an Airbus A320 landing gear, which is one of the structures in question. It is a large structure, with many differing sections: this will require a large array of AE sensors for monitoring, which in turn will produce a huge amount of data.

The complicated loading cycle to which the structure is submitted necessitates a

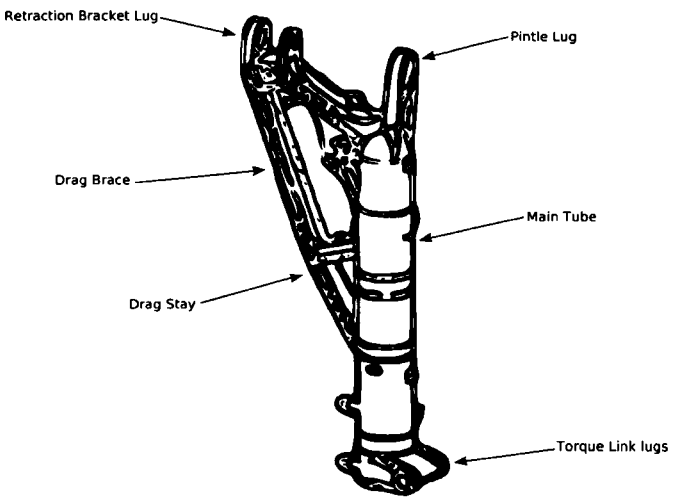


Figure 1.1: Schematic of the Landing Gear from an Airbus A320, showing the main tube and various lugs

large number of moving parts. These will cause acoustic emissions which will be detected by the AE monitoring system. Such mechanisms include fretting, rubbing, bearing noise and load–train noise. These acoustic emissions will be large in number, far outweighing the number from a potential fracture. One of the main objectives of this project will therefore be to discriminate between AE relating to crack growth, and unwanted or 'noise' AE.

Examining Figure 1.1, it is clear that wave propagation will not be simple in the structure due to the complicated geometry. The simple location algorithms employed routinely in AE testing are unlikely to be sufficient here: another objective of this project will be to properly locate acoustic emissions in complex structures.

Unlike other methods such as ultrasonic inspection, AE tests are not repeatable. AE events occur due to some physical mechanism in the structure, resulting in permanent change: the AE engineer gets only one chance to record the data. An important issue therefore is reproducibility of the AE equipment: how much impact changes in the AE equipment affect recorded results, and whether observed differences in the data can be attributed to changes in the equipment or in the AE mechanisms present.

Before undertaking to develop algorithms and techniques to deal with AE data, it is necessary to develop a good understanding of the AE process. This will enable

comprehension of received waveforms, and should give insights into what kind of model is appropriate to meet the objectives. This is the subject of Chapter 2.

The rest of this thesis is organised as follows. Chapter 2 discusses the physics of AE, with an emphasis on wave propagation. Chapter 3 outlines the proposed model for detecting fatigue fractures. Chapter 4 deals with the issue of performing feature extraction on captured AE waveforms, whilst Chapter 5 is entitled visualisation, and discusses methods for visually examining AE data. Chapter 6 tackles the source location problem. Chapter 7 contains examples of the methodologies developed in this thesis applied to several real AE tests, and discusses the future direction of this work.

1.3.2 Aims and Objectives

1. Study wave propagation in complex structures.
   Wave propagation should be studied in order to gain an understanding of the nature of AE signals, how they vary, and what kind of signal processing may be required.

2. Overcome problems with location AE in complex structures.
   Location is a potential strength of AE: the ability to do so on a complicated structure would be immensely useful.

3. Devise a system capable of handling large quantities of AE data.
   Any methodologies developed should be simple enough to run in real time, and the quantities of data acquired during an AE test must be considered.

4. Employ novel signal processing techniques and pattern recognition to distinguish between background 'noise' AE signals and those from fractures.
   There is significant potential for using machine learning tools to solve some of the problems encountered in AE - this potential should be investigated

5. Devise methodology of presenting AE data to the user in a simple form.
   This will involve using dimension reduction techniques for visualisation of data.

Chapter 2

The Physics of AE

This chapter examines the AE process, starting with an overview of the AE forward model. This can be broken into three sections: the source; wave propagation and detection. Different types of AE sources are examined, including fractures, frictional and artificial sources. An overview is given of wave propagation mathematics, resulting in discussion of the use of dispersion curves. The effects of the different parts of the AE forward model on detected AE waveforms is discussed, and methodologies for approaching the source identification problem are addressed. The careful consideration of the AE process presented in this chapter leads to decisions about methodology in several following chapters.

2.1 The AE Forward Model

The AE process will now be presented as a forward model, with reference to the work of Wilcox et al. [14]. The aim of examining the AE process in this way is to examine how different aspects of the model influence the signal measured at the AE equipment. Understanding of the AE process and the identification of uncertainties therein is used to build the data-driven method (Chapter 3) and is considered in the proposed source location algorithm (Chapter 6).

The forward model as described by Wilcox et al. [14] has three main stages as illustrated in Figure 2.1: the source, wave propagation and detection. The following text describes each of these stages and details the factors which contribute to each.

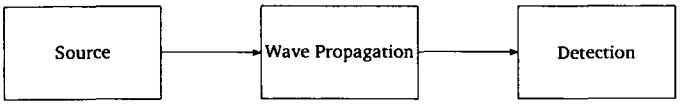


Figure 2.1: A forward model for acoustic emission

The Source The source part of the forward model is related directly to the *cause* of the event. Perfect knowledge of the source would make acoustic emission an infallible non-destructive technique: unfortunately information about the source is usually obscured by the effects of wave propagation and detection. In the AE forward model, the source is characterised by the following:

Position Detection of the position of the source is the greatest strength of AE; using multiple transducers, one can estimate the position of the source from the time differences of the signals received (See Chapter 6). The precise position of the source has a significant effect on the received signals due to the nature of the wave propagation (see next section). Chetwynd et al. [15] considered the effect of uncertainty with respect to the source position, and found significant variation of the received signals from small changes in the source-sensor distance.

Depth The source depth is the vertical distance from the surface of the structure to the point of the AE event. In the common situation where AE is used to monitor a plate-like structure, the depth of the source has a significant impact on the received signals. This is due to the ability of sources of different depths to excite different *modes* of the wave propagation equation. This was discussed by Hamstad et al. [16] and Dunegan [17], who show that source depth can contribute to different modal amplitude ratios (see Section 2.4).

Orientation and beam pattern The orientation of the source is important in a similar way to the depth: different source orientations can excite different wave modes [18, 19]. For example, it can be demonstrated that for plate-like structures, in-plane sources produce largely symmetric wave modes, whilst out-of-plane sources produce largely antisymmetric wave modes. This results in the *envelope* of the received waveform being highly dependent on the orientation of the source.

Time signature Consider the force exerted by some AE source on the surrounding material; one is concerned with the function of the magnitude of the force over time. Of particular interest is the rise-time of the source, since this defines the bandwidth of the signal and hence influences the detectability. Due to the nature of the wave propagation in most interesting structures, the time signature of the source can be an elusive part of the forward model, and one is forced to make assumptions given the type of source (we postulate on the nature of fracture sources later). Spencer et al [20] shows a promising method of numerically calculating the source time signature from experimental data using a finite difference approximation to the wave equation.

Frequency content Clearly the frequency content of the source is related to the time signature; however the frequency content of recorded signals is often dominated by the response of the transducer.

Wave propagation The energy release of an AE event generates a stress wave throughout the structure: information about the source of the AE event is transmitted through the structure in the form of ultrasound. There follows a concise description of wave propagation, categorised into bulk, Lamb, Rayleigh and Stonely waves. A more mathematical description and discussion of Lamb waves follows in Section 2.3.

Material properties The speed of sound propagation in a structure is dependent of the material properties of that structure: the material properties E (Young's modulus) and v (Poisson's ratio) can be related to wave-speeds via the Lamé constants: See appendix A.1.

Reflections and mode conversions Waveforms recorded by AE equipment are likely to contain reflections from structure boundaries in all but the simplest of cases. This is the first of three ways in which the *geometry* of a structure influences the observed acoustic emission.

Attenuation The further the distance an acoustic wave has to travel before reaching a receptor, the smaller the amount of energy transmitted to the receptor. this is due to the 'spreading' of the energy (over the surface of a sphere for 3D cases, over the circumference of a circle for 2D cases) and also due to some of the energy being dissipated as heat as the particles in the medium are displaced. Good knowledge of the attenuative properties

of a material are important for deciding upon sensor-spacing for an AE test.

Wave type The type of wave that propagates in a structure is highly dependent on the geometry of that structure, not only in terms of reflections and attenuation, but also in terms of the *type* of wave which can propagate.

- Bulk waves
  Bulk waves, as the name implies, occur in structures where the dimensions of the propagation medium are (much) larger than the wavelength of the wave. This means that the structure's geometry only interferes with the wave propagation at boundaries, in the form of reflections. Bulk waves are non-dispersive, meaning that all parts of the wave travel at the same speed: c_p for longitudinal waves and c_t for shear waves (these are directly related to the material constants, see appendix A.1.). Bulk waves are familiar to the reader: speech consists of longitudinal bulk waves, and ultrasonic bulk waves are commonly used in medical imaging. The nature of interesting engineering structures means that bulk waves are rarely encountered in an acoustic emission context.
- Lamb waves
  Lamb waves occur in plates where the thickness of the plate approaches the wavelength of the wave. Lamb waves are useful in a NDT context because they propagate long distances (the energy is spread over the circumference of a circle, not the surface of a sphere, and so attenuation is lower). Lamb waves are *dispersive*, meaning that different parts of the wave travel at different speeds; dispersion contributes to the uncertainty in the model, effectively 'amplifying' uncertainty associated with source position [15].
- Rayleigh waves
  Rayleigh waves travel across the surfaces of bulk materials: they are similar to lamb waves in that the structure directly interferes with the wave propagation.
- Stoneley waves
  Stoneley waves occur where two different materials meet: they commonly occur along welds and other such features. With welds being a prime defect site in many structures, Messier Dowty landing gear are

manufactured by machining from a solid piece. Whilst Stonely-like waves may interfere with the waves received during an AE test, they are not the primary from of wave propagation during an AE test on landing gear, and the reader is referred to Rose [21].

Detection Once the energy emitted at the source has propagated to the sensor, the detection section of the AE forward model converts tiny surface displacements to a voltage, and eventually a series of floating point numbers on a computer.

Acoustic couplant An acoustic couplant is used to maximise energy transfer from the structure to the transducer. It does this by filling the small air gap between the structure and the transducer, allowing vibrations in the structure to be transferred in to the transducer. A good acoustic couplant should match the acoustic impedance of the structure and the sensor: water based gels ('Sonagel') of the type used in medical imaging are commonly used, but some oil-based couplants are also suitable for engineering purposes.

Spatial transducer aperture The transducer agglomerates surface displacements over its sensitive area into a single voltage. Wilcox et al. [14] gives a mathematical model for the aperture effect.

Transducer frequency response The frequency response of a transducer depends on the physical dimensions of the active element, any physical damping used and any analogue circuitry in the transducer. To achieve broad-band responses, manufacturers often use multiple active elements with complementary bandwidths: the sensitive range of a transducer is provided as a data sheet by the manufacturer.

In-plane and out-of-plane sensitivity Both in-plane and out-of-plane surface motions are coupled into the transducer by the acoustic couplant. Theobald and Dar [22] discuss a methodology for characterising the in-plane and out-of-plane sensitivity of a transducer.

2.2 Sources of AE

This section describes different AE sources in the context of the preceding section. It was stated in Chapter 1 that one challenge in this project was to be in distinguishing

between different sources of AE: some of these are described here, with discussion of how their characteristics might give rise to differences in detected signals.

Fatigue fracture sources are discussed first, followed by frictional AE sources. The section concludes with artificial sources, of which good use is made throughout the remainder of this thesis.

2.2.1 Fatigue Fractures as AE Sources

It has already been stated that energy releases at growing cracks in structures create acoustic emissions which are detectable. This section considers the mechanisms by which energy is released, and how it might influence detected AE waveforms.

Stress Release Mechanisms

Stress is released by a fatigue fracture in three main ways:

- **Crack extension** As the crack grows, stress is released at the growing tip
- **Plastic region extension** The plastic region around a crack tip grows with the crack. The extension of this region correlates to a release of stress
- **Crack Face Rubbing** On closing of the crack, the crack faces rub together, causing AE.

The literature provides no definitive answer as to which stress release mechanism (if any) dominates the generation of AE waves. Hamel et al. [23] gives careful consideration to potential stress release mechanisms for crack tips, and concludes that experimental data can only be explained by the presence of several mechanisms. The presence of crack extension mechanisms and crack face rubbing (or 'grinding') are discussed by Berkovits and Fang [24], who also postulate that the mechanisms should be discriminable using the captured waveforms.

Despite a variety of mechanisms for stress release mechanisms, note that all the mechanisms are *localised*, since they all in the immediate vicinity of a growing crack.

Opening Modes

Fatigue fractures are classified into three types [25]: tensile opening (mode i); shear sliding (mode ii) and shear tearing (mode iii), as shown in Figure 2.2. There are

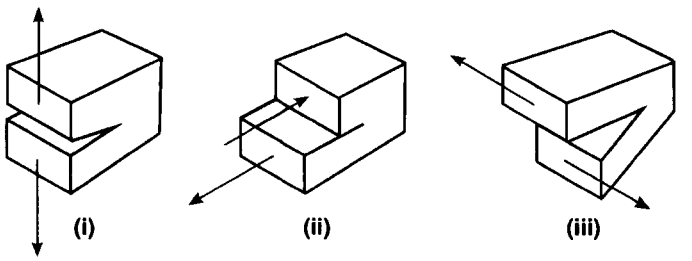


Figure 2.2: Crack opening modes: (i)Tensile, (ii)Shear sliding and (iii) Shear tearing

several things of note here: different opening modes will lead to different prevailing stress release mechanisms: a mode ii or iii crack will be more prone to crack face rubbing than a mode i crack. Also, without changes to the loading regime, the opening mode will remain constant, meaning that this element of uncertainty will remain constant for the duration of any simple test: i.e. whichever mechanism of stress release dominates, this will be constant for the duration of a test.

Paris' Law

The relationship between crack growth rate and stress intensity at a crack tip is given by the Paris law [26]. The Paris law states:

$$\frac{da}{dN} = A(\Delta K)^{\alpha} \tag{2.1}$$

Where: a is the crack length, N is the number of cycles, ΔK is the stress intensity range and A and α are material constants.

Berkovits and Fang [24] discusses the relationship between the acoustic emission *count rate* and the Paris law. Various AE researchers including Hamel et al. [23] and Lindley et al. [27] propose the following relationship:

$$\frac{d\eta}{dN} = B(\Delta K)^{\beta} \tag{2.2}$$

where η is the AE count and B and β are constants. Substituting equation 2.1 and 2.2 for ΔK,

$$\frac{d\eta}{dN} = C(\frac{da}{dN})^{\gamma} \tag{2.3}$$

where $C = \frac{B}{A^{\gamma}}$ and $\gamma = \frac{3}{\alpha}$. The implication of this is that the AE count rate is proportional to crack growth. Whilst this has no effect on individual waveforms *per se*, it implies that a *series* of waveforms relating to a growing crack will have ever increasing count rate. This is an important phenomenum, and is utilised is the methodology Chapter to detect crack-like activity (though the parameter measured is the cumulative energy rate).

Crack Orientation

The orientation of a crack plays an important role in determining the structure of emitted AE waves, as discussed by Pullin et al. [19]. For plates (or plate-like structures), the ratio of the amplitudes of the A_0 and S_0 modes is determined by the orientation of the source (see Section 2.4 for a description of wave modes). Sources which are in-plane generate mostly S_0 waves; out-of-plane sources generate mostly A_0 waves. An experiment devised by Gorman and Prosser [18] can be used to demonstrate this, whereby a plate with an AE transducer affixed to the top surface has on edge machined to a fixed angle. Sources are generated on this edge, and the edge is re-machined to a new angle. The changes in modal amplitude ratios are observed to correlate with the angle of the plate edge.

Finite Element Models

Several authors have attempted to model fatigue fractures (and the resulting AE wave propagation) using finite element modelling. Hamstad [28] thus considers the differences between monopole and dipole sources, and warns researchers to be aware of the differences: A pencil lead fracture is a *monopole* source, whereas growth of a fatigue fracture is a *dipole* source (see Section 2.2.3). Lee et al. [29] modelled crack growth by the release of Finite Element nodes under stress at the crack surface. This technique was used to study the relationship between crack depth and the A_0, S_0 ratio. Whilst finite element modelling is a useful tool in understanding the wave

propagation resulting from crack extension, modelling of all the uncertain aspects discussed here would be impossible.

2.2.2 Frictional AE Sources

Frictional sources of AE have not received a great deal of attention in the literature. Boness et al. [30] studied the relationship between AE energy and wear mechanisms for a sliding contact. Fault detection of bearings [31, 32, 33] has proved successful, but no attempt is made at determining or understanding the underlying frictional mechanisms.

A glance through the tribology literature reveals the following tribological mechanisms which may give rise to acoustic emissions [34, 35]:

- **Adhesion.** At the junction between two contacting surfaces, bonding occurs between surface asperities. Relative motion of the surfaces causes the bonds to break, and new bonds to form.
- **Impact wear.** If the impact energy between two surfaces is high, energy is absorbed through elastic and plastic deformation, leading to subsurface cracking and debris formation.
- **Erosion.** Errosive wear is caused by the impingement of particles of a surface. It is often a problem in transport of powders and slurries.
- **Abrasive wear.** The motion of wear debris or hard asperities relative to a surface causes abrasive wear.
- **Oxidative wear.** If a material is capable of fast oxidation, then oxide layers can be removed by friction of a contacting body, causing further oxidation of the revealed surface.

Comparing this to the list of stress releases from a fatigue fracture above, it is clear that there are many more different mechanisms by which a frictional AE source can release energy. Also, the frictional source is not localised: acoustic emissions can be generated anywhere along the contact. The nature of the position of the source exhibits a high degree of *uncertainty*.

Further, the nature of tribological contact is that the dominant mechanism for wear (and therefore the dominant source of acoustic emissions) can change with time, due to differing lubrication conditions, material hardening, the presence of third bodies and degradation of surface treatments. The mechanisms compete and interact, meaning that the very nature of the source (in terms if frequency content and time signature) is uncertain.

2.2.3 Artificial Sources of AE

Artificial acoustic emissions can be generated by various means, and are extremely useful for studying the behaviour of ultrasonic waves in the controlled conditions of the laboratory. There follows a description of each of four types of artificial AE used in this project.

The Hsu-Nielsen source

The Hsu-Nielsen pencil lead break is used extensively by AE researchers for sensor sensitivity checks, and as a means of quickly generating AE data.

A special propelling pencil (see Figure 2.3) is gently levered against the surface of the structure of interest until the pencil lead fractures. This generates a monopole, point, step source, with orientation perpendicular to the surface.

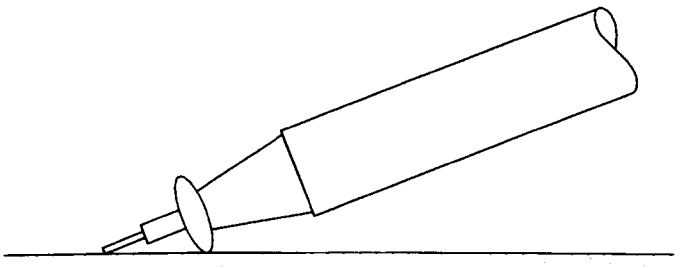


Figure 2.3: Schematic showing a pencil used for generating a Hsu-Nielsen source. The pencil is levered against the surface until the lead fractures, releasing strain energy into the surface

The Hsu-Nielsen source (so called because it was devised separately by Hsu [36] and Nielsen [37]) is often used to ensure that AE transducers are correctly attached and

working (as per ASTM standard E97694 [38]) prior to a test; this is some times incorrectly referred to as 'transducer calibration'. The source is particularly useful as it is cheap and portable, despite poor reproducibility.

A Pulsed Transducer

Using an AE sensor to *transmit* ultrasound is a simple method of generating artificial acoustic emission, giving rise to a highly repeatable source. The technique provides an out-of-plane monopole source, similar to the Hsu-Nielsen technique, but with a finite aperture (the Hsu-Nielsen technique approaches a point source). The bandwidth of the transmitted signal is obviously limited to the bandwidth of the transducer being used, but the effects of this will not be noticed if this matches or exceeds the bandwidth of the transducer being used as a receiver. This technique was utilised in Section 7.1 to generate fracture like sources in a test environment.

Thermoelastic Expansion by Laser

Artificial acoustic emission can also be generated by means of a pulsed unfocused laser. The laser pulse momentarily heats the surface of the plate (pulse durations are typically $\approx 10ns$) which then cools quickly creating a stress wave in the structure. Wu et al. [39] discussed the application of laser generated ultrasound in plates. This method of generating artificial AE events was utilised in Section 6.3 to collect training data for a machine-learning location algorithm. Thermoelastic expansion is an extremely repeatable source.

Focused laser

Focussing the laser beam onto the plate gives a point source of high intensity. This gives rise to a high amplitude artificial AE event, but unfortunately damages the surface of the material by ablation.

2.3 Wave Propagation

The aim of this project is to be able to distinguish between different sources of AE. As has been shown above, the signal recorded at the AE equipment is a combination of the source properties, wave propagation and properties of the detection equipment. The engineer has some control and (presumably) a good understanding of the detection side through sensor-selection etc., yet the nature of the source remains obfuscated by the nature in which the acoustic waves propagate. A good understanding of the wave equation will hopefully shed some light on this part of the AE process.

This section begins with the derivation of the vibration of a taut string as given by Rose [21], and by extension the standard wave equation. Vibrations are then constrained to some geometric boundary, specifically to the plate conditions, which gives rise to Lamb waves and the dispersion curves.

This chapter then closes with a consideration of how to deal with the problem of identifying AE sources in the light of the material presented.

2.3.1 The Wave Equation

Consider a small element ds of a taut string, subject to some force $q(x,t)$ as shown in Figure 2.4. A relationship between the displacement u and the applied forces along the length of the string can be obtained by resolving the forces in the u direction,

$$-T\sin\theta + T\sin(\theta + \frac{\partial\theta}{\partial x}\,dx) + qds = \rho\,ds\frac{d^2u}{dt^2} \tag{2.4}$$

For small deflections, $ds \simeq dx$, $\sin\theta \simeq \theta$ and $\theta \cong \frac{d^2u}{dt^2}$

$$-T\theta + T(\theta + \frac{d}{dx}\frac{du}{dx}\,dx) + q\,dx = \rho\,dx\frac{d^2u}{dt^2} \tag{2.5}$$

dividing through by dx and gathering like terms leaves

$$T\frac{d^2u}{dx^2} + q = \rho\frac{d^2u}{dx^2} \tag{2.6}$$

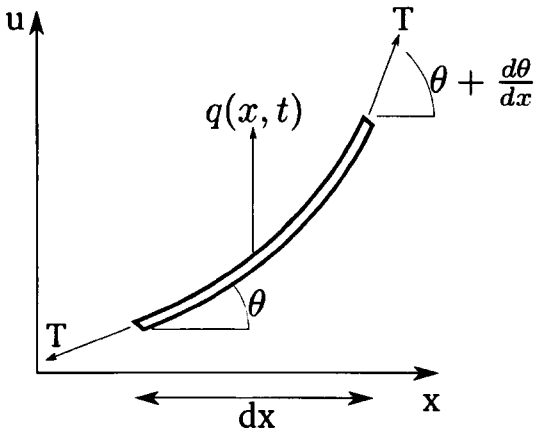


Figure 2.4: A taut string under tension T, subjected to a force q. Reproduced from [21]

Without a forcing term, this becomes

$$u_{xx} = \frac{1}{c_0^2}\ddot{u} \tag{2.7}$$

where u_{xx} is the second derivative of the displacement, u in the x direction, and $\ddot{u}$ is the second derivative of u with respect to time. $c_0 = \sqrt{T/\rho}$ is the velocity of the wave. Generalising this to three dimensions, we arrive at the Standard Wave Equation (SWE), which can be written as

$$\begin{aligned}
(\lambda+\mu)\frac{\partial}{\partial x}\left(\frac{\partial u_x}{\partial x}\frac{\partial u_y}{\partial y}\frac{\partial u_z}{\partial z}\right) + \mu\nabla^2 u_x + \rho q_x &= \rho\ddot{u}_x \\
(\lambda+\mu)\frac{\partial}{\partial y}\left(\frac{\partial u_x}{\partial x}\frac{\partial u_y}{\partial y}\frac{\partial u_z}{\partial z}\right) + \mu\nabla^2 u_y + \rho q_y &= \rho\ddot{u}_y \\
(\lambda+\mu)\frac{\partial}{\partial z}\left(\frac{\partial u_x}{\partial x}\frac{\partial u_y}{\partial y}\frac{\partial u_z}{\partial z}\right) + \mu\nabla^2 u_z + \rho q_z &= \rho\ddot{u}_z
\end{aligned} \tag{2.8}$$

where λ and μ are the Lamé constants for the material (see appendix A.1). A full derivation of the SWE is given in Rose [21]. Notice how equations 2.8 and 2.7 take the same form: the second derivative with respect to space, plus some forcing term, is proportional to the second derivative with respect to time.

The SWE is the basis for the LISA simulation software [40], where it is discretised

upon a grid in time and space, and the derivatives approximated by local differences in the grid (hence LISA is known as a *finite difference* method).

2.3.2 Bulk Waves

Without any boundary conditions, equation 2.7 (and by extension, equation 2.8) represents a freely propagating, or *bulk* wave, with solutions such as:

$$u = (A_1 \sin(kx) + A_2 \cos(kx))\,(A_3 \sin(kc_0t) + A_4 \cos(kc_0t)) \tag{2.9}$$

which can be quickly shown to obey equation 2.7. This is the type of wave which is most likely to be familiar to the reader: it constitutes speech, music and earthquakes, and is used extensively in medical imaging.

2.3.3 Mode Conversion

In a bulk structure, where the material in free to vibrate in three dimensions, there are two possible types of wave: longitudinal waves, where the direction of particle displacement is the same as the direction of wave travel, and shear (a.k.a. transverse) waves, where the direction of motion is perpendicular to the direction of wave propagation. Shear waves are often sub-categorised into horizontal- and vertical-shear.

Where one of these waves meets a boundary, it can be reflected as either the same *or other* type of wave. In fact, this is true of the various modes of guided waves discussed below. The type of wave reflected and the direction in which it travels can be determined by Snell's Law:

$$\frac{\sin(\theta_1)}{\sin(\theta_2)} = \frac{v_1}{v_2} \tag{2.10}$$

where θ_1 and θ_2 are the angles between the normal (line perpendicular to the boundary) and the incident and reflected waves respectively; v_1 and v_2 are the velocities of the waves.

Mode conversion can occur at the boundary of a structure - the interface between steel and air, or at joints in the structure where acoustic coupling between the

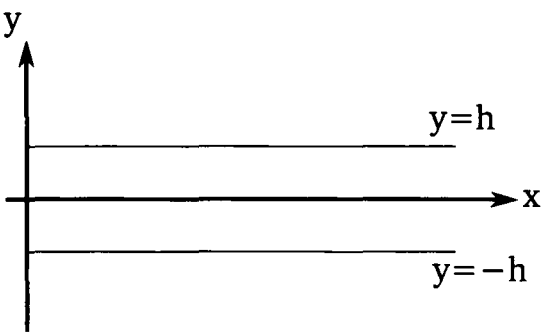


Figure 2.5: Schematic showing the boundary conditions for Lamb waves

sections is not absolute. In some structures, mode conversion can occur at welds.

In general, the phenomenum of mode conversion makes interpreting AE signals more difficult, and in this thesis, all analysis is performed on the early-arrival part of any signal, so as to avoid dealing with reflected signals as far as possible.

2.3.4 Guided Waves

If one is to constrain the propagation of a wave to some finite geometry (i.e. add boundary conditions to equation 2.8), one obtains a guided wave. When the boundary conditions comprise a plate, the resulting waves are known as *Lamb waves* after Sir Horace Lamb who solved the problem as early as 1917 [41]. Guided waves also include waves in rods, tubes, beams and meshes, as well as waves at the surface of a bulk medium. Since many interesting engineering structures are plate-like, Lamb waves are considered here in more detail.

Lamb waves

Matters are simplified by considering only the plane-strain condition, such that motion in the z direction is ignored. In order to constrain the vibration to a plate, the following boundary conditions are applied (2.11)

$$\sigma_{yy} = \sigma_{yx} = 0 \, at \, y = \pm h \tag{2.11}$$

This is illustrated in Figure 2.5. A full derivation is omitted here for brevity, and

the reader is referred to the original text [41], and detailed explanations in more modern texts such as [21] and [42]. Viktorov [43] gives an excellent practical and mathematical guide to applications of Lamb waves. Extensive manipulation of (2.11) and (2.8) yields:

$$\frac{\tan(qh)}{\tan(ph)} = \frac{4k^2pq}{(q^2-k^2)^2} \tag{2.12}$$

which is the dispersion relation for symmetric modes, and:

$$\frac{\tan(qh)}{\tan(ph)} = -\frac{(q^2-k^2)^2}{4k^2pq} \tag{2.13}$$

which is the dispersion relation equation for antisymmetric modes, where $p^2 = \omega^2/c_L^2 - k^2$, $q^2 = \omega^2/c_T^2 - k^2$, and $2h$ is the thickness of the plate. Numerical solution of these equations leads to the dispersion curves as shown in Figure 2.6. The dispersion curves plot both phase- and group-velocity against the frequency-thickness product. This is possible because high frequency waves in a thin plate behave in the same way as lower frequency waves in a thick plate, hence only the frequency-thickness product is required. Dispersion curves can be used to interpret the non-stationary multi-modal nature of AE waves.

The dispersion curves are well covered in the AE literature as a means for interpreting AE waveforms. Pullin et al. [44] demonstrated the application of dispersion curves to AE waves in plates experimentally; Hamstad [45] showed that the distance to source can be calculated from an understanding of waveform structure and the dispersion curves. Prosser et al. [46] showed that the dispersion curves can be utilised with a time-frequency analysis of the signal. The next section discusses the use of dispersion curves, and demonstrates the effect of dispersion on AE measurement.

2.4 Dispersion and AE Measurement

In order to illustrate the effects of dispersion on the received AE waveforms, consider the four signals shown in Figures 2.7 to 2.10 in conjunction with the dispersion curves in Figure 2.6. The figures show the time-trace of each signal (Voltage against time, top) as well as the time-frequency content of the signal (bottom). The continuous wavelet transform was used to calculate the lower plots (see Section 4.2.1).

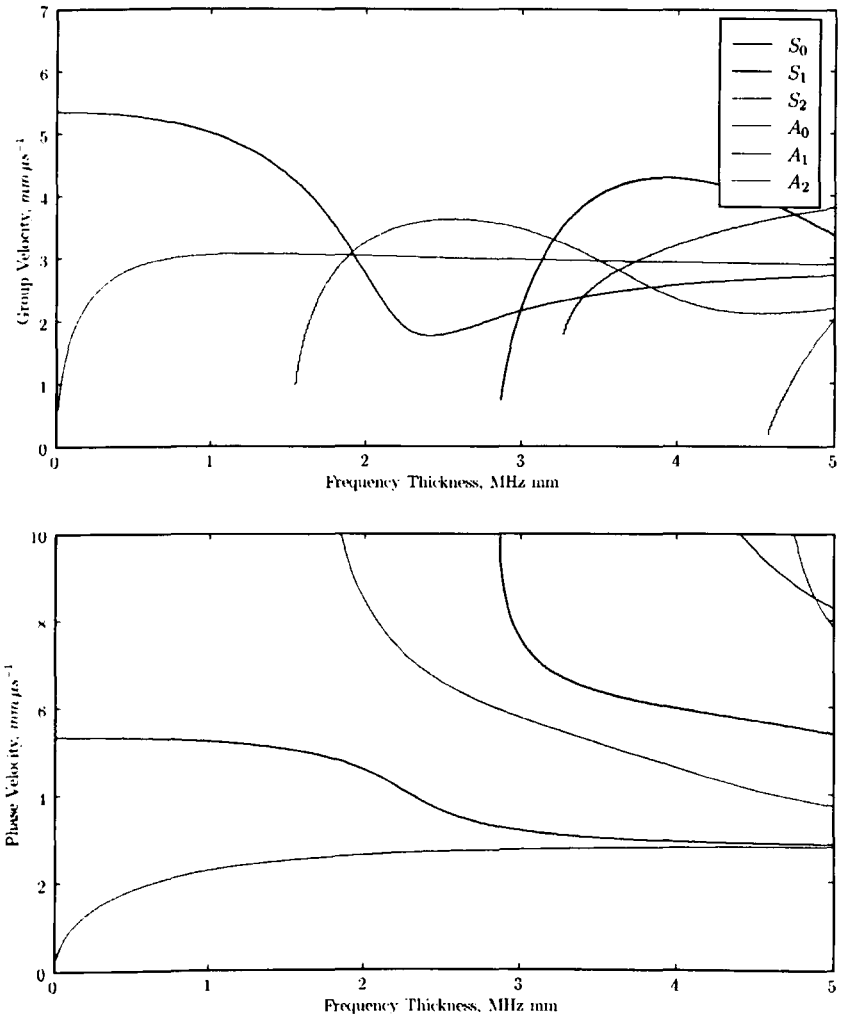


Figure 2.6: Dispersion Curves for an Aluminium Plate

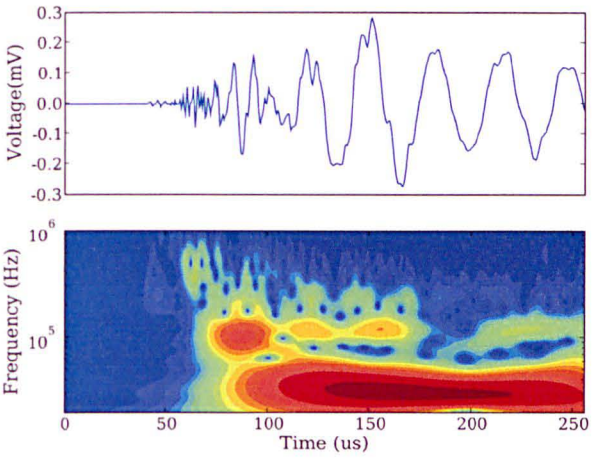


Figure 2.7: Pencil lead fracture on the surface of an aluminium plate at 150mm

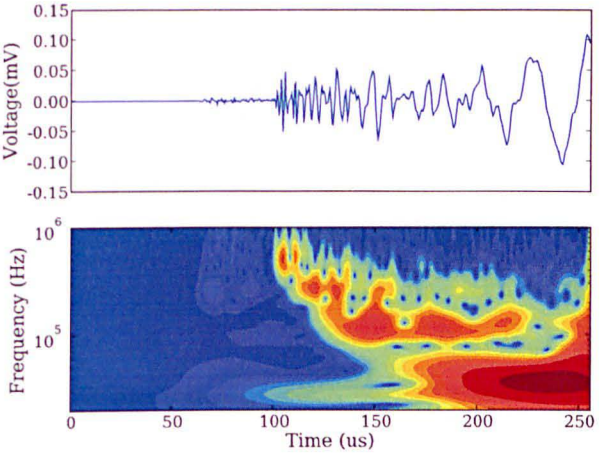


Figure 2.8: Pencil lead fracture on the surface of an aluminium plate at 300mm

Figure 2.7 and 2.8 were generated by activating the Hsu-Nielsen source on the upper surface of the plate (out-of-plane) at 150mm and 300mm from the transducer position respectively, whilst Figures 2.9 and 2.10 involved the activation of the Hsu-Nielsen source on the *side* of the plate (in-plane) at the same distances.

The depth of the aluminium plate was 1.2mm, and the bandwidth of the transducer

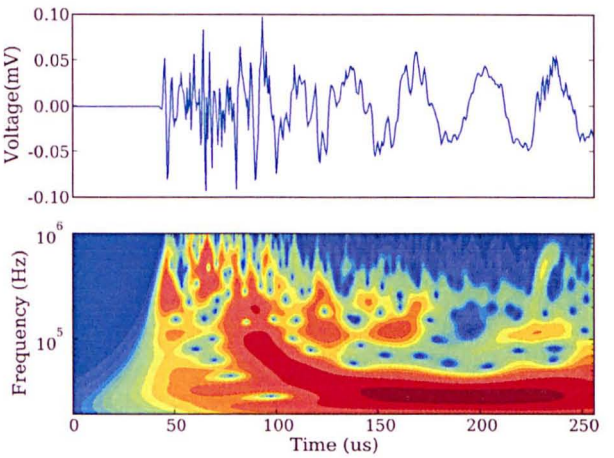


Figure 2.9: Pencil lead fracture on the *side* of an aluminium plate at 150mm

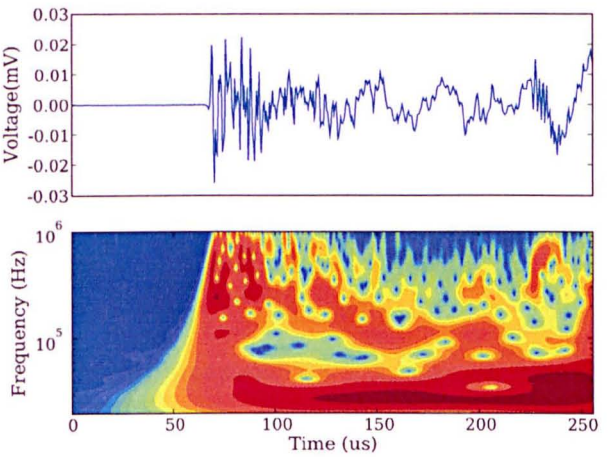


Figure 2.10: Pencil lead fracture on the *side* of an aluminium plate at 300mm

used (A Physical Acoustics UT1000) was 100kHz – 1MHz, giving a MHzmm bandwidth of 0.12 - 1.2 MHzmm. Inspecting the dispersion curves, it is clear that only two modes can propagate in this region, the A_0 and S_0. Also, the speed of the S_0 mode is much higher than that of the A_0 mode, with group velocities of $\approx 5mm\mu s^{-1}$ and $\approx 3mm\mu s^{-1}$ respectively. It is simple to identify the arrival times of each component of the waves for each figure: the S_0 mode should arrive at $30\mu s$ and $60\mu s$

for the 150mm and 300mm propagations, whilst the A_0 mode should arrive at $50\mu s$ and $100\mu s$. These timings are particularly clear in figure 2.8, a low-amplitude wave arrives (slightly later than expected) at around $60\mu s$ (this is the S_0 mode) and a large amplitude wave arrives at $100\mu s$ (this is the A_0 mode).

The difference between the figures generated by in-plane and out-of-plane sources are large. The out-of-plane source generates a small S_0 mode, which can be distinguished before the arrival of the larger A_0 mode. The roles are reversed for the in-plane source, where the S_0 mode is large and obscures the arrival of the A_0 mode.

There are also differences within the same source type: the two waves recorded at different distances differ in more than a simple time-lag. Consider figures 2.7 and 2.8: the gap between the S_0 arival and the A_0 arrival increases with distance from the source: this is to be expected because of the speed of the two wave packets. However the wave packets themselves appear different because of the effect of dispersion. These effects also give rise to marked differences between figures 2.9 and 2.10, though the two-mode structure is not as clear.

Whilst this demonstration of the effects of dispersion on AE waveforms leads to some understanding of waves generated in plates, it does not account for the complex structures which are the target of this project. Examining such a structure (illustrated in Figure 1.1), it is clear that the main part of the structure is tubular. This would not pose a huge problem (dispersion curves can be calculated for simple tubes [21]) were it not for the thickness changes, holes, lugs and other geometrical intricacies which are apparent.

Having seen that the wave propagation section of the AE forward model is extremely complex, the next section discusses approaches that one might take in distinguishing between sources of AE.

2.5 Approaching the Source Characterisation Problem

An ideal way to approach the problem of identifying fracture-related AE sources would be to find to find a unique property of the source (such as the frequency content, or time signature) which distinguished fractures from other sources. This

is an unfeasible approach since the information about the source is obscured by the nature of the AE forward model. Further, it is difficult to say with any confidence that there *are* any unique features of the source which are useful in this manner, since positively identifying the source condition is extremely difficult in all but the simplest of conditions.

Spencer et al [20] proposed a method for source characterisation whereby the wave propagation was modelled by finite difference simulation (known as a Local Interaction Simulation Approach, LISA), built on the work of Lee [40]. The method involves solving the difficult inverse-problem using an evolutionary approach, and has been shown to successfully identify sources in plate structures using a laser vibrometer, which is a small-point, contact-less measurement device.

A second approach might be to use a neural network classifier (or similar machine learning tool) to distinguish between fractures and other sounds. One would need to collect a database of AE waveforms relating to different types of source (i.e. fracture, fretting, rubbing) and train a classifier to distinguish between them. In this approach the wave propagation section of the forward model must be accounted for by the neural network. This approach is severely limited since a trained classifier would only be useful for the particular geometries used in the training set. To train a classifier to generalise across various geometrical conditions would require collecting data for every conceivable permutation of source condition and wave propagation path – clearly a rambunctious proposition.

The approach taken in this work is to create a *data driven* approach to identifying fractures in AE data, which takes into account our understanding of the physical phenomena. Rather than attempt to classify each individual waveform, a method is proposed whereby an analysis of patterns, trends and clusters in sequential waveforms leads to the identification of crack like behaviour. The method utilises the understanding of the AE model as described in this chapter, particularly with reference to the amplification of source uncertainties by the wave propagation, and the time dependent nature of different AE source mechanisms. The next chapter outlines this proposed method.

2.6 Chapter Summary

For AE tests of structures on the scale of landing gear, it is clear that modelling of the AE process is not a viable option as there are too many unknown factors – source nature, propagation type, location uncertainties, mode conversions etc – which make the modelling difficult. This is a major motivation for the methods outlined in the remainder of this thesis.

Furthermore, identifying fractures on a signal-by-signal basis is also an unappealing direction - the amount of uncertainty associated with each signal is high due to dispersion, unknown source origins etc. the next chapter discusses possible ways in which to identify fractures based on a sequence of AE signals.

Chapter 3

A Data-Driven Method for AE

In the previous chapter, the physical aspects of AE were discussed, and the argument was put forward that in order to classify acoustic emission signals, one would need some extensive (physical or neural) model of the AE process, which is impracticable. In this chapter a data-driven method for detecting fractures in large (and small) scale AE tests is presented. We first present an outline of the method, and in subsequent sections present details of how each component in the algorithm works, together with some background.

3.1 Method Outline

The first stage of the proposed methodology involves organising the AE data into clusters as shown in Figure 3.1. Additionally, feature extraction is performed on the data so that it can be easily stored and compared at a later stage.

The basic element of acoustic emission data is known as a *hit*. When the voltage at any sensor is above some user-determined threshold, the AE system records the voltage at that sensor for a fixed amount of time. A memory buffer is used to record data prior to the threshold crossing. The resulting recorded signal is oft referred to as a *waveform*. The time (T in figure 3.1) of the threshold crossing is recorded, as is the channel number (Ch) to which the sensor is attached.

Timing parameters are defined in order to group *hits* into *events*. An event is a

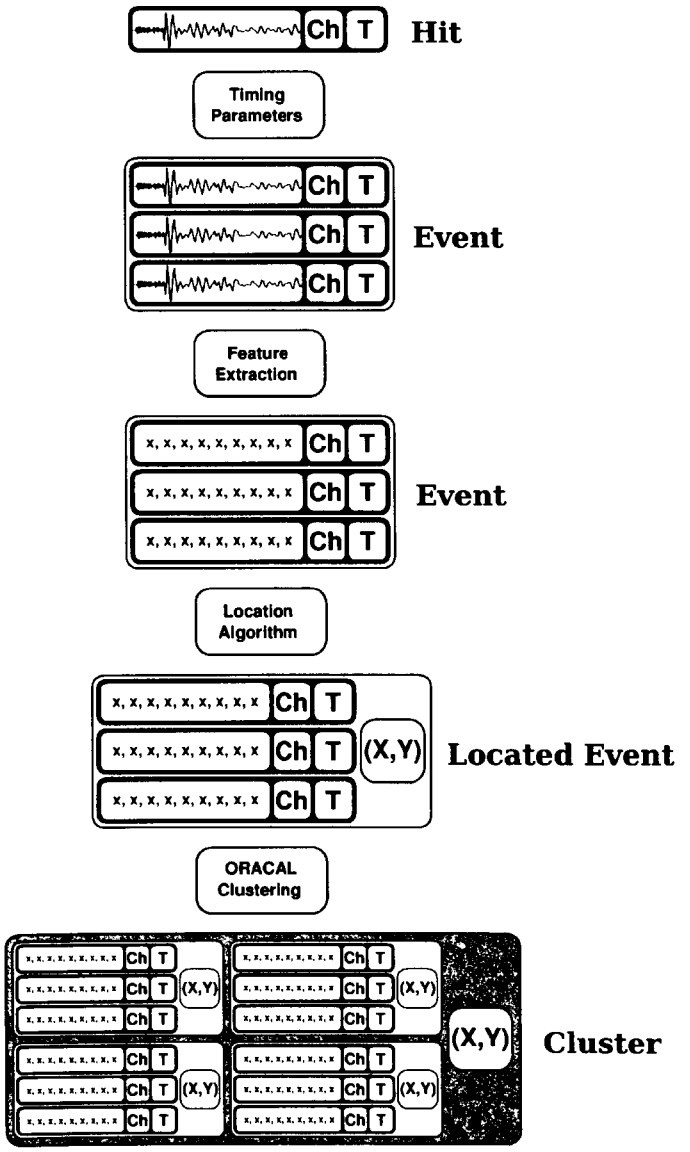


Figure 3.1: Flow diagram showing the organisation of AE data, including the feature extraction, location and clustering stages

group of hits recorded on different channels, which were detected sufficiently close together that they can be assumed to originate from the same source. Occasionally hits are recorded which do not fall into events: these must be discarded.

Feature extraction is explained in Chapter 4. Here, it is sufficient to say that the waveforms in the event are reduced to a low-dimensional set of numbers in such a

way that the numbers describe the 'nature' of the waveform. This is a necessary and important step not only because of data storage requirements; it also allows comparisons to be easily made beteen waveforms which is crucial in later stages of the methodology including visualisation and analysis of the feature space distribution.

The next stage is to assign a location to the event. The location algorithm is explained fully in Chapter 6. This is an important process because 1) it is necessary to group the events together according to their spatial location, and 2) it is necessary to correlate the AE data with the real world (i.e. so that the engineer knows which component to change, or where to perform further analysis)

Once the data is organised into events, the events need to be grouped together into clusters. Clusters are the lowest level in our hierarchy of data: it is the clusters that will be assessed in order to determine whether or not a fracture is present. The remaining stages in the methodology involve assigning labels to clusters based on rate of energy, feature space distribution and novelty, combining these labels or 'clues' in some way (data fusion) and visualising the data. Visualisation is the topic of Chapter 5, and we continue this chapter with a discussion of the clustering algorithm.

3.2 Spatial Clustering

Clustering is the division of a data set into a series of subsets, based on some notion of similarity of the data. In our case, the similarity is going to be a distance, measured across the surface of the structure being monitored, in meters. Since each event has been assigned a location in the previous stage, it becomes possible to group together events based on their locations. There are many different clustering algorithms, the reader may consult any number of appropriate texts for more information, [47] is a good example. There follows a brief review of some clustering algorithms, and an explanation and discussion of the ORACAL algorithm which has been developed for this application.

3.2.1 Grid Allocation

A simple method of partitioning data is currently employed by the commercial AE software. The space of interest is divided into a $n \times m$ grid. Data is allocated to the box in the grid into which it falls; boxes can then be ranked according to the number of data points in them. This simple method is somewhat inflexible (one must pre-define the grid) and does not allow for segments of the data which may span grid boundaries.

3.2.2 K-Means Clustering

K-Means clustering is simple clustering algorithm, involving the division of the data into k groups, according to the data's proximity to k centres; one must define k before running the algorithm. Despite its simplicity, k-means clustering is often a very useful tool. Bishop [48] describes the k-means algorithm and discusses its importance as a simplified form of the powerful Expectation Maximisation (EM) algorithm.

Theory

The k-means algorithm proceeds as follows:

1. Define k centres, points of the same dimension as the data, allocated at random positions within the data space.

2. Associate each datum with the nearest centre.

3. Move each centre to the mean of the data associated with it.

4. Repeat steps 2 and 3 until convergence (i.e. the centres remain stationary in step 3 and no data change association in step 2)

A more detailed description of k-means clustering is given in [47] and [49].

Discussion

Selection of the number of centres k is somewhat arbitrary, unless the engineer has some prior knowledge of the number of expected groups in the data. Some methods do exist for selecting the number of clusters (e.g. running the algorithm for increasing k, and selecting the solution where a further increase in k introduces little improvement in the average distance to centre). Such methods are computationally expensive and require the use of another user-defined setting. These problems make k-means clustering (and related algorithms) unsuitable for our application: the number of clusters is unknown, and will increase as the amount of data collected increases.

3.2.3 Radius Clustering Algorithms (RACAL)

There are several radius-based clustering algorithms available, such as Self Organising Oscillator Networks (SOON) [50] and Quality Threshold (QT) clustering [51].

RACAL algorithms are suited to the task of clustering spatial AE data because they do not require the specification of k, the number of clusters: the number of clusters arises naturally from the data. RACAL algorithms *do* require the specification of the cluster diameter, D. This is a natural quantity to encode in the algorithm in our case: the distance relates to the expected size of the cluster upon the structure. One can define D with knowledge of the structure, the sensor spacing, and understanding of the location algorithm.

The next section describes the Quality Threshold (QT) clustering algorithm, originally developed by Heyer et al. [51] in order to cluster genome data.

QT Clustering Theory

The QT Clustering algorithm consists of two stages: the candidate cluster stage and the candidate selection stage.

Candidate Cluster Stage

1. Define a maximum cluster diameter, based on knowledge of the problem at hand.

2. Create a candidate cluster: starting with the first data point, add its nearest neighbour. Continue adding the nearest datum (i.e. the one which increases the diameter of the cluster the smallest amount) until the pre-defined diameter limit is reached.

3. Repeat step 2, starting with the second, third, etc datum until there are N candidate clusters. Note that at each point the data from the previous cluster are *not* removed, so that the N candidate clusters have significant overlap.

Candidate Cluster Selection The candidate cluster with the most data is selected as a cluster, and the first stage is run again with the remaining data. The process iterates until all the data is divided into clusters.

Discussion

The nature of the QT clustering algorithm is such that massive numbers of candidate clusters are created repeatedly. This is clearly wasteful of computational resources and does not suit an online application. In the next section, the online RACAL (ORACAL) algorithm is proposed, which utilises the distance parameter in a way similar to the QT algorithm, but does not involve the re-calculation of many candidate clusters and is suited to online applications.

3.2.4 Online Radius Algorithm (ORACAL)

The ORACAL method described here extends the RACAL algorithms above to our application by realising that the data is sequential. Further, the algorithm is designed to be computationally efficient and work 'on-line'. This is in contrast to the QT method of repeatedly generating candidate clusters described above, which can be very computationally demanding.

Theory

1. Define a maximum cluster diameter, based on some knowledge of the problem.

2. At the arrival of the first datum, allocate one cluster centred on that datum.

3. For subsequent data:

   - If a datum falls within the diameter of the nearest cluster, it is added to that cluster. The cluster centre is re-calculated as the mean of all the data within it.
   - If a datum is not near enough to any of the existing clusters, a new cluster is created, centred on the datum.

4. Clusters can *merge*: if the re-calculation of a cluster's centre leads to that centre being within the specified distance of another cluster centre, then the two clusters become one larger cluster.

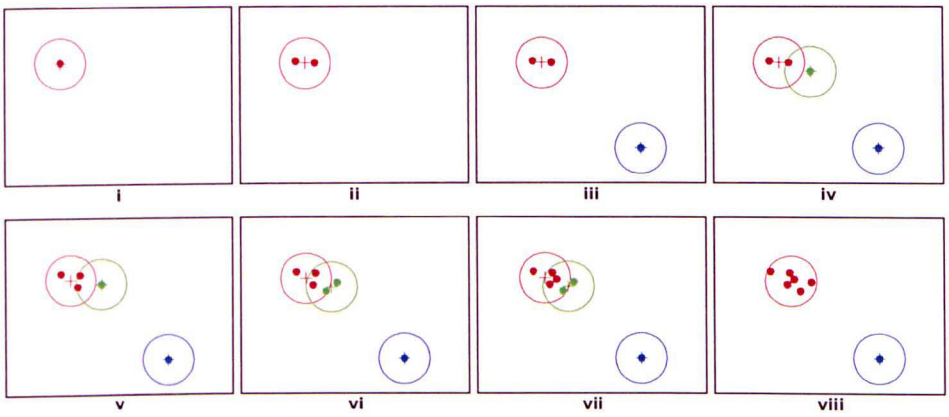


Figure 3.2: Illustration of ORACAL, showing the addition of seven data (small dots), resulting in three clusters (centres represented by small crosses, loci represented by solid lines)

Figure 3.2 show the ORACAL algorithm at work on a small data set. The different stages can be described as follows:

i. The first datum always creates a new cluster (red)

ii. Further data which fall within this cluster are added to the cluster; the centre of the cluster is updated to reflect this.

iii. Further data which do not fall within a cluster create a new cluster (blue).

iv. A further datum falls just outside the red cluster, creating a new (green) cluster. The loci of the clusters are allowed to overlap.

v. A further datum falls within the region shared by the red and green clusters, and is associated with the nearest (red) cluster.

vi. A further datum falls within the region shared by the red and green clusters, and is associated with the nearest (green) cluster.

vii. Another data point is associated with the red cluster, and now the update to the centre of the cluster means that the *centre* of the green cluster is contained by the red cluster, therefore:

viii. The red and green clusters merge to become one (red) cluster.

Discussion

The ORACAL algorithm retains the benefits of the RACAL methods described above, but has the added advantage that the cluster structure need not be re-calculated with the addition of every data point. Outlying data will not affect the cluster structure, since the arrival of an outlier generates a new distinct cluster; this was not the case for the k-means algorithm, where the effect of an outlier was to move the centre of the nearest cluster toward that outlier.

The ORACAL algorithm does rely on the seemingly arbitrary setting of a cluster radius. However, in experience with real data (see Chapter 7) this did not cause any problems; for these tests, the cluster radius was set to 15mm. The cluster radius must clearly be able to account for any measurement error in the location algorithm, and is thus dependent on the method used: using a simple trigonometric system, the errors in location are amplified depending on the sensor arrangement (see Chapter 6), whereas using some other methods as proposed later should minimise the location error.

Once the data has been clustered with the ORACAL algorithm, the clusters are examined for 'clues' which identify them as potentially representing fractures. The next three sections describe some of these clues.

3.3 Feature Space Distribution

When examining the fatigue fracture as a source of Acoustic Emission in section 2.2.1, some *uncertainties* were identified, including crack length, depth, opening mode and whether the AE was generated by crack face rubbing, extension etc. In section 2.2.2, uncertainties associated with frictional sources were identified. The first 'clue' by which the clusters are to be ranked depends on the following hypothesis:

> *The nature of the uncertainties in the extracted features for fracture sources and frictional (unwanted) sources is such that the frictional sources produce a wider range of AE signals i.e. the feature-space variance associated with the fracture events is lower.*

If this hypothesis holds, then clusters which pertain to a fracture source (including fretting, crack face rubbing and crack extension) will contain signals which are very similar, whilst those pertaining to a frictional source will contain signals which are less similar.

The hypothesis is backed up by observational evidence. A visualisation of AE data in Pullin et al. [52] as well as several small scale fatigue test investigations showed that events which could be attributed to fracture activity (usually via spatial filtering) were much more tightly grouped together than those attributed to other AE mechanisms.

In order to quantify this behaviour, the variance of the feature data in the cluster (x^c) is compared with the variance of the whole data set (x) thus:

$$p(fracture|features) = 1 - \frac{\sum_i^{N^c} \frac{(x_i^c - \overline{x}^c)^2}{N^c}}{\sum_j^{N} \frac{(x_j - \overline{x})^2}{N}} \tag{3.1}$$

Since $x^c \in x$, the variance of the cluster x^c will always be less than the variance of the whole data set, and so $p(fracture|features)$ is always a value between 0 and 1. This is the motivation for identifying the measure as a conditional probability in equation 3.1; one should be aware that this identification is by no means rigorous. Using this equation blindly is fraught with danger; if the number of events in a cluster (N^c) is 1, then the variance of that cluster is 0, and $p(fracture|features) = 1$. Clearly there has to be a significant number of events in the cluster in order to get a sensible

result. Further, for sensible data sets, p can never reach 1, since both the variance of the whole set of data and that of the cluster must be finite.

A further caveat is that equation 3.1 must only be applied to signals from the *same channel*. To do otherwise would be to compare the uncertainty not only of the source, but also in the wave propagation paths, which is nonsensical.

Clearly the clue is sensitive to the variability of signals in other clusters during the test: in order to get a consistent result, there needs to be a large number of non-crack related signals in order to properly estimate the variance of all signals in the test. This is usually not a problem!

3.3.1 Examples

Two examples of observational evidence for the feature-space-distribution criterion are presented; the first is from the fatigue testing of a Compact Test (CT) specimen, and the second is from the fatigue testing of a small landing gear component, the lever link.

CT Specimen

Part of this project included the fatigue testing of several CT specimens, and subsequent analysis of the recorded AE data. The results of just one typical test are summarised here for brevity. Figure 3.3 shows a PCA projection (see section 5.2) of features from AE signals recorded during a CT test. Two sensors were used, one on either side of the crack, and events relating to crack activity were identified by time-of-flight differences and have been highlighted in red. The crack related signals appear much closer together than the other signals; since PCA retains as much variance as possible during the projection, this is a sensible way to asses the variance of sub-sets of the data. Analysing the data *before* projection according to equation 3.1 confirms this with a $p(fracture|features)$ value of 0.903, significantly higher than that of the next highest value , 0.638, associated with a cluster of points located near a loading pin.

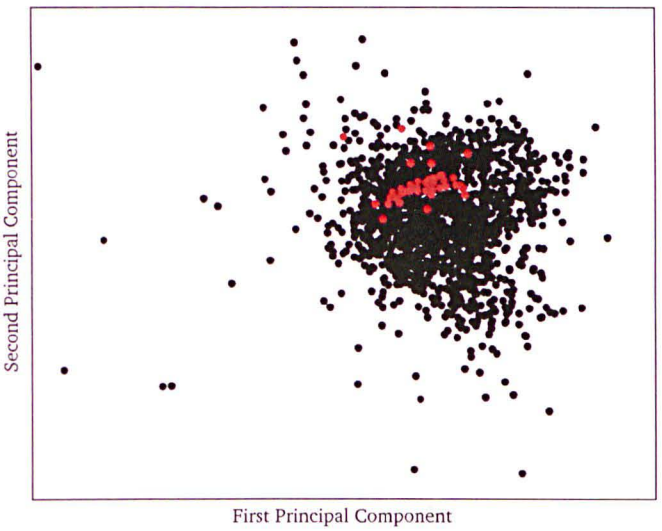


Figure 3.3: Principal Component plot of the data from the fatigue testing of a CT specimen. Data from the crack region (identified by time of flight) is coloured green. Note that the axis scales are irrelevant and have been omitted.

Lever Link

During a CT test as above, there are few sources of extraneous AE activity: only the rubbing of the loading pins contributes to activity hiding the crack activity. A small component from a landing gear, a *lever link* was tested in a similar fashion, as described in [52], where there were significantly more frictional sources due to the loading nature of the component. Figure 3.4 shows a PCA projection of recorded signals, with colour representing position on the component (found by time-of-flight location). Crack related signals were identified by position (post-test analysis revealed several initiated cracks) and were all found to appear in the small group (dark blue) in the upper left corner of the figure. Clearly, the variance of the plot is dominated by non-crack activity. More details can be found in [52].

It is clear from the plot that the fracture related signals appear to be far more compact than those from other sources. This is reflected by the $p(fracture|features)$ values for the areas where the fracture was identified, being 0.862. There was one cluster in the test not identified as a fracture which also had a high value (0.849), a potential false positive. All other clusters (of significant size, i.e. over 50 events) had a value of no more than 0.578. This demonstrates that the proposed feature

based clue is potentially powerful in detecting fracture related clusters, though it may need to be combined with further information.

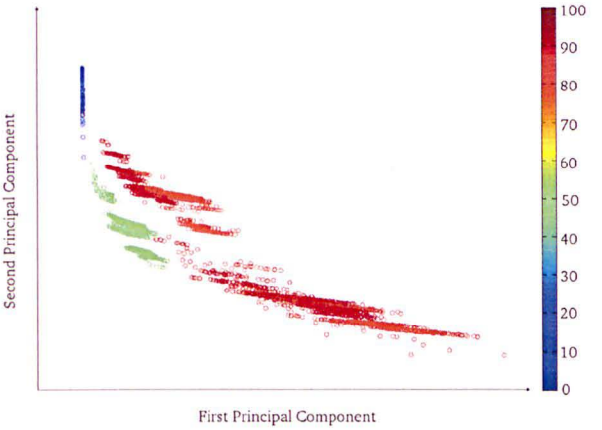


Figure 3.4: Principal Component plot of the data from the fatigue testing of a lever link component. The colour scale denotes positions along the component (in millimetres). Data from the crack regions (around 10mm) appears in the dark blue patch in the top left of the plot.

Discussion

It is clear that there is insufficient statistical evidence in this section to fully support the sole use of this clue for finding fractures. Nonetheless, it has proved to be a useful indicator in the presented tests as well as the more thorough test on landing gear structures in Chapter 7. A full statistical analysis of $p(fractures|features$ is proposed for further work.

Furthermore, this 'clue' is clearly dependent on the particular type of features extracted from the signals. In the cases presented here, the traditional features were used (see Chapter 4). In the work in Chapter 7, a quick study revealed that the particular type of feature used made little difference to the results. For the tasks in Chapter 4 (where the feature types were compared on a classification type task) there seemed to be no significant difference between the feature types.

3.4 Energy Trends

The next 'clue' in determining whether or not a cluster represents a fracture is determined by the energy rate of that fracture. Recall section 2.2.1, where it was determined that the rate of acoustic emissions is proportional to crack growth. This is reflected in the literature, where rates of energy recorded by an AE system have been related to crack growth [53], and can be used to predict a component's remaining fatigue life [54].

The energy rate of a cluster therefore gives some indication of the type of AE source associated with that cluster. One would expect the energy from a non-fracture related mechanism, present throughout the test, under unchanging conditions of loading, to accumulate at a constant rate. The evidence for a crack source suggests otherwise; the count rate for crack propagation is known to increase with crack length. This means that any significant deviation of linear accumulation of energy for a given spatial cluster constitutes another indicator of the presence of a crack.

3.4.1 Quantification

In order to quantify the deviation from a linear energy accumulation, it is possible to fit a simple model to the time-energy data relating to each cluster. The model parameters relating to each cluster can then be objectively compared. Let $\mathbf{t}$ be the time of each event in the cluster, and $\mathbf{E}$ be the energy of those events, cumulatively. We propose a model of the form:

$$\mathbf{E} = \mathbf{a} + \mathbf{bt} + \mathbf{ct^2} \tag{3.2}$$

Figure 3.5 shows energy data from three spatial clusters, with approximations to the data as per the model in equation 3.2. In figure 3.5(i) the coefficient c is near zero. In 3.5(ii) c is positive and in 3.5(iii), c is negative. According to our hypothesis, 3.5(ii) is most likely to indicate a growing fatigue crack.

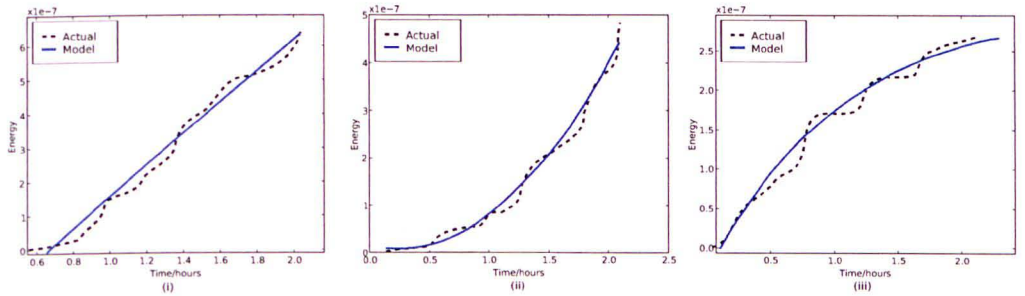


Figure 3.5: Plots of the energy from three different clusters of AE signals, showing zero (i) rising(ii) and falling trends (iii). Parameterised models of the data are shown.

Least-squares model fitting

There follows a derivation of the least squares formula, in which the data is *pre-processed* according to the model in 3.2.

In order to fit the model in equation 3.2 to the data, it is written in the form

$$\mathbf{E} = \mathbf{X}\mathbf{w} + \varepsilon \tag{3.3}$$

where $\mathbf{w} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ is a vector of the parameters or *weights* of the model, and $X = \begin{bmatrix} \mathbf{1} & \mathbf{t} & \mathbf{t^2} \end{bmatrix}$ is a matrix containing a column of ones, the time data, and the time data squared. An error term ε is introduced. The values of a, b, and c are found by minimising the mean squared error 3.4. The derivation of this simple linear model is well known [55], but is included here for completeness.

The sum of the mean-squared errors is:

$$J = \sum_i \varepsilon_i^2 = \varepsilon^{\mathbf{T}}\varepsilon \tag{3.4}$$

Substituting 3.3 into 3.4, we obtain:

$$J = (\mathbf{E} - \mathbf{X}\mathbf{w})^{\mathbf{T}}(\mathbf{E} - \mathbf{X}\mathbf{w}) = \mathbf{E}^{\mathbf{T}}\mathbf{E} - \mathbf{w}^{\mathbf{T}}\mathbf{X}^{\mathbf{T}}\mathbf{E} - \mathbf{E}\mathbf{X}\mathbf{w} + \mathbf{w}^{\mathbf{T}}\mathbf{X}^{\mathbf{T}}\mathbf{X}\mathbf{w} \tag{3.5}$$

We desire the value of $\mathbf{w}$ such that J is minimum, which we shall denote $\mathbf{w}_*$:

$$\frac{\partial J}{\partial \mathbf{w}^{\mathbf{T}}} = 0 = -X^T\mathbf{E} + \mathbf{X}^{\mathbf{T}}\mathbf{X}\mathbf{w}_* \tag{3.6}$$

rearranging:

$$\mathbf{w}_* = (\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathbf{T}}\mathbf{E} \tag{3.7}$$

By solving equation 3.7, one obtains a value of c for each clusters. The values of c can be transformed to a probabilistic interpretation by outlier analysis: since *most* of the cluster present in a test will not represent a fracture, one can perform an outlier analysis [56], which is simple since c is a univariate variable. The *discordancy* of c is given by:

$$\zeta_{cluster} = \frac{c_{cluster} - \bar{c}}{\sigma} \tag{3.8}$$

where $c_{cluster}$ is the value of c for the cluster of interest, $\bar{c}$ is the mean value of c for all clusters, and σ is the standard deviation of the same.

If one is to make the assumption that the distribution of c across clusters is Gaussian, then the discordancy of a cluster can be easily transformed into the probability of a fracture by:

$$p(fracture|c) = 1 - \frac{1}{\sqrt{2\pi}\sigma}\exp(-0.5\zeta^2_{cluster}) \tag{3.9}$$

This is convenient when performing data fusion, see section 3.6.

3.5 Novelty

The nature of the landing gear certification test toward which this project is orientated means that novelty detection is likely to be an excellent method for detecting damage. The tests are of a long duration (typically six months) during which it is possible to collect a plethora of AE data. Novelty detection involves collecting this data, and deciding upon some safe limit based on the data. Further observations which fall outside this limit are classed as novel, and can be acted upon accordingly.

The novelty detection approach was first used by Tarassenko et al. [57], to detect masses in mammograms. Subsequent work with novelty detection has seen it become a useful tool in fields such as vibration of structures [58], classification of image

regions, [59] and biomedical data-processing [60].

Novelty detection requires the collection of a set of data which is labelled as *safe* or *normal*. This data is then used to train some (statistical or neural) model which describes the nature of the system. Subsequent data is then compared to this model, and determined either (a) to fall within the bounds of the model, i.e. the data does not deviate from the safe data and is considered not representative of damage; or (b) to fall outside the bounds of the model, and represent a damaged state.

The uncertain nature of acoustic emission makes it difficult to build a model based upon signal subspaces (i.e. feature extraction, see Chapter 4). Instead, we present here a method for applying novelty detection to the *spatial* distribution of signals. This approach is obviously highly dependant on the reliability of the location algorithm (see Chapter 6).

The method proceeds as follows:

1. The data designated as 'normal' is divided into overlapping time windows.

2. The spatial distribution of the data is modelled using a Gaussian Mixture Model (GMM, see Bishop [55]).

3. Subsequent (test) data is also divided into windows (of the same length as that for the training data) and modelled using a GMM.

4. The *parameters* of the mixture models are compared using a statistical analysis such as outlier analysis [56].

5. The mixture models can give a degree of novelty for any point, easily assigning a novelty index to each cluster.

As crack events are likely to emerge as novel features in the spatial distribution of recorded events, novelty detection offers the possibility of another clue or indicator to the emergence of a fracture. The mean position of a cluster can be related the the current spatial density model and a degree of novelty obtained.

Novelty detection in general is proposed as a seam of future work, see section 8.2.1.

3.5.1 Other Possible Indicators

The framework is capable of including more indicators of fracture activity as understanding of the AE process brings more indicators to light. For example, one possibility might be *load cycle correlation* - do the signals in this cluster all come at the same point in the load cycle? or are they scattered across the cycle?

3.6 Decision Making

After carefully sectioning the data into clusters, and finding clues with which to describe the clusters, it is necessary to decide which (if any) of the clusters represents a fracture. A probabilistic framework is ideal for this, as it allows us to make statements about each of the clues such as:

> "Based on the novelty clue, there is a 95% chance that the cluster at position (x, y) relates to a fracture"

Further, the use of probability gives a good mechanism for combining clues via *data fusion*. This section describes the application of data fusion to the problem of identifying clusters related to crack growth.

3.6.1 Data Fusion

Hall and Llinas [61] describe two separate application of data fusion, one for determining *locations* and one for performing *identification*. The problem at hand is one of identification (we wish to specify whether or not a cluster relates to crack growth), and the data which we wish to fuse consists of *clues* (c.f. sensors in the data fusion literature). Further, Hall describes three sub-types of identity fusion, as shown in Figure 3.6

(i) **Data level fusion.** The raw data regarding each clue are combined, and feature extraction is performed on the fused data. Some transformation from the feature data is made to an identity class, perhaps by a mechanism such as an artificial neural network.

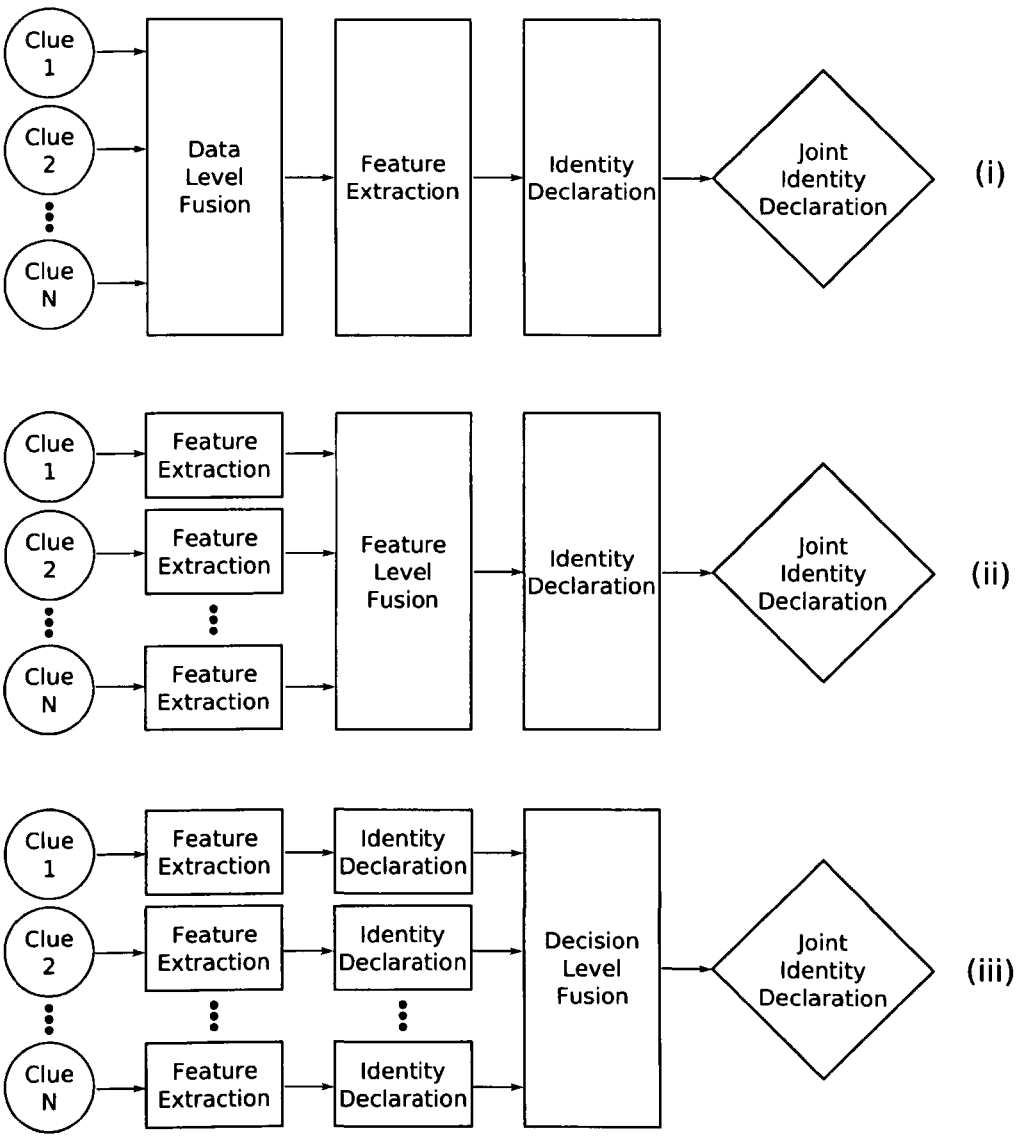


Figure 3.6: Three types of *identity* data fusion, showing data level fusion (i), Feature level fusion (ii) and Decision level fusion (iii)

(ii) **Feature level fusion.** Feature extraction is performed independently upon each clue, and the feature vector from each is concatenated together before an identification is made.

(iii) **Decision level fusion.** Feature extraction and identification are performed separately for each clue, and a combined decision is made by use of weighted decision methods, Bayesian inference or the Dempster-Schafer method.

One of the challenges regarding data fusion is *correlation*, or association: ensuring that data being fused corresponds to a single target. This is a problem specifically regarding identification of aeroplanes (via fusion of e.g. radar and infra-red information) and applies equally to many data-fusion applications. For our application, correlation is trivial since each clue refers to a unique cluster, and has been omitted from figure 3.6.

The natural fusion method for the application at hand is decision level fusion. Each clue involves essentially a different feature extraction[1] (see sections 3.3, 3.4 and 3.5), and so raw data fusion is impossible. Further, each feature set (for example equation 3.1) provides a way to produce independent identifications, and so decision level fusion is a natural choice.

The identity problem in this instance of data-fusion is a binary classification problem: a cluster either represents a crack or it does not. Each clue which has been described in this chapter presents a method for determining whether or not a particular cluster represents a fracture: the feature-space clue provides a number between 0 and 1 which we treat as a probability; under a Gaussian assumption the energy-trend clue provides a direct probability; the novelty clue can provide a probability in a Bayesian setting.

The simplest method of combining the clues is via a simple majority vote: if more than half of the clues have a probability of fracture of more than a given threshold for a particular cluster, then the combined identity of the cluster is a fracture. More sophisticated methods of data fusion include Bayesian data fusion (see e.g. Friedrich et al. [62]) and Dempster-Shafer theory data fusion (see e.g. Worden and Staszewski [63]). Since all of the clues are probabilistic, it is natural to use a Bayesian setting for data fusion.

[1] note that the features here correspond to the description of the clues as presented, not to the feature extraction method in Chapter 4

The actual implementation of data fusion is included in plans for future work see section 8.2.2.

Chapter 4

Feature Extraction

Feature extraction is the process of representing the object of interest (i.e. our acoustic emission waveform) by some short list of numbers (a vector). In other applications, feature extraction might involve extracting numbers from images (e.g. in face recognition), web pages (in search engine algorithms such as Google) or sounds (as in voice recognition).

In an acoustic emission context, the process of feature extraction has two important roles: 1) To *compress* the information due to storage requirements, and perhaps more importantly 2) as a basis for *comparing* signals for visualisation, classification etcetera. It is the features of the waveforms that are compared to form one of the indicators for the data-driven method in section 3.3.

Despite the current availability of large digital storage devices, storage of the complete waveform is still impracticable. A signal consisting of 2^{16} samples, at a sample rate of 5MHz represents an acoustic signal of just over 3ms, and (assuming a fidelity of 16 bits) requires 131 kilobytes. Since large AE tests may have signals arriving on over 120 channels, at a rate of several waves per second, over a period of several months, the data storage requirements for the complete waveform would be colossal. For example, a test during this work using only six sensors and lasting only ten hours produced 4GB of data - scaling this up to 120 channels and a 3 year test gives over 210 Terrabytes of data – this may be feasible for huge operation such as Google, but is massively inconvenient for laboratory use, especially since there is (as will be demonstrated) huge redundancy in the data. Some kind of compression

is therefore required if one is to keep a permanent record of an AE test.

The ability to compare signals is important for any algorithm which attempts to discriminate between them. Treating each signal as a long vector and finding the Euclidean distance in this very high-dimensional space is unlikely to work well in practise because of phase differences between the signals; it is also very computationally expensive. Further, reducing the number of dimensions of the data to two or three allows for visualisation of the data.

4.1 The Traditional Way

The technique of acoustic emission has been around for some time, and pre-dates the modern computer with gigabytes of memory. In order to store information from an AE event therefore, researchers and practitioners made use of a feature extraction technique as outlined in Figure 4.1. The motivation was reduced memory footprint - instead of storing several thousand floating point numbers in order to describe the samples of a waveform, only a few numbers were needed to be stored. Most notable amongst these were amplitude, rise time, duration, count and energy. Further parameters were later added as the field of AE evolved and matured. These include:

- **Initialisation frequency** The frequency of the section of the waveform leading up to the peak
- **Counts to peak** Threshold crossing before the peak (clearly linked to initialisation frequency above)
- **Peak Frequency** The centre of mass of a Fourier transform of the waveform

A number of the features outlined above (particularly those known to be more useful, such as the rise-time) are dependent on the user-defined threshold. This is undesirable, since two identical waveforms recorded with different thresholds will give rise to different feature sets.

This threshold-dependent method of extracting features from waveforms is still in use by AE practitioners today. Two feature extraction algorithms are proposed based

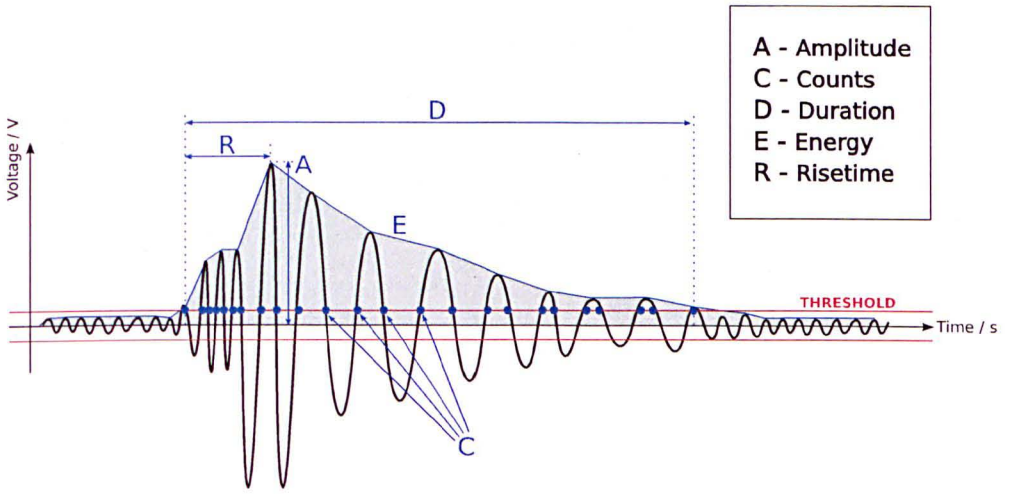


Figure 4.1: How Traditional AE features are calculated

on wavelet transforms, with the aim of extracting features which better describe the nature of the signal, and are threshold-independent.

4.2 Gaussian Mixture Modelling of the CWT

Consider a typical AE waveform as shown in Figure 4.2(i). The signal is composed of several bursts of energy; some of these are related to Lamb wave modes, some are effects of the transducer, and some are reflections from features of the geometry of the specimen under examination. An ideal feature extraction algorithm will encode this information in some manner.

Consider 4.2(ii). This is a spectrum of the waveform shown above [1]. The power spectrum shows the frequency content of the signal, but contains no information about timing. Conversely, examination of the waveform as in Figure 4.2(i) will give timing information, but no information about frequency. In order to create a feature set which takes into account the multi-modal nature of AE waves, it is necessary to examine the time-frequency content.

The short-time Fourier transform (STFFT, [65]) is one way to examine the time-frequency content of a signal. This involves windowing the data (in time), and taking

[1] Spectrum generated by the Welch method, see Welch [64]

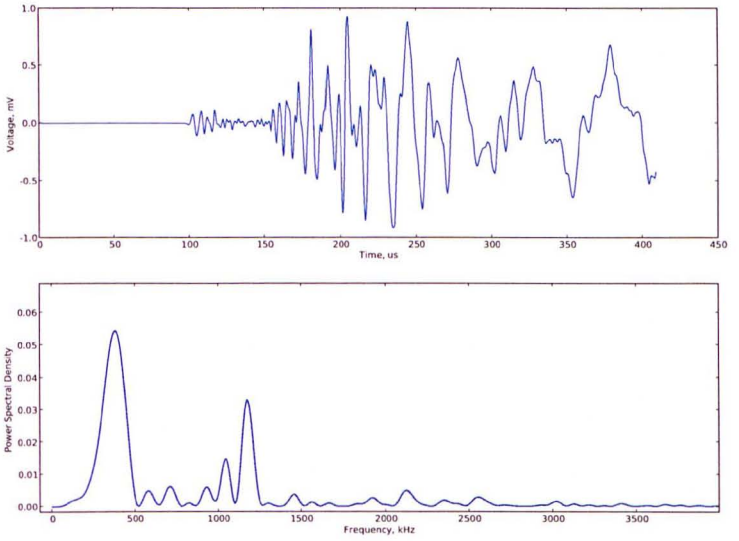


Figure 4.2: (i) Waveform of an artificial AE event, generated by a Hsu-Nielsen source on a 1.5mm aluminium plate. Received at a distance of 100mm by a Physical Acoustics UT1000 transducer. (ii) the Power spectrum of the same waveform

a Fourier transform of the windowed data. The window remains fixed, however, for all frequencies - for a full discussion of this see Addison [66]. The author prefers the *wavelet* transform, where the window and frequency are changed together, appropriately.

4.2.1 The Continuous Wavelet Transform

The wavelet transform has been used extensively in recent AE research [67]. It has been used to examine the source characteristics of the signal [68], and the location of signals based on modal structure [69]. Most of this work has been focused on the application of the Continuous Wavelet Transform (CWT) which provides a map of time-frequency[2] content of the signal, as demonstrated in Figure 4.4.

The Fourier transform breaks a signal down into a series of (complex) sinusoids,

[2]Technically, the wavelet transform provides a decomposition into time-*scale* space, but scale can easily be converted to frequency by the formula $f = f_0 f_s a^{-1}$, where f_0 is the 'centroid' frquency of the wavelet family, f_s is the sampling frequency, and a is the scale.

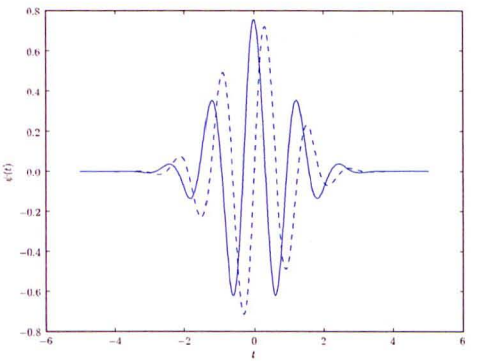


Figure 4.3: The Morlet wavelet function, showing the real (solid) and imaginary (dashed) parts

as shown in equation 4.1, where $x(t)$ is the signal, and $e^{-j\omega t}$ forms a *basis* for the signal. In the Continuous Wavelet Transform (equation 4.2) the basis for the signal is ψ^*:

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt \tag{4.1}$$

$$T(a,b) = \int_{-\infty}^{\infty} x(t)\psi^*_{a,b}(t)dt \tag{4.2}$$

where $\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)$, and $*$ denotes the complex conjugate. $\psi(t)$ is the wavelet function (often referred to as the mother wavelet in its untransformed state).

The wavelet function is usually oscillatory but short in time [66]. The choice of a wavelet function (of which there are many) is somewhat arbitrary, and throughout this section a Morlet wavelet is used, see Figure 4.3. This is convenient for our application: the Morlet wavelet has real and imaginary parts, and by use of the *magnitude* of the transform it is possible to disregard the phase information, leaving a smooth function of scale and time: this gives an elegant visual representation of the data, see Figure 4.4.

The equation for the Morlet wavelet (named after Jean Morlet [70]) is

$$\psi(t) = c_\sigma \pi^{-\frac{1}{4}} e^{-\frac{1}{2}t^2}(e^{i\sigma t} - e^{-\frac{1}{2}\sigma^2}) \tag{4.3}$$

which (ignoring the constant parts) is a complex sinusoid ($e^{i\sigma t}$) multiplied by a window ($e^{-\frac{1}{2}t^2}$). Typically, (and throughout this work) $\sigma = 5$. Clearly, changing the *scale* of the wavelet will affect both the sinusoidal and windowing section of the

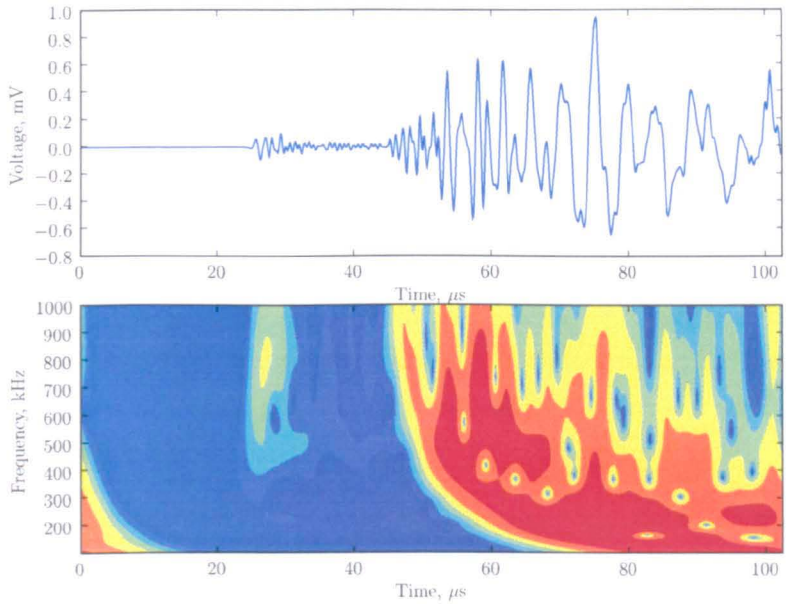


Figure 4.4: The waveform of Figure 4.2, and the corresponding wavelet map, taken at wavelet scales corresponding to frequencies between 100kHz and 1 MHz.

wavelet equally: the wavelet appears the same at all scales.

Figure 4.4 shows a wavelet transform of the waveform shown in Figure 4.2. The multiple modes, arriving at different times and frequencies are clearly visible. In performing feature extraction from the waveform, it is desirable to capture this modal structure. In the next section, a feature extraction method is proposed whereby the CWT data is modelled by a series of Gaussian functions, with the intention that each function will represent one important part of the data.

4.2.2 A Gaussian Mixture Model

In order to estimate the CWT surface using a mixture of Gaussians, a model is proposed of the form:

$$\psi'(a,b) = \sum_{n=0}^{N} A_n e^{([a\,b]-\mathbf{c}_n)\Sigma_n([a\,b]-\mathbf{c}_n)^T} \tag{4.4}$$

where A_n and $\mathbf{c}_n$ represent the magnitude and centre of the n^{th} basis function respectively. Σ_n is a 2×2 diagonal matrix, the elements of which represent the widths of the directions of a and b (i.e. across scale and time). The coefficients A_n, $\mathbf{c}_n$ and Σ_n are found by optimising the cost function

$$Z = \sum_a \sum_b \left(\psi(a,b) - \psi'(a,b)\right)^2 \tag{4.5}$$

Since Z is likely to have multiple local minima, some stochastic search algorithm is appropriate: the work here was performed by particle swarm optimisation [71], though many other search algorithms are appropriate such as differential evolution [72] or simulated annealing [73]. An explanation of the particle swarm algorithm is given in appendix A.2.

Once Z has been optimised, the values of A_n, $\mathbf{c}_n$ and Σ_n which occur at the optimal solution are used as a set of features.

Increasing the number of parts in the mixture

Increasing N, the number of basis functions in the mixture, has the advantage of increasing the accuracy of the representation of the CWT, with the trade-off of increasing the dimensionality of the feature data set. Figure 4.5 shows the CWT of a signal (top) and representations of that signal made using increasing N. For low N, the structure of the signal is well represented, but the finer detail is missing. As N increases, the finer detail is modelled. In order to apply this method for feature extraction, one must select N such that signals are well represented, without increasing the dimensionality of the feature space too far.

The described method of feature extraction is applied to some real AE data, and compared to the traditional features in section 4.5

Gaussian mixture modelling of the CWT is a method of feature extraction which gives a set of features containing information about the *time-frequency* nature of AE signals. This gives a potentially more powerful measure by which to compare signals, as was demonstrated in [75]. The method does have some drawbacks: it can take a long time for the optimisation routine to converge, making this a potentially slow technique; it also requires the engineer to select N, the number of parts in the mixture, which controls the complexity of the representation. The next section

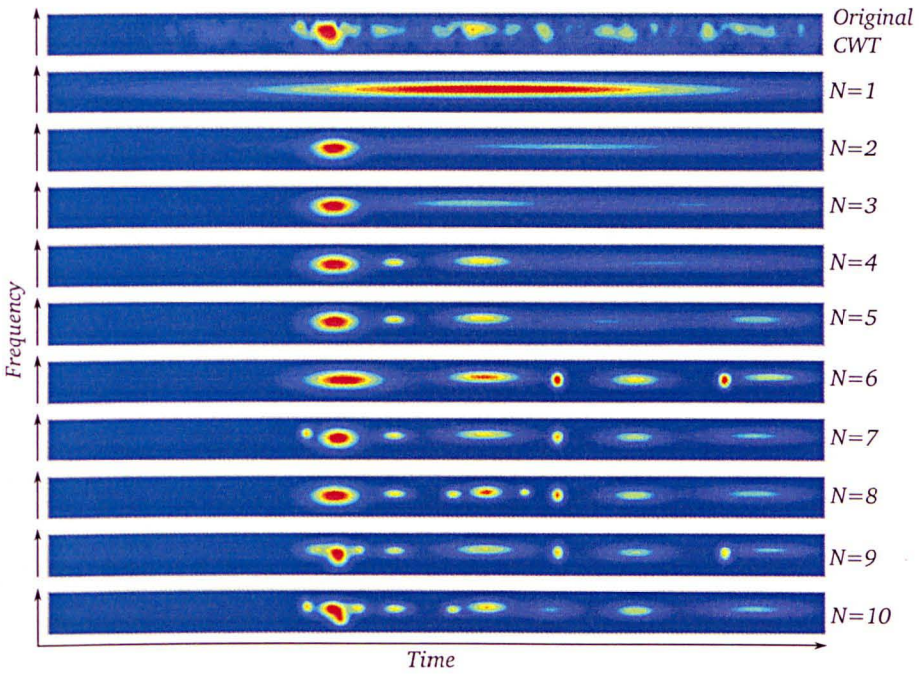


Figure 4.5: The CWT of a signal captured during fatigue fracture of a box girder specimen (discussed in [74]) and representations of that using a mixture of Gaussians. Each subsequent plot is a representation of the data using a model of increasing complexity as described in the text.

presents a feature extraction method based on the discrete wavelet transform which whilst still requiring the selection of the complexity of the representation, results in an extremely quick feature extraction method.

4.3 Discrete Wavelet Coefficients

The continuous wavelet transform results in an abundance of redundant data: in the above section, this data was 'compressed' by approximating the transform with a Gaussian mixture model. This section describes the Discrete Wavelet Transform (DWT), which enables examination of the time-frequency nature of signals without introducing any redundant data, and without any data loss (i.e. the transformed data contains exactly the same number of points as the original signal, and the signal can be *perfectly* reconstructed from the transformed data)

Dyadic Grids

The DWT is so called because the transform is only taken at discrete points in a and b, such that the translations in time are proportional to the scale of the wavelet:

$$\psi_{m,n}(t) = a_0^{-m/2}\psi(\frac{t - nb_0a_0^m}{a_0^m}) \tag{4.6}$$

Usually, a_0 and b_0 are taken to be 2 and 1 respectively, which means that the subscripts m and n define a series of points in the time-frequency space which form a *dyadic grid*: for successive discrete scales, the size of the translations doubles. A dyadic grid across scale (a) and translation (b) is illustrated in Figure 4.6 for logarithmic spacings of a.

When examining a signal of finite length, the scale of the first wavelet ($m = 1$) is chosen such that there are half as many translations as there are points in the signal. Thus the second scale is chosen such that there are half as many points again, etcetera. It is therefore helpful to have a signal of length 2^n, $n \in \mathbb{Z}$, else some technique such as zero padding is required.

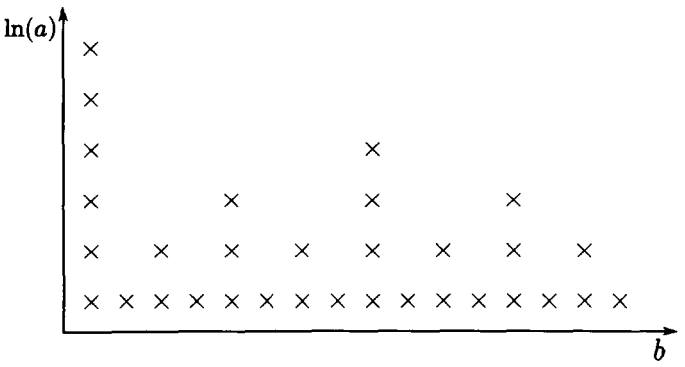


Figure 4.6: A dyadic grid across translation (b) and log scale (a)

Frames

If the original signal is to be recovered from the wavelet coefficients, it is natural that the total energy in each is matched. This is referred to as the formation of a *tight frame* (see Addison [66]), which means they must conform to

$$\sum_m \sum_n T_{m,n} = \int_{-\infty}^{\infty} x(t)^2 dt \tag{4.7}$$

Orthogonality

In order to achieve orthogonality, the wavelets must satisfy

$$\int_{-\infty}^{\infty} \psi_{m,n}(t)\psi_{m',n'}(t)dt = \begin{cases} 1 & \text{if } m = n \\ 0 & \text{if } m \neq n \end{cases} \tag{4.8}$$

A wavelet function $\psi_{m,n}(t)$ which fulfils (4.8) and (4.7) on a dyadic grid (4.6) is known as a *orthonormal dyadic discrete wavelet.* Many such wavelets exist, and the choice is a somewhat arbitrary one for the engineer: throughout this work, the Daubechies 10 wavelet is used.

Decomposition

The equation describing the DWT is similar to that for the CWT (equation 4.2):

$$T_{m,n} = \int_{-\infty}^{\infty} x(t)\psi_{m,n}(t)dt \tag{4.9}$$

however the complexity of orthonormal dyadic discrete wavelets means that the solution is rather involved, and the reader is directed else where (e.g [66]) for more information. The solution turns out to be a series of recursive filters, such that each successive filtering splits its input into two equally sized parts: *approximation* coefficients and the *detail coefficients*. The approximation coefficients are passed again through the filter to obtain the next level of representation, or *decomposition*.

Decomposition of an AE signal

Figure 4.7 shows the detail coefficients for a full decomposition of an AE signal. Subsequent plots show coefficients $T_{m,n}$ for increasing decomposition levels m (corresponding to *decreasing* frequency). The horizontal axis represents the translation parameter n. It is worth noting that each subsequent plot contains half as many data points as the one above it. The counterpart to Figure 4.7 is Figure 4.8, which shows *reconstructed* signals, using only the coefficients at each level. The number of points in each sub-plot here is always 4096, which is the length of the original signal. The term 'coefficients' is sometimes confused in the literature: some authors refer to the values in Figure 4.7 as the coefficients. Some authors make extensive use of the decomposition shown in Figure 4.8, see e.g. [76] for further information.

The signal was generated on a 1mm aluminium plate using a Hsu-Nielsen source at a distance of 200mm. The transducer used was a Physical Acoustics UT1000, which is sensitive between 100kHz and 1MHz; the sampling frequency was 40 MHz. The Daubechies 10 wavelet was used. Inspecting the dispersion curves for an aluminium plate (Figure 2.6), only two wave modes should be present below the frequency-thickness of 1MHz mm: the fast travelling S_0 mode, and the slower A_0 mode.

The first two levels of the decomposition, $m = 1 \ldots 2$ represent frequencies which are outside the bandwidth of the transducer. Nonetheless, some high-order Lamb wave activity is visible. The magnitude of this information is small, and has little impact on the construction of the signal.

The next two levels ($m = 3 \ldots 4$) show the components of the signal representing the S_0 mode, and the high frequency components of the A_0 mode (which looking at the dispersion curves, will travel the fastest and therefore arrive first).

Decomposition levels $5 \ldots 8$ show the rest of the A_0 mode, with lower frequency components arriving later in accordance with our understanding of the dispersion curves. The remaining levels ($m = 9 \ldots 11$) simply show the very low frequency components of the signal, which have little bearing on its construction.

Notice the edge artefacts caused by the wrap-around boundary condition (particularly observable in Figure 4.8), similar to the effect on the continuous transform (Figure 4.4).

Once the signal has been decomposed, one can use one of the levels of detail coefficients as a set of features to describe the waveform. Analogously to the problem faced in the previous section, it is necessary to select a level of decomposition, i.e. select the complexity of the representation. In the next section, some example AE data is collected and the effectiveness of each of the levels of decomposition is compared in describing the data.

4.4 Decomposition level comparison

In order to examine the usefulness of each level of decomposition as a feature set, three sets of experimental data were used. The sets of data were devised such that separability would be easy to achieve with the first, harder with the second, and hardest with the third. the separability of the data is compared for each level, and against the traditional features. This work was also presented in [77].

4.4.1 Experimental Procedure

Each experiment involved a series of Hsu-Nielsen pencil lead fractures on a composite plate. The Hsu-Nielsen source is often used in Acoustic Emission techniques in order to simulate AE source (see section 2.2.3). The composite plate was 3.5mm thick and consisted of 12 plies laid in a 0-90-0 configuration. The plate was of sufficient size that reflected waves (from plate edges) would not interfere with the recorded signal. A single AE sensor (Physical Acoustics Corp. UT1000) was used to monitor

the plate, attached to a Physical Acoustics PCI-2 system.

Eight positions on the plate were chosen, denoted A_1, A_2, A_3, A_4, B, C, D and E; the transducer position is denoted T. Points A_1 to A_4 are co-linear along the weave of the composite, with $\overline{TA_1} = 150mm$, $\overline{TA_2} = 200mm$, $\overline{TA_3} = 250mm$. Distance $\overline{TA_4} = \overline{TB} = \overline{TC} = \overline{TD} = \overline{TE} = 300mm$. TA_4 and TE are perpendicular, along the weave of the composite. Points A_4, B, C, D and E are evenly spaced, such that the angles $\angle A_4TB = \angle BTC = \angle CTD = \angle DTE = \pi/8$. Figure 4.9 shows the setup of the experiments. Three separate sets of data were recorded:

(i) A 0.5mm 2H pencil was broken at points A_1, A_2, A_3 and A_4. 50 measurements were made at each point.

(ii) A 0.5mm 2H pencil was broken at points A_4, B, C, D and E. Again, 50 measurements were taken at each point.

(iii) Fifty of each 4H, 2H, HB and 2B pencils were broken at point A_4.

4.4.2 Data Analysis

In order to examine the suitability of each of the sets of features, one can measure how well the sets of features fit the labels assigned them, i.e. examine whether a set of features allows one to distinguish between different sources of AE. This was done via two methods: first, the data was visualised using a dimension reduction technique, second, a K-Nearest-Neighbour (KNN) algorithm was run in order to quantify the separability of the data.

Dimension Reduction

In order to provide a subjective measure of the effectiveness of the each set of features, the data was visualised using the PCA technique described in section 5.2. PCA has been used by several AE researchers to investigate data strucure (usually using traditional features) such as Pullin et al. [52], Rippengill et al. [78] and Manson et al. [79].

K-Nearest-Neighbour

Low-dimensional visualisation is not enough to allow an assessment of the utility of given features. For example, the PCA reduction may well superimpose classes when the data is projected into two dimensions where the data is actually perfectly separable in the higher-dimensional feature space. In order to quantify the utility of the features one needs to assess the separability of the classes in the higher-dimensional space. In order to do this, one can apply a standard classification algorithm and then look at the probability of correct classification possible with each feature set.

The choice of which classification algorithm to use requires a balance between the effectiveness of the classifier itself and the computational cost of applying it. The k-Nearest-Neighbours (kNN) algorithm is a simple supervised classification algorithm [80] which provides such a balance. In order to classify an unknown point, one examines the class membership of the k nearest neighbours of that point, and takes a majority vote as to the label. For example, taking an unknown point from experiment C, whose three nearest neighbours are labelled as HB, HB and 2H, results in a label for that point of HB. In this work, k was always set at 3. The neighbours of a point are defined as those which has the smallest Euclidean distance from the point. While better classifiers are available, the kNN classifier is quick to apply and has the nice property that it's error rate is never worse than twice the optimum (Bayes) error rate [81].

In order to quantify the separability of the data given at each wavelet level, leave-one-out validation was performed, and a score was allocated to each data set corresponding to the percentage of correct classifications using this method.

4.4.3 Results

Experiment A results

Figure 4.10 compares the different levels of decomposition with the features collected by traditional parameters.

In experiment A, the AE waveforms were recorded from pencil breaks at different distances from the sensor. The signals will therefore differ because the propagation

of the waves (as Lamb waves) in a plate is dispersive and the waveforms will spread out as they travel further. Figure 4.10 compares the different levels of decomposition with the features collected by traditional methods. The optimum feature set is that corresponding to the eleventh level of DWT decomposition and is based on a set of four features. The classification rate of 97% (for the eleventh decomposition, with 4 parameters) is slightly better than that of 96% for the traditional feature set which has 6 parameters[3]. The classification rates are uniformly high, reflecting the separability of the data. A poor performance is only really obtained for the third decomposition level which is probably mainly constituted by high-frequency noise.

Figure 4.11 shows a visualisation of the data using PCA. The top left plot shows the data using traditional feature extraction, and subsequent plots show the data using the DWT method described here. The four-class structure is visible in all the plots, particularly for $m = 10$, which corresponds to the best performance in the KNN test.

Experiment B results

Figure 4.12 compares the different levels of decomposition with the features collected by traditional parameters.

This experiment involves distinguishing between pencil lead fractures situated radially around the sensor, the main physical reason for the signal differences here will be the variation in speed of wave propagation in the different directions. Figure 4.12 compares the different levels of decomposition with the features collected by the traditional approach. The wavelet features outperform the traditional features at every level apart from the third, with the best level being the eighth decomposition (of length 32). The poor performance at third level is again due to the fact that this level will be dominated by high-frequency noise. The best classification rate is 92% compared with 84% for the traditional parameters.

Examination of the PCA plots of the data (Figure 4.13), shows that five class structure is visible, particularly for $m = 10$. The linear PCA projection has failed to some extent to capture the separability of the data, but it is clear that some clusters are present.

[3]The parameters used were amplitude, rise-time, count, duration, absolute energy and initiation frequency

Since the points A and E are equidistant from the sensor and are aligned along the weave of the axis, one might expect the corresponding data to be inseparable: this is not the case. The separability of these data is due to the aperture effect of the transducer: since the signals are travelling 'under' the transducer from different directions, and the *top* layer of the composite weave is interacting with the transducer, the recorded signals are different, as can be seen visually in Figure 4.13 and also recorded numerically in Figure 4.12.

Experiment C results

Figure 4.14 compares the different levels of decomposition with the features collected by traditional parameters. This set of data represents different hardnesses of pencil leads and is intended here to simulate the different source types associated with different damage and non-damage sources. It proves to be the most difficult to separate using either the wavelet or traditional features. Figure 4.14 compares the different levels of decomposition with the features collected by the traditional approach. The traditional parameters resulted in a 62% classification rate. Some wavelet levels such as the eighth and tenth were higher, whilst the third and fourth levels performed poorly. An improvement is thus obtained by using a DWT feature set with 8 dimensions. (The best results are obtained using a 32-dimensional feature set; however on the basis of preferring lower dimensions, the 8-dimensional set would probably be selected here.)

Again, the data is plotted using PCA in Figure 4.15. The data *is* arranged according to the class label, though some other structure appears in the data. The classes do not appear to be obviously separable, perhaps due to the linearity of the PCA projection. Visualisation of this data is used to demonstrate some non-linear dimension reduction algorithms in the next chapter.

4.4.4 Discussion

This section has demonstrated the applicability of DWT to feature extraction of AE waveforms for a simple classification problem. Principal Component Analysis was used to examine the data, and leave-one-out k-nearest-neighbour was used to generate an objective measure of the data's separability.

The technique does raise the question of which level of decomposition to use for feature extraction. In the simple studies shown, it was not clear that any particular level is better than another, other than the first few levels (which are of high dimension, and consist of very high frequency content) and the last level (of dimension 2) which perhaps is not capable of fully describing the data due to its small dimension.

It may initially appear desirable to choose the decomposition level which maximises the compression (i.e. has the smallest number of coefficients) whilst allowing for successful separation of the data. There is a danger, however, that this procedure will remove information that will become useful in future. A compromise is needed, though I hesitate to make concrete recommendations without further investigation of real data.

Furthermore, this Chapter has focussed on the separation of signals for a set of simple laboratory tests: of real interest to the engineer is to test the procedures on crack-related signals on a real structure. Whilst this work does demonstrate that the use of the presented techniques is viable, it may be unwise to make detailed decisions (such as which level of decomposition to choose for a feature set) based on the data presented here. Nonetheless, the presented practise gives a framework for this work to be done in future.

4.5 Feature Extraction Comparison

In the preceding sections, two feature extraction methods were presented based on wavelet transforms of AE waves. In section 4.3, the features based on the fast wavelet transform (*DWT features*) were compared to the traditional features for some laboratory data. In this section, all three feature extraction methods (Traditional features, Gaussian Mixture Model features and DWT features) are applied to a data set from a fatigue test.

The data used here was first presented in Rippengill et al. [78]. It consists of a series of waveforms recorded during the fatigue testing of a box girder bridge component.

All three feature extraction methods presented in this chapter were performed on the data. In order to visualise the data, a PCA projection (see 5.2) was performed on each data set. K-means clustering was also performed for each feature set, with

k set to 6, as per the original analysis in [78]. The purpose of Figure 4.16 is to examine any correlation between the feature extraction methods; nine sub-plots are shown, with each column showing a PCA projection of a different data set. The rows of sub-plots feature the data coloured according to the labels acquired on the separate feature sets. Thus, for example, the third plot in the first column shows a PCA projection of the *traditional* feature set, but with the points coloured according to the DWT features.

In this work, the number of centres in the GMM feature set was fixed at three, whilst for the DWT features the fifth decomposition layer was used. That signal sampling rate was 2MHz, and there were 2048 samples in each signal.

Note that there is overlap of the classes, even on the diagonal where the PCA projection and the colours of the markers are from the same feature set. This is because the k-means clustering took place in the original (high dimensional) feature space, and the PCA projection naturally cannot preserve in 2D the full structure of the data in this space; nonetheless, it is a useful guide. Note also that it is not necessary for the feature sets to be of the same dimension, since they are being compared only by their 2D representations.

Consider first the sub-plots on the diagonal. In accordance with the original analysis, the traditional features separate reasonably into three groups. Using the GMM features, the structure in the PCA projection does not match the colours so well, indicating that much structure was lost in the projection. In the bottom right (DWT) plot, the structure of the projection matches reasonably with that of the clustering. Colours are not preserved across these plots.

It is now possible to subjectively compare the three feature projections using a single row or column: we proceed by considering the first column. The cluster of signals in the bottom right of the projection (purple in the top left sub-plot) matches well with clusters from the other feature sets, although the DWT features have divided this set into two (red and cyan). This is significant because these signals were identified as crack activity during the original analysis.

Two further groups show good correlation between the feature sets: the cluster coloured red in the top plot matches with that coloured cyan in the second, and a set coloured blue in the top (traditional features) plot matches with a cluster coloured magenta in the bottom plot.

Whilst the above analysis is somewhat subjective, it is gratifying to know that there is some agreement between feature sets. Since the data is unlabelled, it is impossible to perform a rigorous examination of the feature extraction methods, though further work later in this thesis hopefully gives credibility to the use of the DWT feature extraction method. Whilst the GMM method is intuitive and easily understood, it is somewhat impracticable due to the optimisation routine required for each of the signals.

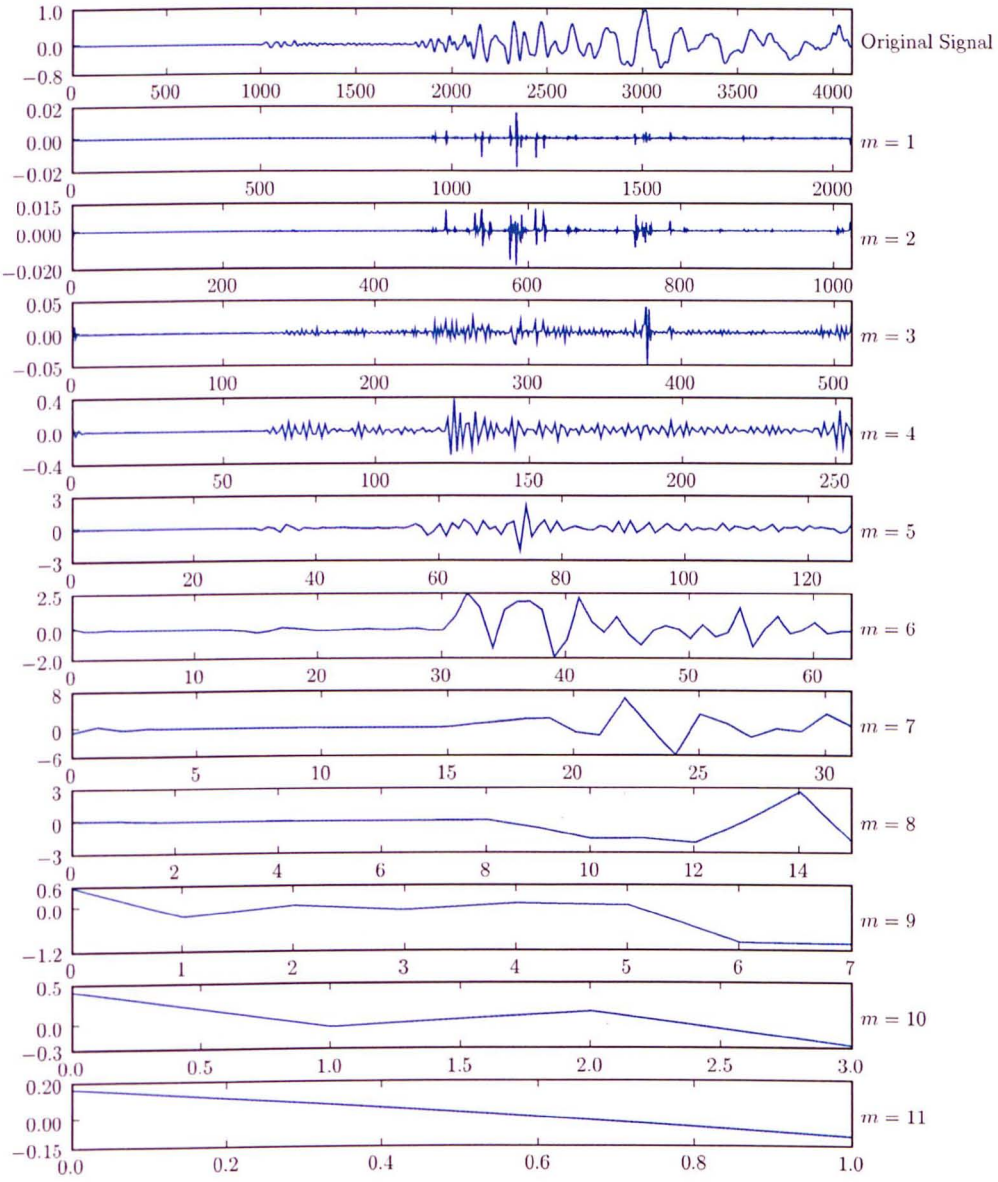


Figure 4.7: Decomposition of an AE signal

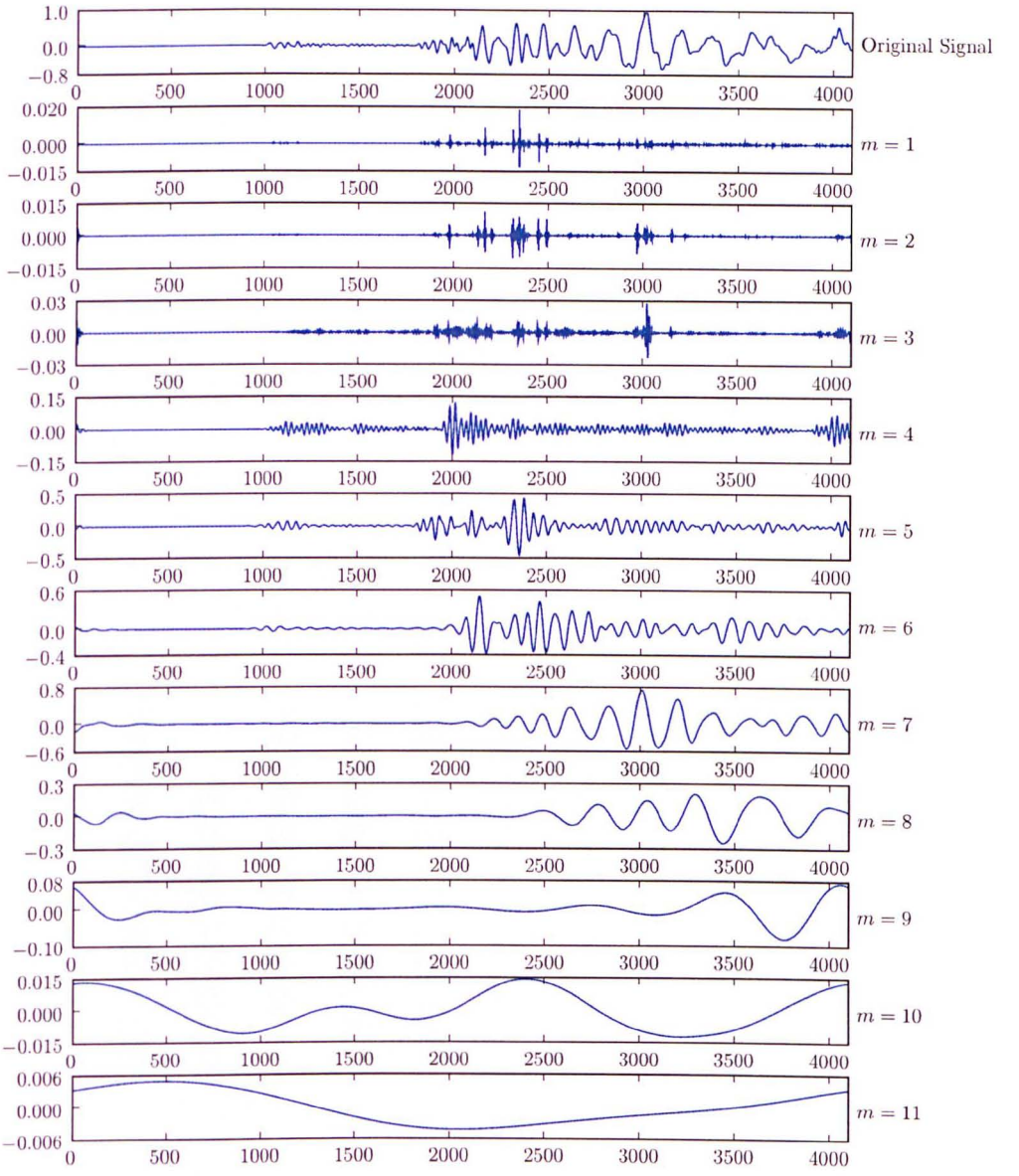


Figure 4.8: Reconstruction of an AE signal

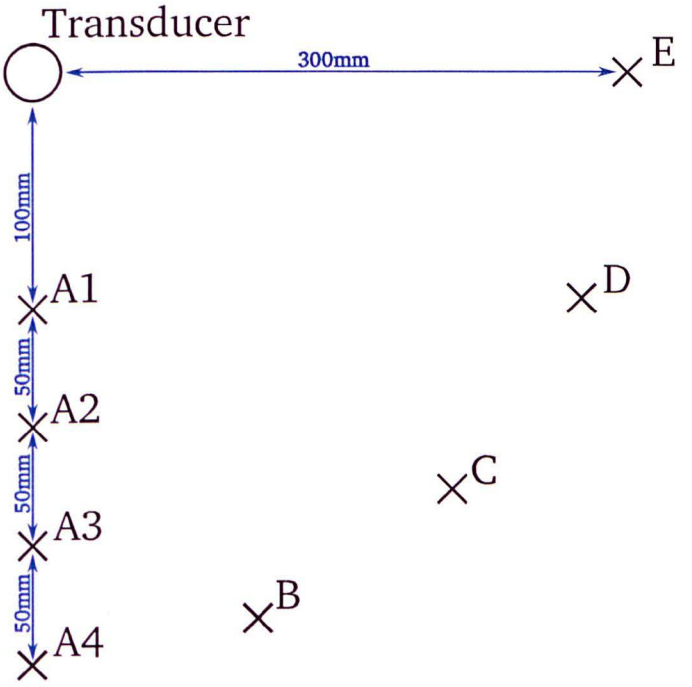


Figure 4.9: Schematic showing the experimental setup for the collection of AE data. the edge of the plate specimen is at least 300mm away in any direction to avoid any reflections.

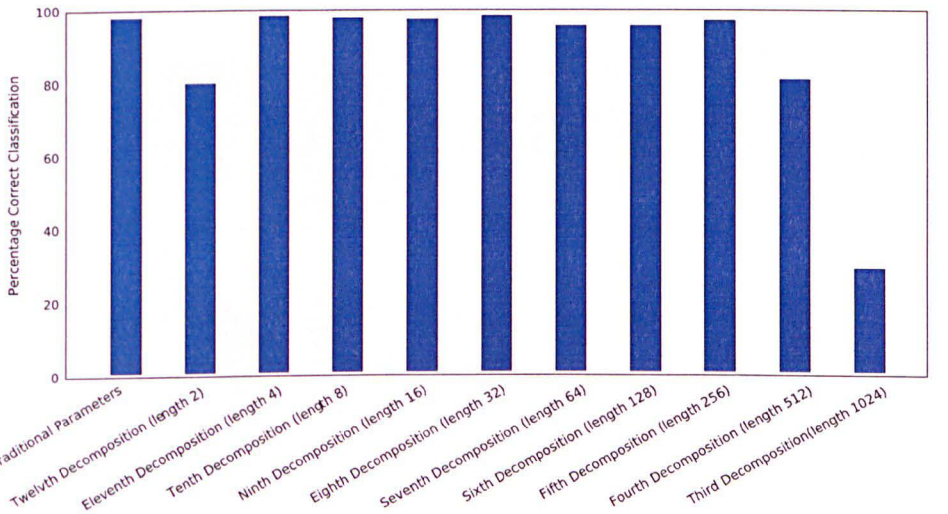
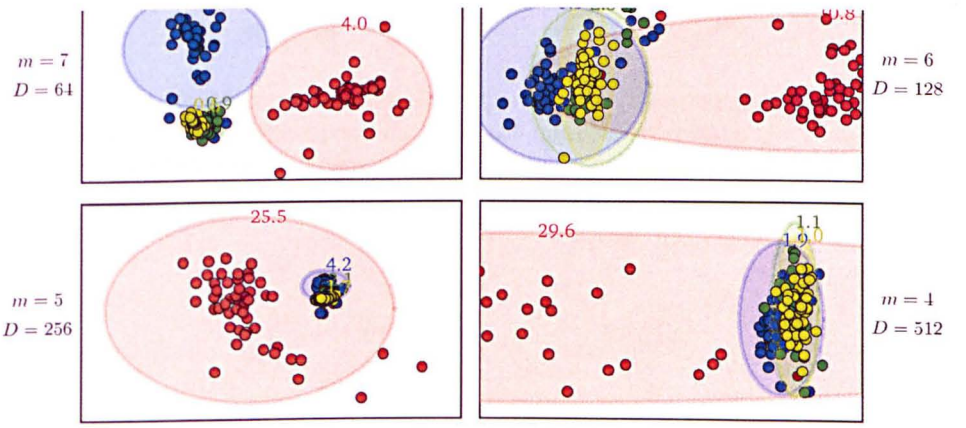


Figure 4.10: Results for experiment A, comparing the effectiveness of the traditional parameters and the various levels of decomposition as features. A wide range of decomposition levels are as effective as the traditional features for this test.


m = 7
D = 64
4.0
m = 6
D = 128
25.5
4.2
m = 5
D = 256
29.6
1.1
m = 4
D = 512

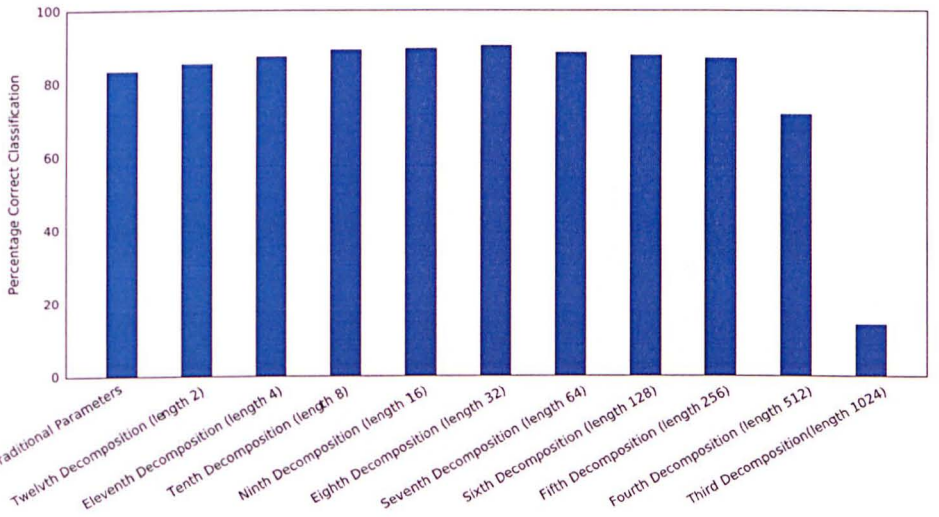


Figure 4.12: Numerical results for experiment B, showing that a wide range of wavelet decomposition levels are more effective than the traditional features for this experiment

IMAGING SERVICES NORTH
Boston Spa, Wetherby
West Yorkshire, LS23 7BQ
www.bl.uk

BLANK PAGE IN ORIGINAL

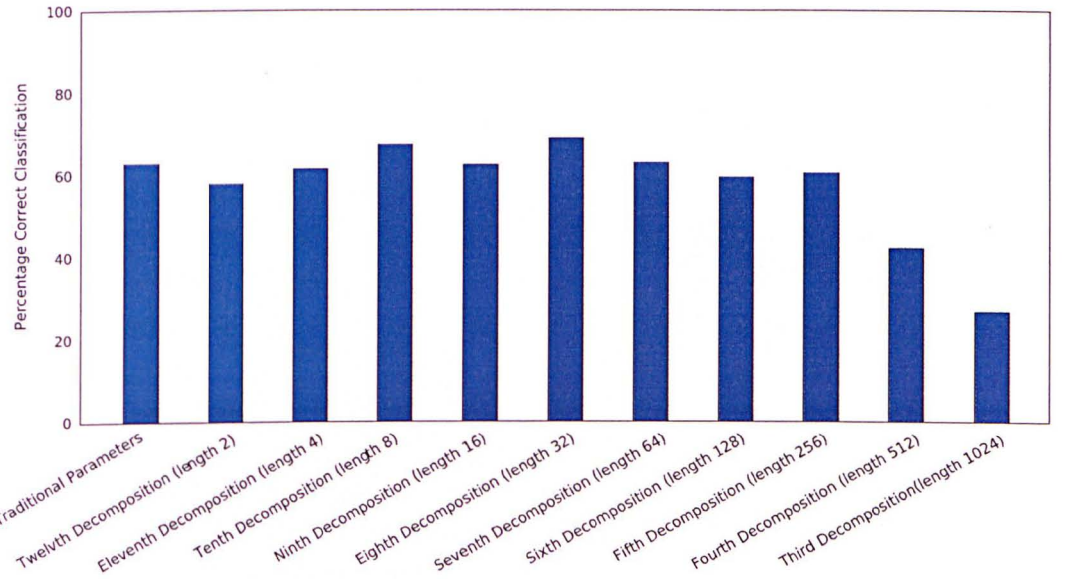


Figure 4.14: Numerical comparison of the results for experiment C, showing that a range of decomposition levels are as effective as the traditional parameters for experiment C.

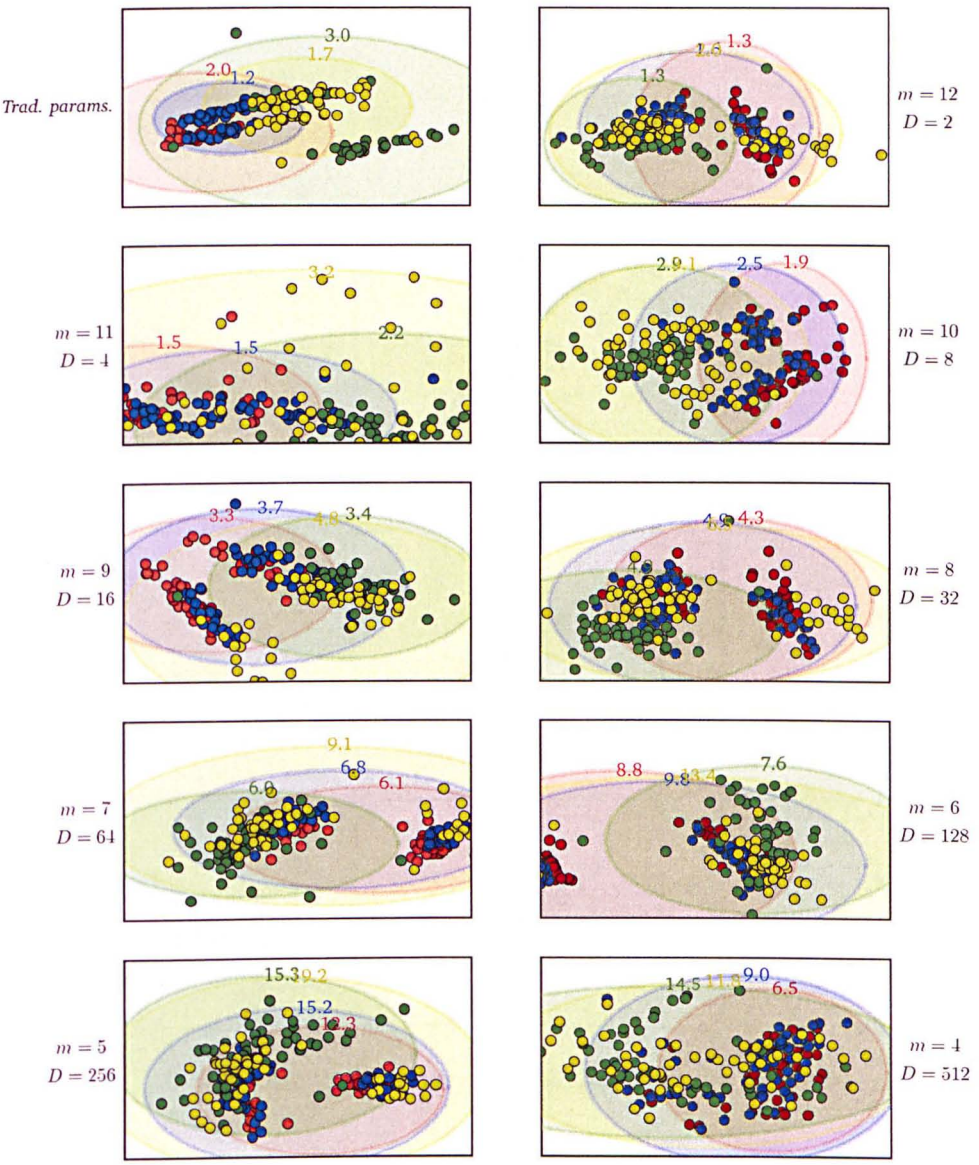


Figure 4.15: Data from experiment C reduced to two dimensions via Principal Component Analysis. The top left plot shows the data using traditional feature extraction, and subsequent plots show the data using decreasing levels of decomposition of the DWT method. The data for each class are surrounded by a ring indicting their spread (two standard deviations), which is labelled at the top of each ring). The dimension of the data after feature extraction, D, is labelled, as well as the level of decomposition, m.

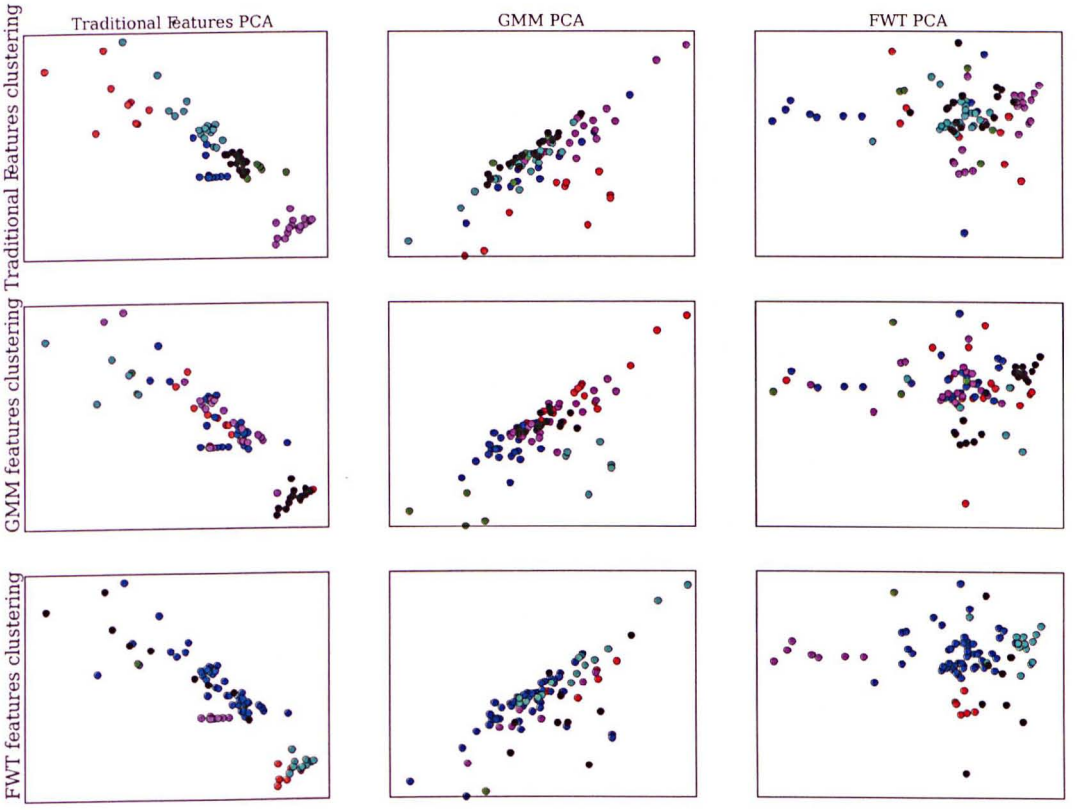


Figure 4.16: A comparison of feature extraction methods

Chapter 5

VISUALISATION

This chapter is concerned with presenting AE data to the user. This is perhaps one of the most important aspects of AE analysis because AE equipment is extremely sensitive, generating vast quantities of data very quickly. Other chapters in this thesis attempt to deal with this quantity of data, rationalising, sorting and classifying it. Here, the emphasis is on how to convey the data to the end user in a way which can be quickly understood. In order to present the data to the engineer, it must be plotted upon either paper or screen, and must therefore consist of only two (or occasionally three for some software impementations) dimensions. The software supplied with AE equipment contains an array of graphing options, however the reduction of the data to 2D in order to utilise these is limited to the simplest of algorithms (PCA, see section 5.2). This Chapter explores a wealth of dimension reduction techniques, applying them to two example data sets.

5.1 Dimension Reduction

The human brain is a formidable tool in recognising patterns and relationships in data. Unfortunately we live in a 3D world, and have difficulty perceiving data of dimension as high as that required in AE analysis. The subject of dimension reduction encompasses a series of tools which can be used to map high dimensional data to a 2 or 3D space, so that it can be interpreted easily.

Typically, the data considered here corresponds to AE signals, from which features

have been extracted in some way (as described in Chapter 4). The purpose of visualising the data might be to see whether the data falls into specific groups, or to asses whether a new set of data corresponds to an old set. Visualisation is also a useful tool in detecting outlying clusters, or seeing whether spatial patterns in the data align with patterns in the feature space.

In this chapter, the running example of pencil lead fracture data *(example 2)* will illustrate how visualisation can be used to distinguish between different source types of AE - in effect, unsupervised source characterisation.

5.2 Principal Component Analysis

Principal Component Analysis (PCA) is perhaps the simplest form of dimension reduction, but nonetheless often proves to be extremely useful. PCA can be derived through many different mathematical interpretations: Nabney [49] shows PCA to be a re-arranging of the *basis* of the data, Roweis and Ghahramani [82] discusses PCA as part of a review of linear Gaussian models, whilst Baldi and Hornik [83] shows that PCA can be performed by a neural network.

PCA is a popular tool because it reduces to an eigenvalue problem, and is therefore (for small data sets) very fast to compute. There follows a simple, practical introduction to PCA.

5.2.1 Theory

Consider an observation $\mathbf{x}$ (e.g. waveform features from a single AE hit as explained in Chapter 4) in a D dimensional space

$$\mathbf{x} = [x_1, \ldots, x_D] \tag{5.1}$$

It is required to map the point $\mathbf{x}$ to a corresponding point $\mathbf{z}$ in d dimensional space where $d < D$, see Figure 5.1

Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]^T$ be an $N \times D$ matrix containing N observations, where each column of $\mathbf{X}$ has zero-mean[1], and $\mathbf{Z}$ be a $N \times d$ matrix containing the corresponding

[1]This is not usually the case for observed data, but it simplifies the notation here and is trivial

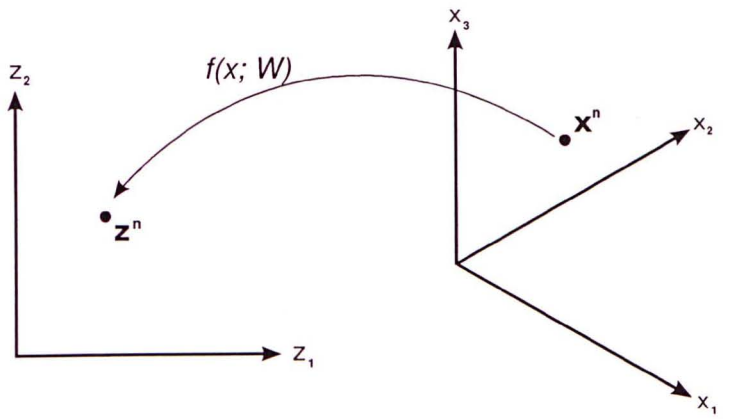


Figure 5.1: Dimension reduction involves mapping observations $\mathbf{x}$ from a D dimensional space to corresponding points $\mathbf{z}$ in a d dimensional space. Here, $D = 3$ and $d = 2$. The map f is parameterised by a set of weights $\mathbf{W}$.

points in the 'latent' lower dimensional space. The covariance matrix $\mathbf{\Sigma}$ of $\mathbf{X}$ is computed as:

$$\mathbf{\Sigma} = \mathbf{X}^T\mathbf{X} \tag{5.2}$$

Since $\mathbf{\Sigma}$ is a square$(D \times D)$ positive semi-definite matrix, it has a full set of D non-negative eigenvectors.

$$\mathbf{\Sigma}\lambda_i = c_i\lambda_i, i = 1, \ldots, D \tag{5.3}$$

The map in PCA consists of a linear projection of the data x by the eigenvectors λ_i which have the largest corresponding eigenvalues c_i, thus $z_i^n = \mathbf{x}^n\lambda_i$ or:

$$\mathbf{Z} = \mathbf{X}\mathbf{\Lambda} \tag{5.4}$$

where $\mathbf{\Lambda}$ is a $D \times d$ matrix containing d eigenvectors of $\mathbf{\Sigma}$. This matrix represents the weights of the map $(\mathbf{W})$in Figure 5.1; in PCA, the map f is a *linear* map.

5.2.2 Simple Example

There follows a complete example, mapping a 2D data set to 1D. Whilst this might seem like a pointless exercise, it is important to remember that if the data were

to implement

to begin in a higher dimension, it would be difficult to observe the effect of the PCA. The method is far more useful in circumstances where the data starts in a much higher dimension , and is projected to 2D for visualisation. The data is shown in table 5.1, and plotted in Figure 5.2. Note that the data is arranged into two clusters, coloured red and blue; it is this structure in the data which we wish to preserve during dimension reduction.

x_1	x_2
0.08	0.27
0.03	0.04
0.06	-0.06
-0.32	0.31
-0.37	-0.21
0.87	1.18
0.97	1.07
0.89	1.24
1.09	0.89
0.97	0.87

Table 5.1: Data set for simple example

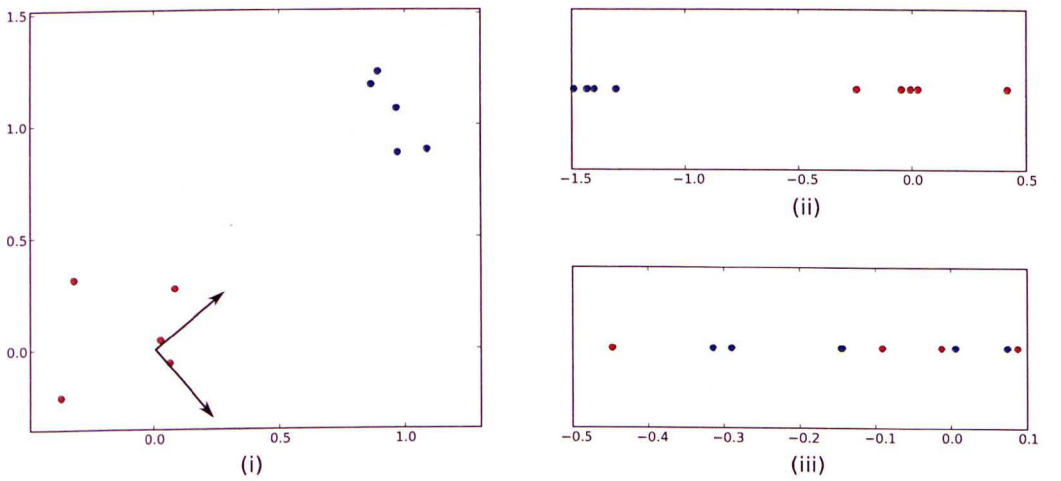


Figure 5.2: A simple example of Principal Component Analysis. (i) the original data and directions of the principal components, (ii) projection of the data along the first principal component and (iii) projection on the data along the second principal component.

The covariance matrix of the data is

$$\Sigma = \begin{bmatrix} 4.86 & 4.97 \\ 4.97 & 5.83 \end{bmatrix}$$

which has eigenvectors $[-0.74 - 0.67]$ and $[-0.67 - 0.74]$ with corresponding eigenvalues 10.33 and 0.35: the eigenvectors are orthogonal. The directions of the eigenvectors are shown in Figure 5.2(i).

The projection of the data along the first eigenvector is shown in Figure 5.2(ii). Note how the structure of the data (the groups coloured red and blue) is preserved. For completeness, we show the projection along the second eigenvector in 5.2(iii), in which it is clear that the data structure is not preserved. The preservation of the data is mirrored in the eigenvalues relating to the two projections, the first being thirty times as large as the second.

5.2.3 Example 1

Throughout this chapter, two running examples will be used. For the first, data will be generated in a high dimensional space $\mathbb{R}^D$, uniformly distributed with variance σ about the points of a $D-1$ simplex. A simplex is a n dimensional analog of a triangle; a triangle (which is a 2-simplex) can be generated in $\mathbb{R}^3$ about the points $[1,0,0]$, $[0,1,0]$ and $[0,0,1]$. Similarly, a tetrahedron can be generated in $\mathbb{R}^4$ as $[1,0,0,0]$,$[0,1,0,0]$,$[0,0,1,0]$ and $[0,0,0,1]$.

To generate the data used in these examples, a 6 simplex was used. Fifty data points were distributed about each of the points of the simplex, using a normal distribution with variance $\sigma = 0.1$.

It is hoped that generating the data thus will provide a reasonable challenge for the dimension reduction schemes presented, and allow the reader to reproduce the results easily.

A PCA visualisation of the simplex data is shown in Figure 5.3. The points are coloured according to the corner of the simplex from which they were generated; these shall be referred to as classes. The PCA algorithm has done a poor job of providing visual separation of the classes: only one class (blue) is separate from the rest. Points from the same class *do* lie close together, but the classes would remain

indistinguishable were the class labels (colours) not known. Since PCA is a linear map, it is unsurprising that the distribution of the points in any class remains (or appears to remain) Gaussian.

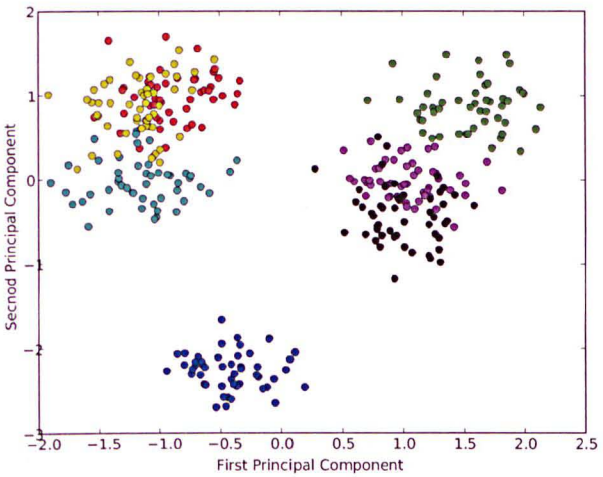


Figure 5.3: Visualisation of the artificial simplex data using PCA

5.2.4 Example 2

In this second example, AE data was generated in the laboratory using a Hsu-Nielsen source on a composite plate. A broadband transducer was used to capture the signals, which were generated using four different hardnesses of pencil: 2B, HB, 2H and 4H. This data was previously used in the feature extraction section, and the reader is guided there for a fuller explanation of the experimental procedure (see 4.4.1, experiment C).

Figure 5.4 shows a PCA visualisation of the data. The data are coloured according to the hardness of the lead used to create the source - it can be seen that there is some separation of the sources. The harder leads (2H and 4H) are more or less separable from the softer (2B, HB) leads. Whilst the plot is informative, this is not an entirely satisfactory visualisation of the data.

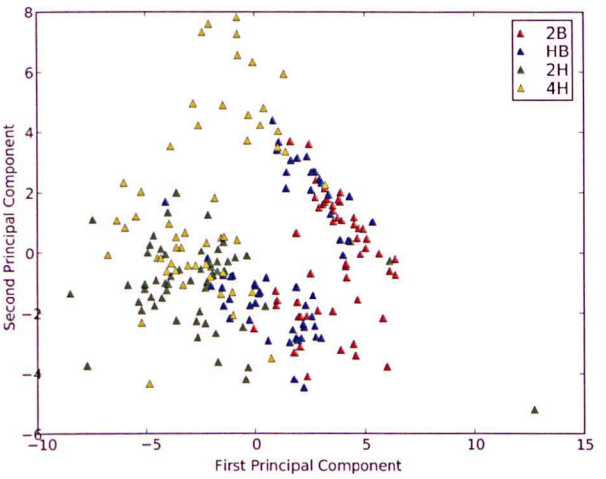


Figure 5.4: Visualisation of the data for example 2 using PCA.

5.3 Latent Variable Models

PCA has been presented as a *projection* of data along the axis of highest variance in a data set. It can also be considered a latent variable model, where the model is a simple plane (in the case of $d = 2$) or hyper-plane (for $d > 2$). The concept is illustrated in Figure 5.5 where a set of 3D data ($\mathbf{x}$) is transformed to a latent space $\mathbf{z}$ by PCA; PCA effectively embeds a plane in the higher dimensional space.

It can be argued that AE data is best visualised by a latent model, since the components of the feature vectors are not independent. This is easiest to illustrate using traditional AE features, but applies equally to wavelet derived signal features. Suppose we are to examine a group of signals from any AE test by their amplitude and energy characteristic; signals with large energy are *more likely* to have high amplitude, because of the physical nature of these measurements. Equally, high frequency signals will have a higher number of threshold crossings. These relationships between the features which are used to describe the signal lead to the feature data lying on some kind of manifold; the data exhibits a *structure* in high dimensional space.

The concept of latent variable models involves approximating the underlying structure of the data with some model, and then arranging the data on a lower dimensional space in such that it represents the data *with* respect to the model in the

data space.

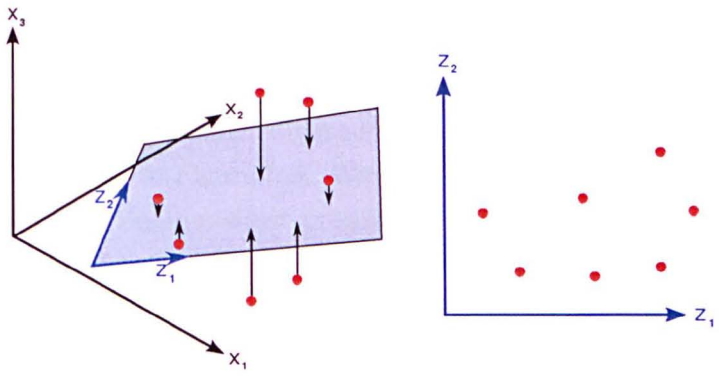


Figure 5.5: Embedding a plane $\mathbf{z}$ in a set of data $\mathbf{x}$ in order to provide a map to a lower dimensional space

The latent model need not be restricted to a plane; indeed this might not be a sufficiently malleable model in order to represent the underlying structure of the data. Various techniques employ more general latent models: Generative topographic mapping uses a manifold represented by a RBF neural network; the Self organising Map [84] uses a series of discrete points in the data space associated with a series of points in a lower dimensional latent space, as does Fast Multi Dimensional Scaling (section 5.6). A recent development in latent variable models is the Gaussian Process Latent Variable Model (GPLVM) by Lawrence [85].

The next section covers one of the more elegant latent variable models, the Generative Topographic Map (GTM).

5.4 Generative Topographic Maps

In this section some theory behind GTM is presented, with an illustrated example; the technique is then applied to the two running examples.

The Generative Topographic Map was first presented by Bishop et al. [86], and later extended in [87]. Applications of the GTM include voice morphing [88], document classification [89] and general data exploration [90].

5.4.1 Theory

The formulae in this explanation of GTM follow the notation of [49].

Consider a map $y(\mathbf{z};\mathbf{W}), y : \mathbf{z} \rightarrow \mathbf{x}$ which maps from a latent (2D) space to a 2D manifold $\mathcal{S}$ embedded in a higher dimensional 'data' space. The map y is parameterised by $\mathbf{W}$ as shown in Figure 5.6. We define a probability density $p(\mathbf{z})$ on the latent space, inducing an equivalent probability density $p(y|\mathbf{W})$ in the data space. $p(y|\mathbf{W})$ is zero away from $\mathcal{S}$, which is unreasonable since the data $\mathbf{x}$ are unlikely to lie precisely on any sensible manifold; we therefore introduce a noise model for $\mathbf{x}$:

$$p(\mathbf{x}|\mathbf{z},\mathbf{W},\sigma) = \frac{1}{(2\pi\sigma^2)^{\frac{D}{2}}} \exp\left(\frac{||y(\mathbf{z}|\mathbf{W}) - \mathbf{x}||^2}{2\sigma^2}\right) \tag{5.5}$$

the density $p(\mathbf{x}|\mathbf{W},\sigma)$ is obtained by integrating out the latent variables $\mathbf{z}$

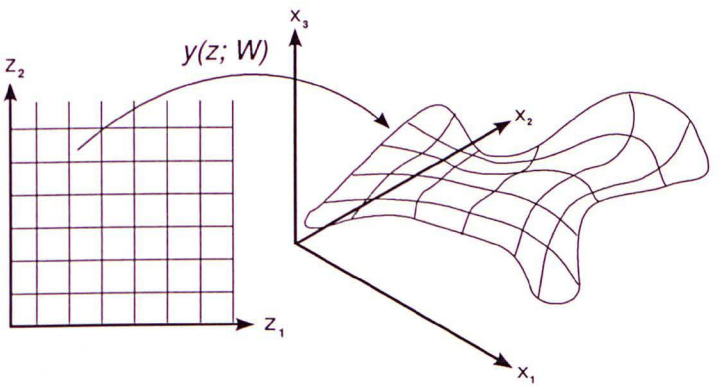


Figure 5.6: A map $y(\mathbf{z};\mathbf{W})$ from a latent space $\mathbf{z}$ to a manifold in a higher dimensional 'data' space $\mathbf{x}$

$$p(\mathbf{x}|\mathbf{W},\sigma) = \int p(\mathbf{x}|\mathbf{z},\mathbf{W},\sigma)p(\mathbf{z})d\mathbf{z} \tag{5.6}$$

if $y(\mathbf{z};W)$ is a linear function, the method is called Probabilistic Principal Component Analysis (PPCA), as described by Tipping and Bishop [91]. Bishop et al. [86] extended the idea for a more general function as described by an Artificial Neural Network; this is called Generative Topographic Mapping (GTM).

If $y(\mathbf{z};W)$ is not a linear function, then 5.6 is analytically intractable. The solution is to allow the $p(\mathbf{z})$ to consist of a series of delta functions centred of a series of

points $\mathbf{z}^m, m = 1, \ldots M$, so:

$$p(\mathbf{z}) = \frac{1}{M} \sum_{m=1}^{M} \delta(\mathbf{z} - \mathbf{z}_m) \tag{5.7}$$

and equation 5.5 becomes:

$$p(\mathbf{x}|\mathbf{W}, \sigma) = \frac{1}{M} \sum_{m=1}^{M} p(\mathbf{x}|\mathbf{z}_m, \mathbf{W}, \sigma) \tag{5.8}$$

which is a Gaussian mixture model centred on $y(\mathbf{z}_m)$, where each element in the mixture has the same (spherical) variance σ.

The dimension reduced data $\mathbf{z}^n$ can be calculated from the model by examining the mean of $p(\mathbf{z}^n|\mathbf{x}^n, \mathbf{W})$:

$$\mathbf{z}^n = \sum_{m=1}^{M} \mathbf{z}^m p(\mathbf{x}^n|\mathbf{x}^m) \tag{5.9}$$

5.4.2 Simple Example

The objective of this example is to illustrate to the reader the principles described above. Consider the 2D data presented in Figure 5.7, which consists of two clusters lying on the function $y = cos(\pi x)$, with some additional random noise. The objective is to remove this underlying structure, and map the data to 1D whilst preserving the separation of the data into two clusters.

Consider 20 points $\mathbf{z}^m, m = 1, \ldots, 20$ evenly spaced across the interval $[0, 1]$; this is our latent space. The function $y(\mathbf{z}; \mathbf{W})$ is represented by a radial basis function neural network, the weights of which are given by the matrix $\mathbf{W}$. The manifold $\mathbb{S} = y(\mathbf{z}; W)$ is shown in Figure 5.8 as a black line. The points $\mathbf{x}^m = y(\mathbf{z}^m; \mathbf{W})$ are represented by black circles. A Gaussian mixture model probability density function (centred on $\mathbf{x}^m$) is shown by the contours. The manifold does pass through some space which is far from the data: here $p(\mathbf{x}) = 0$.

Clearly the manifold has depicted some structure upon which the data lies – in this case it is not the same structure as one might have expected, but it will suffice nonetheless. The Gaussian Mixture Model elegantly depicts the lie of the data. The

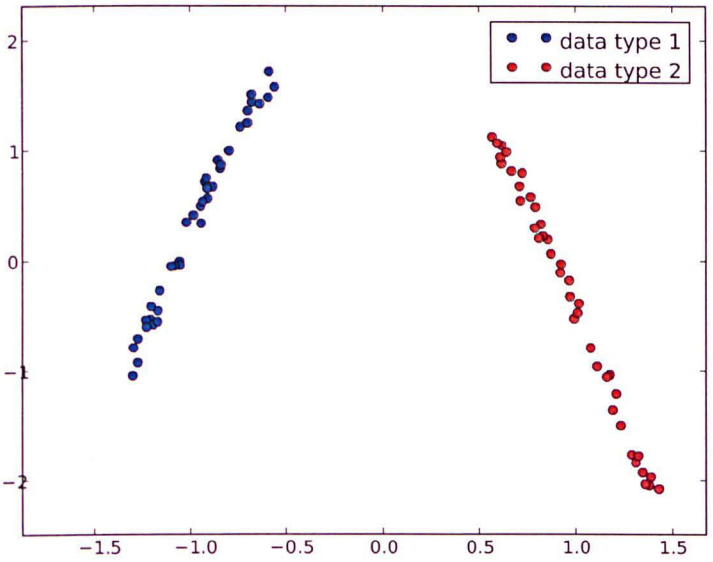


Figure 5.7: Two clusters of data lying on a manifold in $\mathbb{R}^2$

dimension reduced data can now be constructed in $\mathbb{R}^1$, according to equation 5.9; this is shown in Figure 5.9

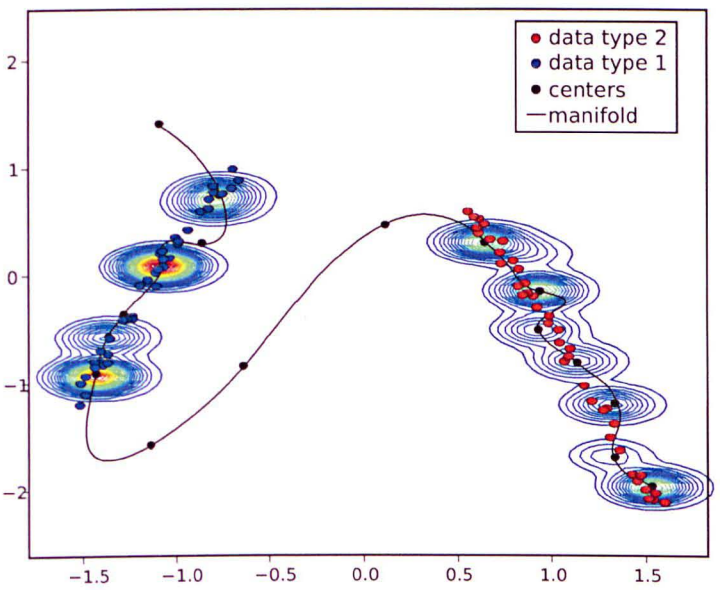


Figure 5.8: The GTM model converged on the data from Figure 5.7. The centres corresponding to a latent space (black dots) lie upon a manifold (black line) which is approximated by a RBF neural network. The probability distribution (a Gaussian mixture model) is shown, centred on the black dots.

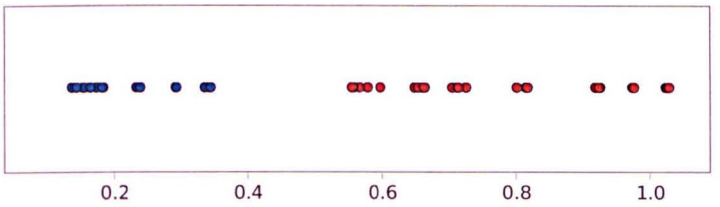


Figure 5.9: The data from Figure 5.7 mapped to $\mathbb{R}^1$ by the model in Figure 5.8

5.4.3 Example 1

Figure 5.10 shows the simplex data described in 5.2.3 reduced to two dimensions by a generative topographic map. The RBF neural network in the map has eight hidden units. The separation of the classes in this figure is the best of all the methods represented in this chapter. The mapping has distorted the structure of the data, however; the classes no longer appear to fall under a Gaussian distribution.

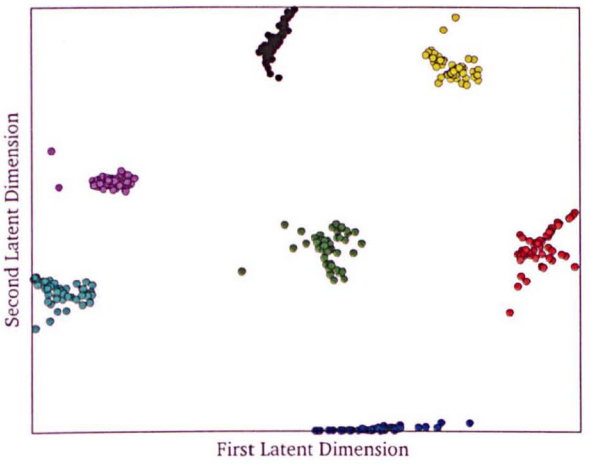


Figure 5.10: The simplex data described in section 5.2.3 reduced to two dimensions via a Generative Topographic Map

5.4.4 Example 2

Figure 5.11 show the data described in 5.2.4 reduced to two dimensions via a GTM. Some separation of the classes is apparent, however the structure of the data on

the page is not consistent with a technical understanding of the experiment: it could reasonably be expected that data from the harder pencil leads (2H and 4H) would be distant from the softer ones (2B and HB). This is in part the case, but a further structure has been imposed upon the visualisation, whereby there is a large separation within classes.

This additional structure is due to the mapping of the neural network in the GTM. If the map becomes too complicated, it will *over-fit* the data, and exert its own structure on the dimension reduced data. This is an issue of generalisation; it is a hard issue to solve since there are no target values for a GTM, and hence no means of measuring the accuracy on a 'validation' set, as is normal when using neural networks for regression or classification (see Chapter 6).

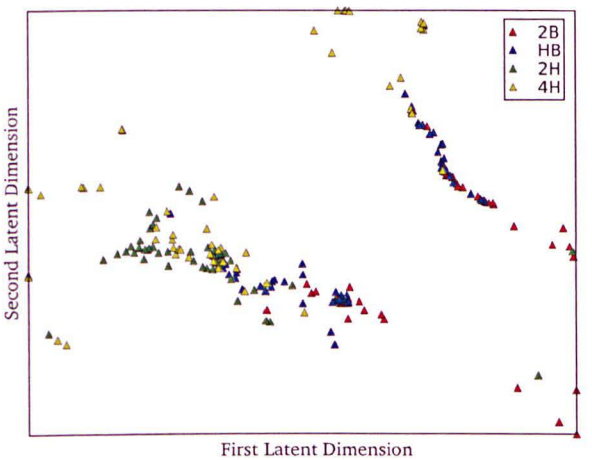


Figure 5.11: A GTM dimension reduction of the artificial AE data generated by pencil lead fractures on a composite plate as described in section 5.2.4

5.5 Multi-dimensional Scaling

Multi-dimensional Scaling (MDS) takes a different approach to dimension reduction to the methods described previously. It is often referred to as a 'distance preservation' algorithm. The concept is simple, one defines a stress term ε which is minimised in order to keep distances in the data space $\mathbb{R}^D$ similar to distances in the latent or

projection space (usually $\mathbb{R}^2$). An example of a stress term (the Sammon stress [92] in this case) is:

$$\varepsilon = \sum_{i=1}^{N} \sum_{j=1}^{N} |\mathbf{x}_i - \mathbf{x}_j| - |\mathbf{z}_i - \mathbf{z}_j| \tag{5.10}$$

where $\mathbf{x}_{i,j}$ are data points in $\mathbb{R}^D$, and $\mathbf{z}_{i,j}$ are corresponding points in $\mathbb{R}^2$. The algorithm is usually initiated by selecting a random set of $\mathbf{z}$, and optimising ε with respect to $\mathbf{z}$. The optimisation can be performed using any one of a number of methods, such as particle swarms [71] or differential evolution [73].

5.5.1 Disadvantages

The main disadvantage of the MDS method for dimension reduction is its computational inefficiency. In order to evaluate the stress function, $N(N-1)/2$ distances must be calculated. When considering the optimisation algorithm, the dimension of the problem is $N \times d$ (i.e. d coordinates must be calculated for each point). This means that the amount of computation required for MDS is unreasonably high for large data sets.

5.5.2 Example 1

This section continues with the running example of the simplex data described in section 5.2.3. Figure 5.12 show data scattered about the points of a 6 simplex in $\mathbb{R}^7$, reduced to two dimensions using MDS. The algorithm used for optimisation was a particle swarm, with 20 particles running for 600 iterations.

Figure 5.12 shows the data grouped into clear classes, though there is much overlap between the classes. In all likelihood, the algorithm would produce a better result if it was left to run for longer, though the computational expense of the method makes this undesirable for any real AE application.

5.5.3 Example 2

Similarly, Figure 5.13 shows a MDS representation of the data described in section 5.2.4. Again, the algorithm used for optimisation was a particle swarm, with 20

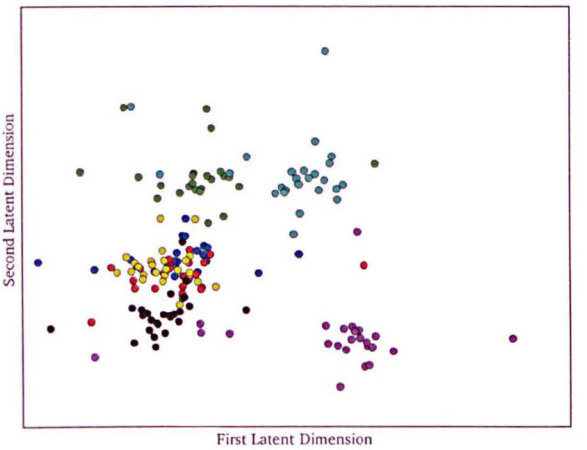


Figure 5.12: Multidimensional Scaling applied to data spread about the points of a 6 simplex. The different symbols represent data associated with different points of the simplex. The optimisation algorithm used was a Particle Swarm.

particles running for 600 iterations. There is a some separation of the classes into the harder (2H, 4H) and softer (2B, HB) signals, though the separation is not perfect. Once more, the representation could perhaps be improved by running the simulation for longer, but the computational expense of doing so outweighs any potential benefit.

5.6 Fast MDS

The computational expense of the MDS algorithm makes it prohibitive for tasks with large data. This motivated Lowe and Tipping [93] to devise the Neuroscale algorithm, where a neural network is trained to map points in a high dimensional space to a lower one, using the Sammon stress between the two spaces as an error function.

Independently, The author devised the Fast MDS algorithm, which shares some

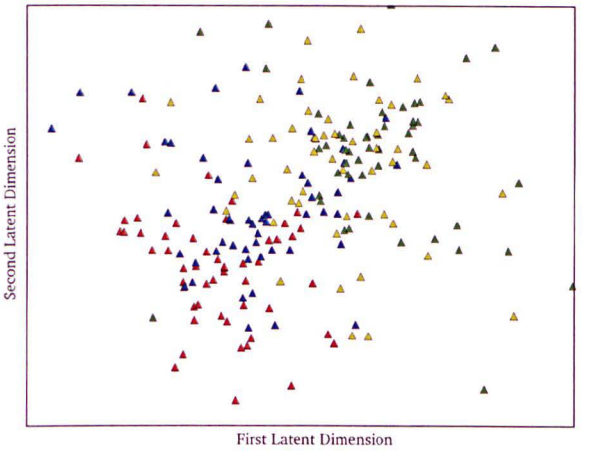


Figure 5.13: Multidimensional Scaling applied to signals collected from the experiment described in section 5.2.4

traits with Neuroscale (the reference points $\mathbf{x}^m$ here are analogous to the centres of the neural network in Neuroscale). The motivation for Fast MDS was to produce an algorithm which was distance preserving like MDS, but without the computational expense. The algorithm is outlined below.

5.6.1 Theory

In Multidimensional Scaling, in order to calculate the Sammon stress of one individual pair of points $(\mathbf{z}^n, \mathbf{x}^n)$, one must calculate the relationship between that point and all other points, as in equation 5.10. This means that for a dataset with N points, $N(N-1)$ distances must be calculated. Since the distances in the high dimensional set $\mathbf{x}$ need only be computed once, this leaves $\frac{N(N-1)}{2}$ distances. However, most of these distances are of little consequence: it simply is not necessary to describe the relationship of one point in a data set by it's precise relationship with *every* other bit of data.

Fast MDS introduces a series of reference points $\mathbf{x}^m, m = 1, \ldots, M$ into the high

dimensional space. The simplest way to assign these is to place them at the centres of M k-means clusters (see section 3.2.2). These points are associated with M points in the latent, lower dimensional space: $\mathbf{z}^m, m = 1, \ldots, M$. The points $\mathbf{z}^m$ are arranged such that the Sammon stress between the sets $\mathbf{x}^m$ and $\mathbf{z}^m$ is minimum, à la Multidimensional scaling. Providing that M is small, this is a trivial operation.

Having established a set of pairs of points $(\mathbf{x}^m, \mathbf{z}^m), m = 1, \ldots, M$, the optimal position of $\mathbf{z}^n$ given $\mathbf{x}^n$ is given by minimising equation 5.11:

$$E_n = \sum_{m=1}^{M} |\mathbf{x}^n - \mathbf{x}^m| - |\mathbf{z}^n - \mathbf{z}^m| \tag{5.11}$$

The optimisation of E_n must be performed for each data point $\mathbf{x}^n$ in order to generate corresponding data points $\mathbf{z}^n$. For a set of N data points which are desired to be visualised in $\mathbb{R}^2$, we now have N small ($2D$) optimisation problems, where evaluating the cost at any stage requires calculating M distances. Compare this with the optimisation problem in MDS, where there is one large optimisation problem, of dimension $2 \times N$, where the number of distances required to calculate the cost at any point is $N(N-1)/2$.

5.6.2 Similarity to a SOM

The Self Organising Map (SOM) [84] is a dimension reduction algorithm similar to the Fast Multi-dimensional Scaling described here. The SOM algorithm takes a set of points in the latent space, and pairs them with a set of points in the data space $(\mathbf{z}^m, \mathbf{x}^m)$. The points $\mathbf{x}^n$ are then presented with the data, and move towards the data in the data space; the update algorithm ensures that points $\mathbf{z}$ which are close in the *latent* space are also close in the data space. Dimension reduction is performed by representing the data $\mathbf{x}^n$ in the latent space at the point $\mathbf{z}^m$ such that $\mathbf{x}^m$ is the closest reference point to $\mathbf{x}^n$. Typically, the SOM requires more reference points, since the latent space is discrete.

5.6.3 Example 1

Fast MDS was applied to the simplex data as described in section 5.2.3: Figure 5.14 shows the result. Seven pairs of points were chosen to reference the data in the high

and low dimensional space ($M = 7$). The method manages to give perfect visual separation of the classes, whilst retaining the individual class distributions (i.e. the spread of points in each class appears spherical-Gaussian).

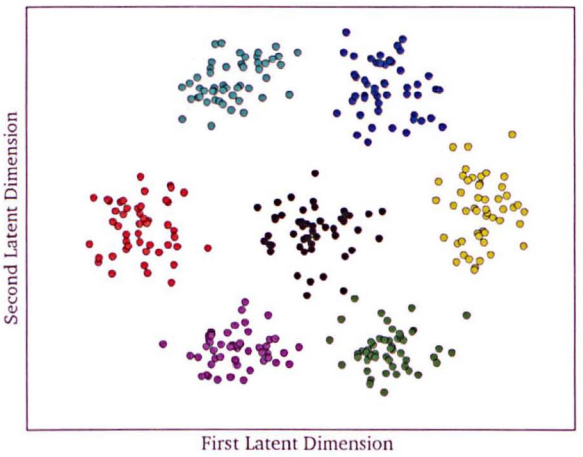


Figure 5.14: Fast MDS applied to simplex data. Classes appear completely separable and class distributions appear uniform

5.6.4 Example 2

Figure 5.15 shows the data as described in section 5.2.4, dimension reduced using fast MDS. This is the most satisfactory visualisation of this data from a practical point of view: signals which one would expect to be similar lie reasonably close together (i.e. signals from HB pencils lead fractures lie together). There is some overlap of the data - the groups are not entirely separable; however the underlying structure of the data is in accordance with our expectation of the experiment: The hardest lead (4H) lies furthest from the softest (2B).

5.7 Kernel PCA

Kernel PCA (KPCA) was proposed by Scholkopf et al. [94], and an optimised version was later presented by Weinberger et al. [95]. In KPCA, the data is first projected into a higher dimensional space using the kernel trick, and PCA is performed in this

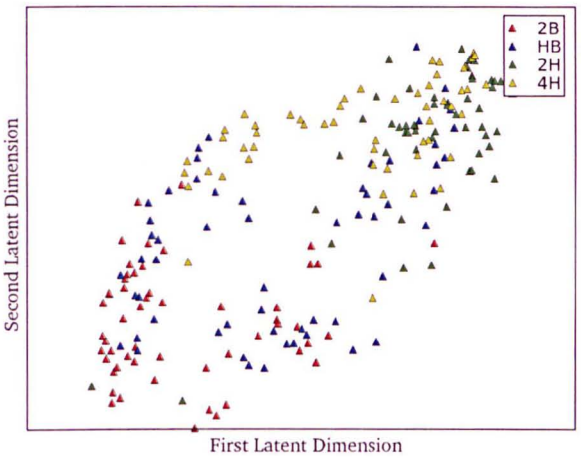


Figure 5.15: Fast MDS representation of the example pencil-lead data

space. KPCA therefore requires no nonlinear programming, and is computationally simple like the original PCA. In fact, the original PCA algorithm is recovered if the kernel is linear, such that $k(x, x') = xx'$.

5.7.1 Theory

Consider a map ϕ from our data space to a feature space $\mathcal{F}$

$$\phi : \mathbb{R}^N \rightarrow \mathcal{F} \tag{5.12}$$

The covariance matrix in $\mathcal{F}$ is given by

$$C = \phi(\mathbf{x}_j)\phi(\mathbf{x}_j)^T \tag{5.13}$$

Performing PCA in this space requires the eigenvalues λ and eigenvectors $\mathbf{V}$ such that

$$\lambda \mathbf{V} = C\mathbf{V}. \tag{5.14}$$

since the solutions $\mathbf{V}$ lie in the span of $\phi(\mathbf{x}_1), \ldots, \phi(\mathbf{x}_m)$, one can consider

$$\lambda(\phi(\mathbf{x}_k)\dot{\mathbf{V}}) = (\phi(\mathbf{x}_k) \cdot C\mathbf{V}) \tag{5.15}$$

Since the dimensionality of $\mathcal{F}$ is very high (possibly infinite) we recognise that there must exist coefficients α_i such that

$$\mathbf{V} = \sum_i \alpha_i \phi(\mathbf{x}_i). \tag{5.16}$$

Recall the kernel trick used in support vector machines[96]:

$$\mathbf{K}_{i,j} := (\phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j)) \tag{5.17}$$

substituting 5.16 and 5.17 into 5.15,

$$\lambda\alpha = \mathbf{K}\alpha \tag{5.18}$$

The KPCA algorithm was applied to the running examples in this chapter. The result of the simplex data is shown in Figure 5.16, and the result of the pencil lead break data is shown in Figure 5.17.

Figure 5.16 shows that the algorithm is capable of separating the classes for this simple case. The nonlinearity of the transform means that the Gaussian structure of the classes is lost, and that the distances between the classes are not at all uniform, but the algorithm has succeeded. In Figure 5.17 however, there is no clear distinction between the four groups. There is some clustering of the 2H and 4H data in the top left corner, but in general it is difficult to distinguish between the groups.

5.8 Manifold Detection

Several algorithms approach the problem of dimension reduction by considering the *local* structure of the data. These include Locally Linear Embedding (LLE)[97] and Isomap [98]. LLE works by treating the relationship between any point and its k neighbours as linear. One solves multiple least-squares problems in order to find a set of weights such that each point can be linearly reconstructed from its neighbours. In the low-dimensional space, LLE solves the problem of finding a set of data such that these *weights* are reconstructed.

Isomap also works on local information, by calculating the *connectivity* of the data;

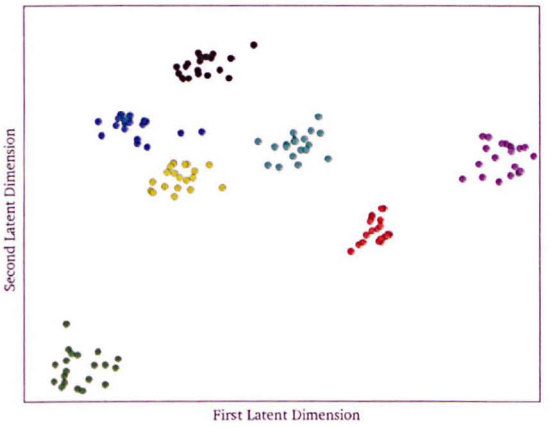


Figure 5.16: Results of the KPCA algorithm on the example simplex data. Good separability of the groups is clear in this case.

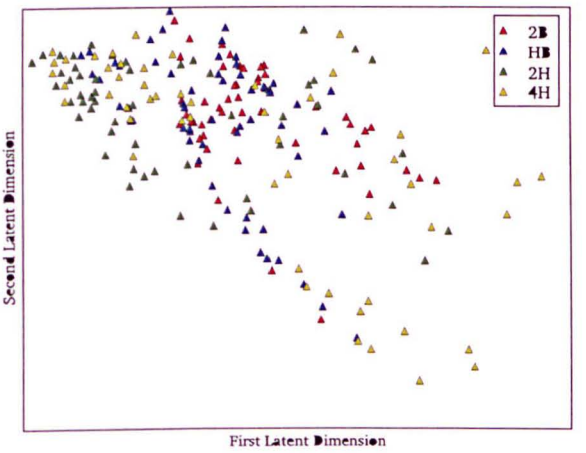


Figure 5.17: Results of the KPCA algorithm on the example pencil lead break data. There is no great distinction between the four groups, however there is some trend toward a group consisting of 2B and HB, and another of 2H and 4H.

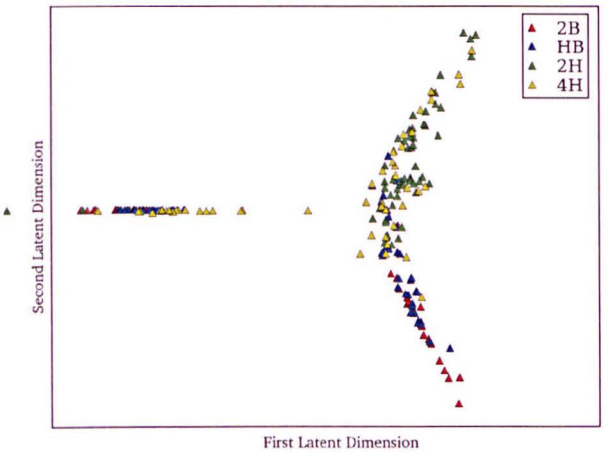


Figure 5.18: An Isomap projection of the pencil-lead data described in section 5.2.4

connectivity is calculated by either some radius or k neighbours. Projection into a lower dimensional space is done in the same manner as for multi-dimensional scaling, but the distances in the high-dimensional space are taken to be *through* the connected 'manifold' of points.

An advantage of the Isomap algorithm is that if some sections of data in the high dimensional space are distinct from the others, this will be represented in the connectivity. When applying Isomap to the simplex data which is a theme of this chapter, the algorithm instantly detects the 7 distinct clusters. An individual low-dimensional representation can then be made of each distinct group (since each group of the simplex data is a rather boring normal distribution, this is omitted here).

Isomap was applied to the second set of example data in this chapter, and the result is displayed in Figure 5.18. Like many of the previous algorithms, the harder pencil lead signals are distinct from the softer ones, but the two hard data sets (2H and 4H) are indiscriminable, as are the softer ones (2B and HB).

5.9 Dimension Reduction For AE

The dimension reduction algorithms presented provide a range of options for visualising AE data. This, combined with suitable location data (see Chapter 7) should provide an excellent starting point for AE data exploration. Indeed, the visualisation technique has lead to successful results for a test of a small landing gear component as part of this project [52].

When considering the choice of dimension reduction technique, there is some trade-off to be made between the power of the technique and its computational cost. This section has presented a series of techniques which are suitable for acoustic emission, and the results of the considered techniques have been presented on two data sets. This should give the reader some insight into the computational limitations of the algorithms. When considering speed, one should consider how each algorithm scales with the number of data to be dimensionally reduced - clearly the more data present, the more computations are necessary.

The selection of dimensionality reduction technique is also dependent on the nature of the test: where the user requires to apply the technique on-the-fly, it may be best to run the simple PCA algorithm. Where the user requires only batch processing, say for daily inspection, it may be convenient to employ a range of techniques and make them all available to the user.

Whilst the use of toy data is useful in the context of this text in order to illustrate the nature, uses and pitfalls of the various techniques, when selecting a technique for the monitoring of a specific AE application, it may be more useful to compare the methods on a set of well understood, relevant AE data.

Chapter 6

SOURCE LOCATION

The ability to calculate the source position from the time-differences between the arrival of an AE wave at a series of sensors is a great strength of the AE method. In Chapter 3, the proposed method for detecting fractures using AE began with location of the source.

Whilst the problem may at first appear simple, for most real engineering structures it is not, as shall be seen in this chapter. The first section in this chapter examines the simplest method for locating AE which is built-in to current AE software. The second section introduces some experimental data in order to illustrate the benefits of onset picking; the data will be used again in order to demonstrate a machine learning algorithm for AE source location, which is the topic of the third section. The chapter finishes by making comparisons on the methods using data collected on a main-tube landing gear component.

6.1 A Trigonometric System

The purpose of an AE location system is to find the location of the event, E, given the position of the sensors and the timing information for the event. The mathematics is non-trivial, largely due to the unknown quantity t_0, the time of the event. A brief explanation is given in [99], but this section aims to explore the mathematics further.

6.1.1 Location Mathematics

Consider three AE transducers on a uniform homogeneous plate at positions $S_i = (x_i, y_i)$, $i = 1, 2, 3$. An AE event occurs at position $E = (x_E, y_E)$ at time $t = t_0$, and the subsequent sound field arrives at the sensors at times $t = t_1, t_2, t_3$. It is the nature of an acoustic emission test that t_0 is unknown, and the only information available is $t_2 - t_1, t_3 - t_1$ and $t_3 - t_2$ which shall be denoted $\Delta t_{2,1}, \Delta t_{3,1}, \Delta t_{3,2}$ respectively.

Under the homogeneous plate assumptions, it is possible to define a speed of sound c which is independent of position or angle. This is quite a limiting assumption: it will not hold either for composite plates, where the speed of each wave mode is dependent on the fibre angle, or for plate-like structures with thickness changes, holes, curvature etcetera. Further, modal wave theory is ignored here, and only the speed of the fastest wave mode is considered. The simple relationship between speed, distance and time gives:

$$\begin{aligned} \Delta d_{1,2} &= c \times \Delta t_{1,2} \\ \Delta d_{1,3} &= c \times \Delta t_{1,3} \\ \Delta d_{2,3} &= c \times \Delta t_{2,3} \end{aligned} \tag{6.1}$$

The quantity Δd is a distance, which defines:

$$\begin{aligned} |ES_1| - |ES_2| &= \Delta d_{1,2} \\ |ES_1| - |ES_3| &= \Delta d_{1,3} \\ |ES_2| - |ES_3| &= \Delta d_{2,3} \end{aligned} \tag{6.2}$$

The source position E therefore lies on the intersection of the three hyperbolae defined by equation 6.2.

A hyperbola is a locus, defined such that the difference between the distances to two *foci* is constant. In this application, the foci are the positions of the sensors, S_i and the constant is Δd. The derivation of the equation describing a hyperbola is given in appendix 4, and results in the same equation as for an ellipse:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{6.3}$$

where: $a = \Delta d/2$, $b^2 = (\frac{|S_i S_j|}{2})^2 - a^2$, and the foci S_i lie on the x axis, centred about

the origin. In order to describe the hyperbola for a general pair of sensors (which do not necessarily lie on the x-axis, as in Figure 6.1), one must rotate and translate the axes x, y into the coordinate system of the sensors, using a rotation matrix.

6.1.2 Solving the Hyperbolae

In order to solve equation 6.3 for the position (x_E, y_E) of the event, one must substitute the hyperbolae for every pair of used sensors and solve. This yields an equation which is analytically intractable, it is therefore preferred to use a minimisation procedure to find the position of the event (x_E, y_E). The function to be minimised is:

$$Z = \sum_{i=0} N_{pairs} |\Delta T_{measured,i} - \Delta T_{model,i}| \tag{6.4}$$

where $\Delta T_{model,i} = \frac{|ES_1| - |ES_i|}{c}$, and $\Delta T_{measured}$ is the measured value. It is clear from (6.4 that the purpose of the optimisation is to match the measured ΔT values against those of the model. The motivation for writing the equation in this manner is that the model may be replaced later by some better understanding of the wave propagation, for example from collected data as in section 6.6.1. There are many different options available for solving equation 6.4 including the Nelder-Mead simplex routine [100].

6.2 Sensitivity to Noise

During a real AE test there are potential errors in both time information, sensor position information, and the estimated speed of sound. This section will demonstrate that the system is fairly insensitive to errors for events located inside the array, but is more sensitive for those events outside the array.

Figure 6.1(a) shows a hypothetical array of three sensors and a source of acoustic emission located within the array. The hyperbolae in solid lines intersect at the point of the event, denoted ▲. The dashed lines show the hyperbolae calculated when there is a 5% error in the ΔT information. In this case, the small errors will give rise to only a small error in the location of the source. Figure 6.1(b) shows the

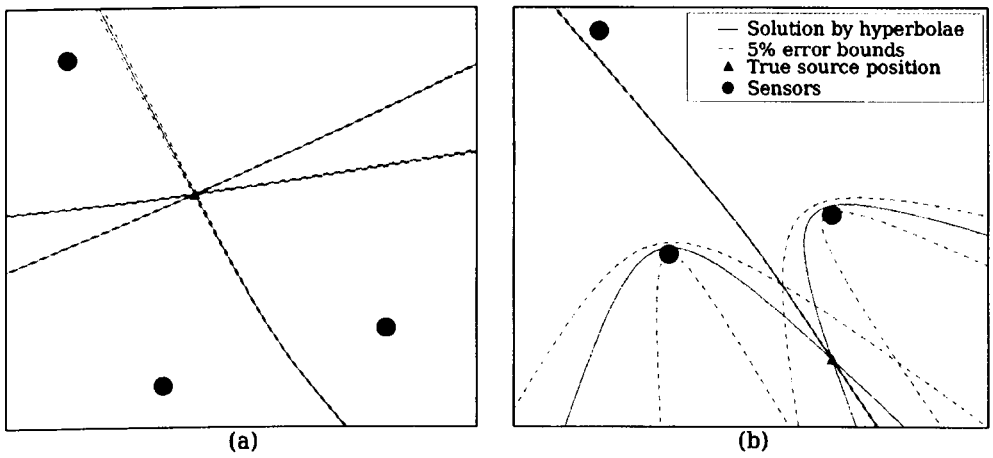


Figure 6.1: Source location by solution of hyperbolae for events inside an array of sensors (a) and outside the array (b)

same array of sensors and the hyperbolae for locating an event outside of the array. The solid lines (representing the solution with zero error in the measurements) have an intersect at the true location of the source, but the dashed lines now show a magnification of the error in terms of position of the solution.

It is apparent from Figure 6.1 that the magnification of the errors is due to the fact that as one move the true location outside the array, the hyperbolae which must intersect at the solution become closer and closer to being parallel. In order to reduce the uncertainty, one would have to add another sensor such that the resulting additional hyperbola ran perpendicular to the existing ones. It is trivial to show that to do so requires adding a sensor which encompasses the true location!

For small errors in ΔT, it is clear that *inside* the array the errors in location will be small. However for the same errors in ΔT for points outside the array, the errors can be much larger.

6.2.1 Simulated Sensitivity Analysis

It is possible to evaluate the sensitivity to noise of the system described above for a series of points around the array of sensors. Given the location of a hypothetical set of sensors, one can easily calculate the ΔT information for any given point.

Corrupting this ΔT with some small error and feeding it back into the location algorithm gives a new location, with some error. The distance between the known (initial) point position and the one calculated by the algorithm from the corrupted data gives an indication as to the sensitivity of the system to errors in the ΔT information for the given point. Performing this procedure on a mesh of points surrounding the sensors, we obtain a contour plot as shown in Figure 6.2. The sensor positions are marked by green circles; the ΔT information was corrupted by a 1% error.

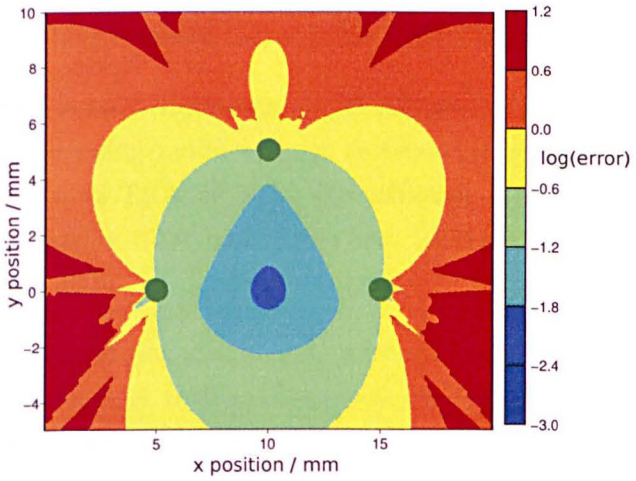


Figure 6.2: A contour plot showing the sensitivity of a location system to 5% discrepancy in the ΔT information. The area inside the array of sensors (denoted by green dots) shows low sensitivity, whilst areas outside the array show a higher sensitivity

Examining Figure 6.2, it is clear that the areas *inside* the array are less sensitive to the corruption of the ΔT information than those *outside* the array. This is in accordance with the example shown above in section 6.2, and with the experience of AE practitioners (for example Holford et al. [101]).

6.3 Experimental Methodology

The following two sections (6.5 and 6.6) make use of a data set collected as part of this project. This section describes the acquisition of the data set.

Four specimens (denoted Plate A ... D) were made from 3mm mild steel sheet, dimensions 200 x 370mm. The specimens were designed to replicate the challenge of locating Acoustic Emissions in complicated engineering structures: a series of holes at the centre of the plate provide geometric obstacles to the wave propagation path, inducing reflections and scattering. Since future work involves fatigue testing of the plates and location of the test acoustic emissions, a further series of holes at the ends of the plates provide a mechanism for the fatigue rig - this further complicates the wave propagation path. Eight Sonox P5 discs measuring 8mm diameter, 0.2mm thickness fitted with wrap-around electrodes were bonded to each the plate using cyanoacrylate. The positions of these sensors are depicted in Figure 6.3. The sensors have a thickness resonance of approximately 3MHz, though they are utilised here somewhat outside their resonant band in the region 100kHz – 1MHz.

An eight-channel Physical Acoustics DISP AE system was utilised to record all waveforms. The recording mode was set to record synchronous waveforms across all channels - known as TRA mode in the software. The sampling frequency was 10MHz. the anti-aliasing filter had bandwidth 100Hz – 3MHz and a pre-amplifier was used with gain 40dB.

Thermoelastic expansion was used to generate artificial AE in the plate by means of a high-power laser pulse. A Brilliant laser was positioned above the plate and fired in synchronism with the AE equipment.

In order to generate artificial AE sources across a grid on the plate, a positioning rig was built. The tolerance in the rig's absolute position was $5\mu m$. The positioner, with a plate specimen attached is shown in Figure 6.4.

6.4 Onset Picking

One cause of noise in the measurement of arrivals of signals (and hence noise in the ΔT information) is the way in which the onset of a signal is defined. The most

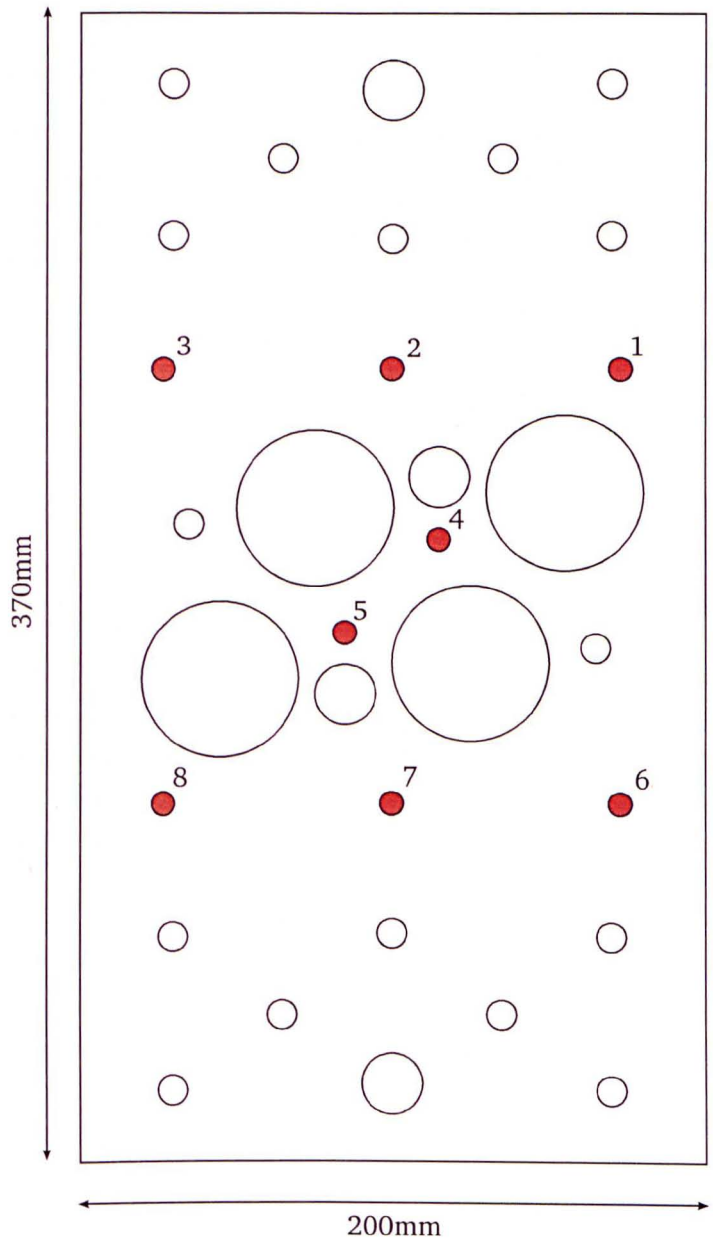


Figure 6.3: Schematic of the plate specimen used. Transducers are represented by numbered solid circles; hollow circles represent holes in the plate.

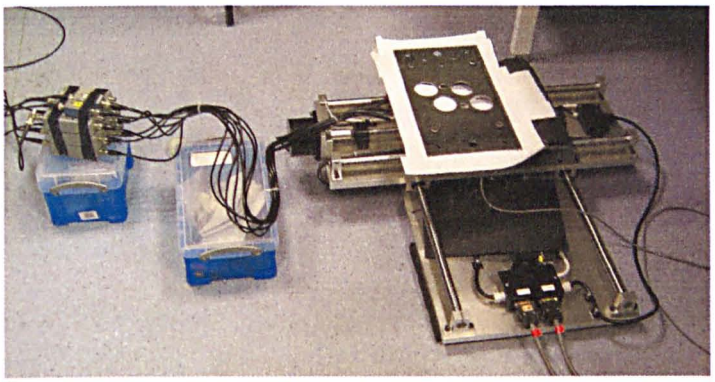


Figure 6.4: Photograph showing the X-Y positioner system holding a plate specimen underneath the laser. The plate is mounted sensor-side down, and paper shields protect cabling from the laser. The AE pre-amplifiers are also shown to the left of the rig.

common way to select the onset of a signal is to choose the time at which the signal first crosses some user defined threshold. Not only is this somewhat arbitrary, it is prone to significant errors due to the modal nature of Acoustic Emission waves.

In a situation where a small wave packet arrives before a larger one (which is often the case in acoustic emission, see section 2.4), if a user defines a threshold which is approximately the same voltage as the peak of the first packet, then uncertainty in the voltage of the waveform will cause the recorded onset time to bifurcate. Two similar waves, with only slightly different voltages, will be recorded at very different onset times.

Kurz et al. [102] devised a method whereby the onset of a signal was picked not by a user defined threshold, but by examination of the waveform statistically. This shows improvement over the threshold crossing when compared to onsets picked by the human eye.

6.5 Onset Picking Implementation

Initial inspection of the data revealed some interesting patterns due to the geometry of the plate, albeit with some anomalies in the data. An onset picking algorithm based on the work of [102] was used to detect reliable onsets of the signals, resulting in the correction of many of these observed anomalies.

Figure 6.5 is an illustration of some of the data collected during the experiment, it shows the ΔT information for the sensors pair (4,6), which shall be denoted ΔT_{46}. The laser was used to generate ultrasound at a series of points arranged in a uniform grid (avoiding the specimen holes) across the entire surface of the plate; each point is represented by a small solid circle in Figure 6.5, which is coloured according to ΔT_{46} A threshold crossing point was used to determine the arrival times.

The areas close to both sensors show a smooth transition of ΔT_{46} with space: over most of the lower part of the plate the ΔT_{46} function changes slowly and smoothly. In the upper left quadrant of the plate, where the path to either of the sensors (particularly sensor 6) is blocked by a series of holes, ΔT_{46} can be seen to change more abruptly and chaotically. Much of this is due to the threshold method used to calculate the onset times of the waveforms, as shall be seen.

Figure 6.6 shows eight waveforms, received on channels 1-8 after a laser excitation of the plate at position (70, 270). The horizontal red lines show the threshold used to detect the waveforms, and thereby determine their onset times. The effect of *dispersion* [21] is evident here: waves which have travelled a long distance show the emergence of a lower amplitude, fast travelling mode from the main wave. This is particularly evident in waveforms from sensors 6, 7 and 8: these were the most distant sensors from the laser source position. This results in the threshold crossing occurring some time after the perceptible arrival of the wave (threshold crossings are marked with vertical red lines).

In order to detect the earlier onset of the wave, one could lower the threshold, enabling the system to detect the lower voltage part of the waveforms. This would increase the risk of a false trigger, however. The problem of onset picking was addressed by [102], who presented a range of methods of picking the onset of a signal *without* a user defined threshold. In this work, we utilise the Akaike Information Criterion picker (AIC), given by:

$$AIC(t) = t \log_{10}(var(x[1;t])) + (T - t - 1) \log_{10}(var(x[t;T])) \tag{6.5}$$

where $var(x)$ is the variance of x, i.e. $\sum_i (x_i - \overline{x})^2/N$, and $x[1;t]$ represents the section of x from 1 to t. $AIC(t)$ is represented in Figure 6.6 by a green line; the minimum of AIC gives the onset of the signal (vertical green line). The distance between the vertical red and green lines gives the 'correction' made by the onset picking algorithm. It is clear that the more dispersive waveforms are heavily corrected by

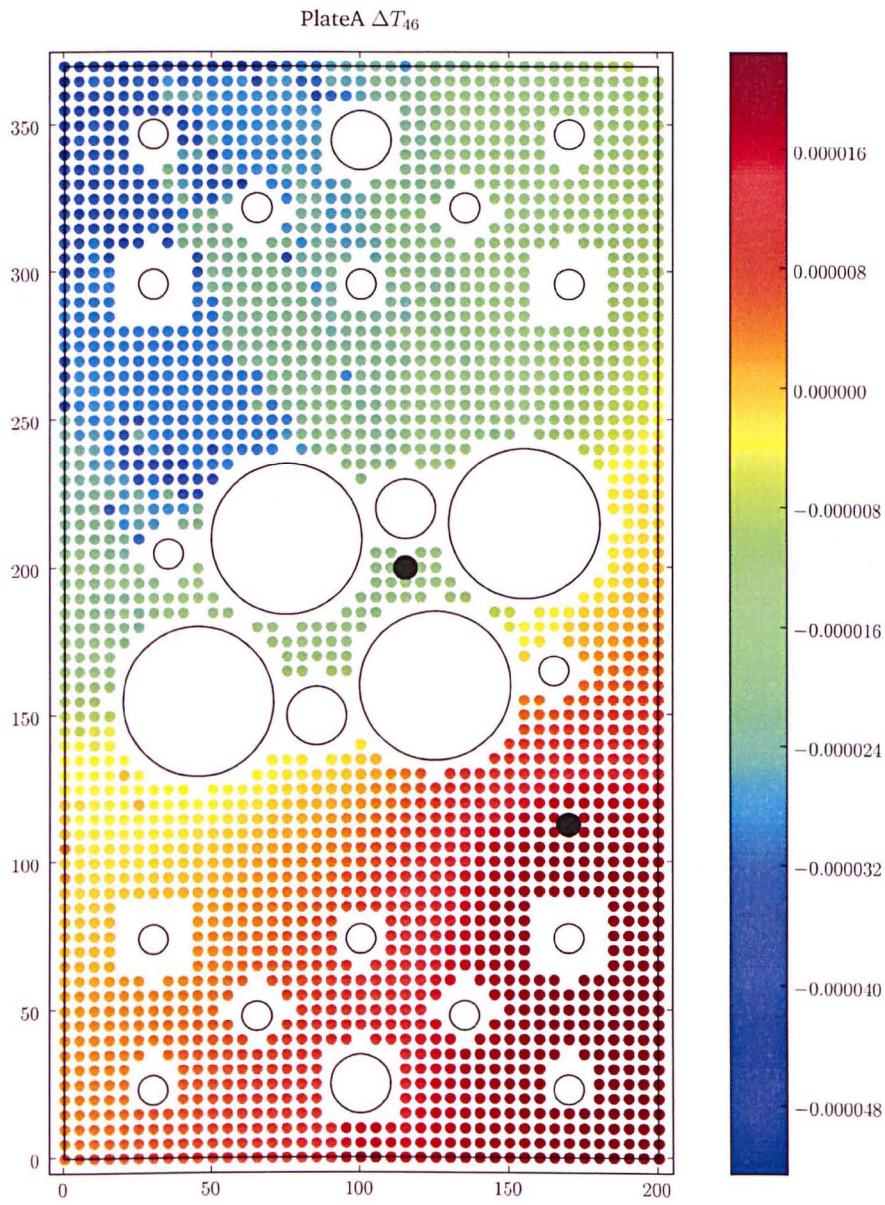


Figure 6.5: Scatter diagram showing ΔT_{46}, the difference between arrival times at sensors 4 and 6, for each point measured on plate A, as calculated by the time of threshold crossing. Each small circle represents one activation of the laser, and is coloured according to ΔT_{46}. The position of sensors 4 and 6 are marked by solid black circles.

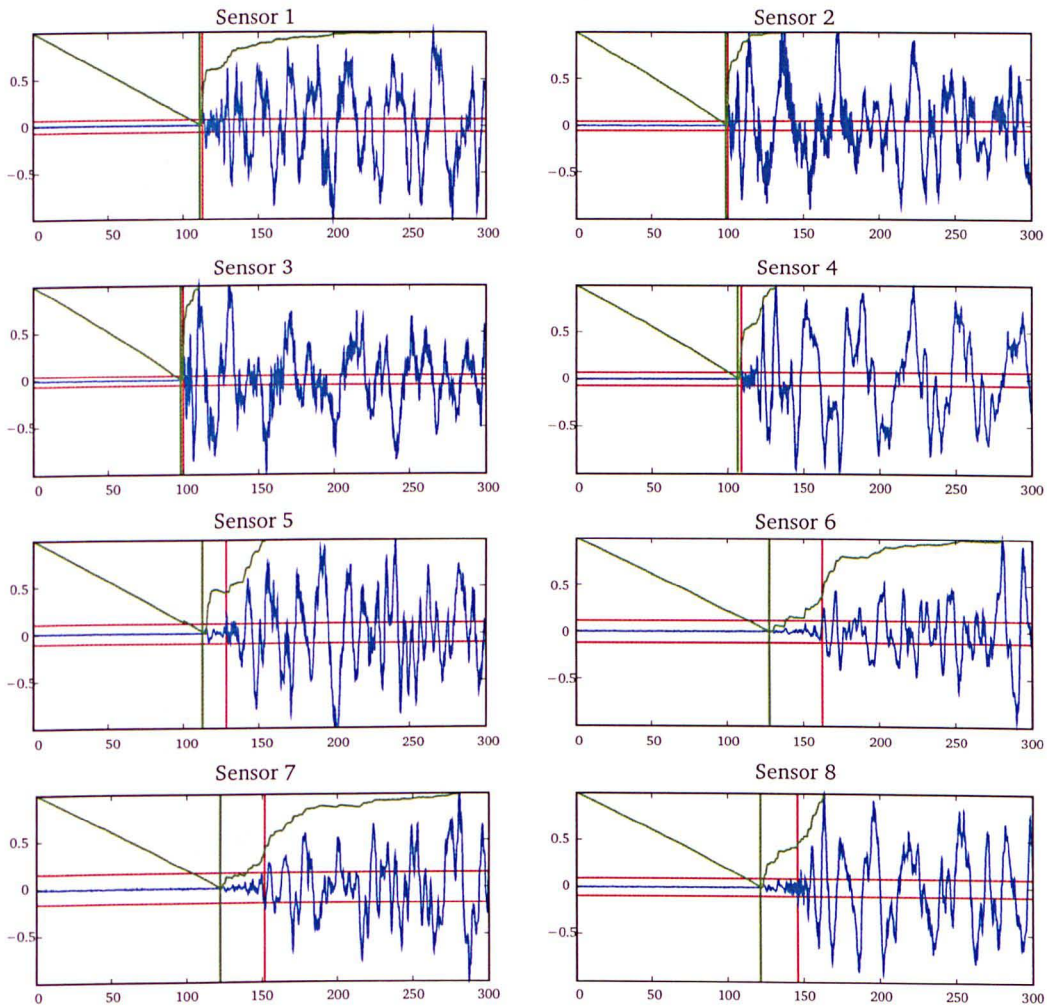


Figure 6.6: Waveforms received at sensors 1-8 due do a laser impulse at point (70,270). The thresholds used are shown in red, and the onset picking equation is shown in green. The onsets as picked by the thresholds and the onset picking equation are represented by vertical lines. The vertical axes have been normalised for clarity.

the onset picking (channels 5–8 in Figure 6.6) whilst those nearer to the source (channels 1–4) are corrected by a smaller amount.

Examining equation 6.5 in more detail: for every time t, the signal $x(t)$ is broken into two parts; that leading up to t ($x[1;t]$) and that after t, ($x[t;T]$). AIC describes the similarity in entropy between the two parts of the signal - when t is aligned with the onset of the signal, $AIC(t)$ is minimum.

This statistical onset-picking has a particularly important role when using machine learning for the task of locating acoustic emission. It enables reliable onset picking of the waveform *regardless of the amplitude of the wave.* Hence when attempting to locate acoustic emissions in a test situation, the amplitude of the test signals need not be the same as those used to train the learning algorithm.

Figure 6.7 is comparable with Figure 6.5, but with the onset times calculated by equation 6.5. It is clear that many of the onset times in the upper left quadrant have been corrected, whilst the pattern in the lower part of the plate remains the same. Some anomalies remain, particularly in the 'hard to reach' corner of the plate, and in the area around (45, 220) which lies in the 'shadow' of several holes.

6.5.1 Summary

It is clear from the data presented here that the onset picking algorithm is extremely useful. The barrier to implementing this is relatively low, since the application of equation 6.5 requires only a small number of calculations: in the experiments presented here, the computational cost is trivial.

6.6 Machine Learning for Location

In the section 6.1, the general process for locating AE sources and problems therein was described. The major flaw in using this system is the assumption that the speed of sound c is constant across the medium.

This problem has been addressed for composite structures by Coverley and Staszewski [103], who devised an algorithm whereby Acoustic Emissions from impact in composite sheets could be located. The algorithm described took account of the changing

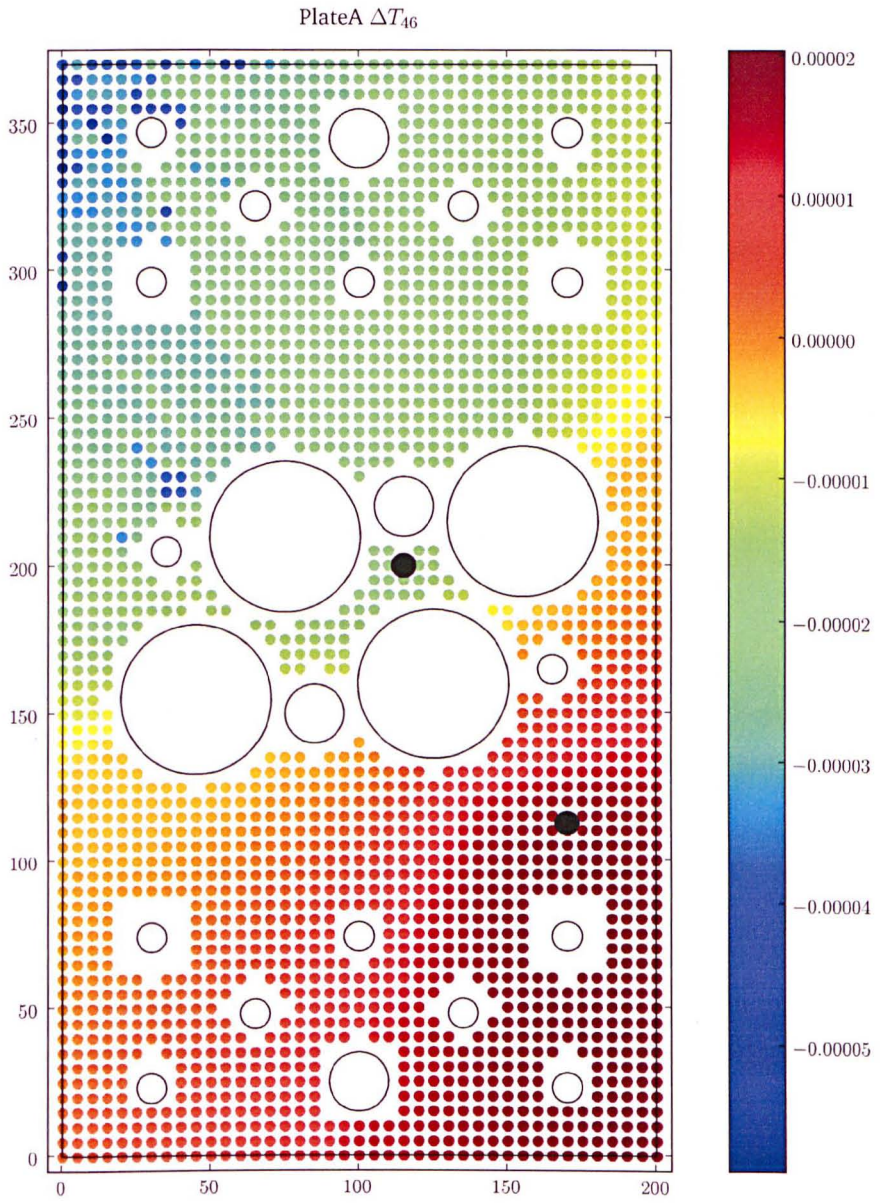


Figure 6.7: Scatter diagram showing ΔT_{46}, the difference between arrival times at sensors 4 and 6, for each point measured on plate A, as calculated by use of the AIC picker defined by Kurz. Each small circle represents one activation of the laser, and is coloured according to ΔT_{46}. The position of sensors 4 and 6 are marked by solid black circles.

wave velocity with direction relative to the fibre. Schubert [104] also addresses the problem by means of Acoustic Emission tomography.

Whilst these methods are capable of dealing with anisotropic materials, they do not account for the biggest problem in location of AE, that of geometry. When locating AE in structures such as landing gear, the propagation path includes holes, lugs, changes in thickness and possibly changes in material properties. AE wave propagation is highly sensitive to the propagation path (see section 2). This was motivation for Baxter et al. [105] to devise a method for location where the information about the different wave-speeds is held in a series of ΔT-*maps*.

This section describes some improved methods for calculating the ΔT, and then go on to describe a method whereby the maps are bypassed altogether, using a kernel method to represent a function which gives accurate locations.

6.6.1 The ΔT Interpolation Method

Baxter et al. [105] proposed a method for learning the relationship between position and the time difference (ΔT) between a pair of sensors. Using these learned ΔT maps, accurate source location of acoustic emissions can be achieved.

The method works by constructing a grid of points across the structure to be investigated, and breaking ten pencil leads (the Hsu Nielsen technique) at each point. The ΔT for each pair of sensors is measured for each pencil lead. Results are mean-averaged across the ten pencil leads. Linear interpolation is used to create *maps* of the ΔT information which can then be used to replace the model section in equation 6.4.

This technique is shown in Figure 6.8: two sensors were placed on a main landing gear fitting at positions $x = 0$ and $x = 40$, and leads were broken at grid points between them[1]. Note that the data is one-dimensional in this case, as there is only one sensor-pair. In order to locate AE in a planar fashion, the functions approximated must be surfaces, and one surface is required for each sensor pairing.

Once the ΔT maps have been obtained, the position of an AE event can be derived

[1]this is a subset of the main landing gear data used in section 7.2

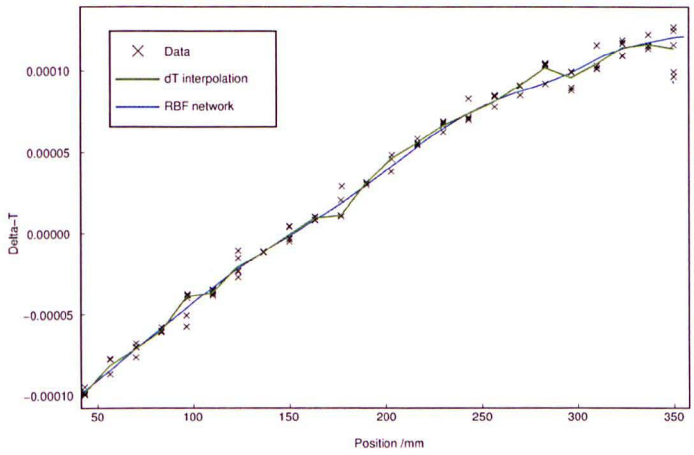


Figure 6.8: The ΔT interpolation method applied to some linear data. The data is shown by black crosses, and the interpolation technique is shown as a green line. A simple RBF neural network is also shown in blue.

using equation 6.6:

$$Z = \sum_{i=0} \Delta T_{measured,i} - \Delta T_{model,i} \tag{6.6}$$

which is the same as equation 6.4, except that the model ΔT is derived from our interpolation method.

6.6.2 Gaussian Processes Regression

Neural networks have long been used as regression tools [55]. More recently, the framework of Bayesian statistics has been applied to the field in order to provide a probabilistic viewpoint. Neal [106] provides a comprehensive Bayesian treatment of Neural Networks, and shows that in the limit of an infinite number of hidden units, a Bayesian Neural network becomes a Gaussian Process (GP), see also MacKay [107] for a discussion. More recently, work by Rasmussen and Williams [108] has provided an excellent framework for using GPs for machine learning tasks. There follows a brief introduction to the ideas of Gaussian Processes in a regression context.

Consider a series of observations $X = \{x_i\}_{i=1}^N$ with corresponding noisy target values $Y = \{y_i\}_{i=1}^N$, $y_i = f(x_i) + \sigma_n$. Assuming that f is a smooth function which maps $x_i \rightarrow f(x_i)$, we wish to infer new target values $f(x_*)$ given unlabelled points x_*. Using the Gaussian Process framework, a Gaussian prior is placed jointly across

$f(x_i)$ and $f(x_*)$.

Consider momentarily only two points $x_1\, x_2 \in X$, corresponding to the target values $f(x_1), f(x_2) \in Y$. Writing

$$p\left(\begin{bmatrix} f(x_1) \\ f(x_2) \end{bmatrix}\right) = \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} k(x_1,x_1) & k(x_1,x_2) \\ k(x_2,x_1) & k(x_2,x_2) \end{bmatrix}\right) = \mathcal{N}(\mathbf{0}, \Sigma) \tag{6.7}$$

places a *prior* Gaussian distribution across the values of $f(x_1)$ and $f(x_2)$, centred around $\mathbf{0}$, with covariance Σ. The elements of Σ are given by the *covariance* function k; only the squared exponential function is considered as a covariance function here (others are discussed by [108]):

$$k(x_1, x_2) = \exp\left(\frac{-|x_1 - x_2|^2}{l^2}\right) + \delta_{1,2}\sigma_n \tag{6.8}$$

Where $\delta_{1,2}$ is the Dirac function. The diagonal elements of Σ are $1 + \sigma_n$, and the other elements specify that the dependence of $f(x_1)$ on $f(x_2)$ is given by the distance between x_1 and x_2. This is illustrated by the top row in Figure 6.9: on the left is the probability distribution across $f(x_1)$ and $f(x_2)$ when the variables x_1 and x_2 are far apart, i.e. $k(x_1, x_2) = 0$; in the middle column, x_1 and x_2 are closer together ($k(x_1, x_2) = 0.5$), and in the right column $x_1 \approx x_2$, hence $k(x_1, x_2) \to 1$.

Suppose now that one wishes to infer the value of $f(x_2)$ given that the value of $f(x_1)$ has been observed; this is simply conditional probability of $f(x_2)$ given $f(x_1)$, which has a standard result which can be found in many textbooks, e.g. [109], see also [108] for the result in the context of GPs. The result is:

$$p(f(x_2)|f(x_1)) = \mathcal{N}\left(k(x_2,x_1)k(x_1,x_1)f(x_1),\ k(x_2,x_2) - k(x_2,x_1)k(x_1,x_1)^{-1}k(x_1,x_2)\right) \tag{6.9}$$

This is again illustrated in Figure 6.9, the vertical lines on the middle row showing the observed values of $f(x_1)$ (set arbitrarily here to 1.2), and the bottom rows showing the probability distribution $p(f(x_2)|f(x_1))$. In the left most column, the observation of $f(x_1)$ has no effect on the predicted value of $f(x_2)$: the prediction is centred around zero with a wide variance. For the middle column, the predicted value of $f(x_2)$ moves slightly toward the observed value of $f(x_1)$, and in the right column, it is moved even further.

Extending the number of unobserved variables to cover the whole of X results in a

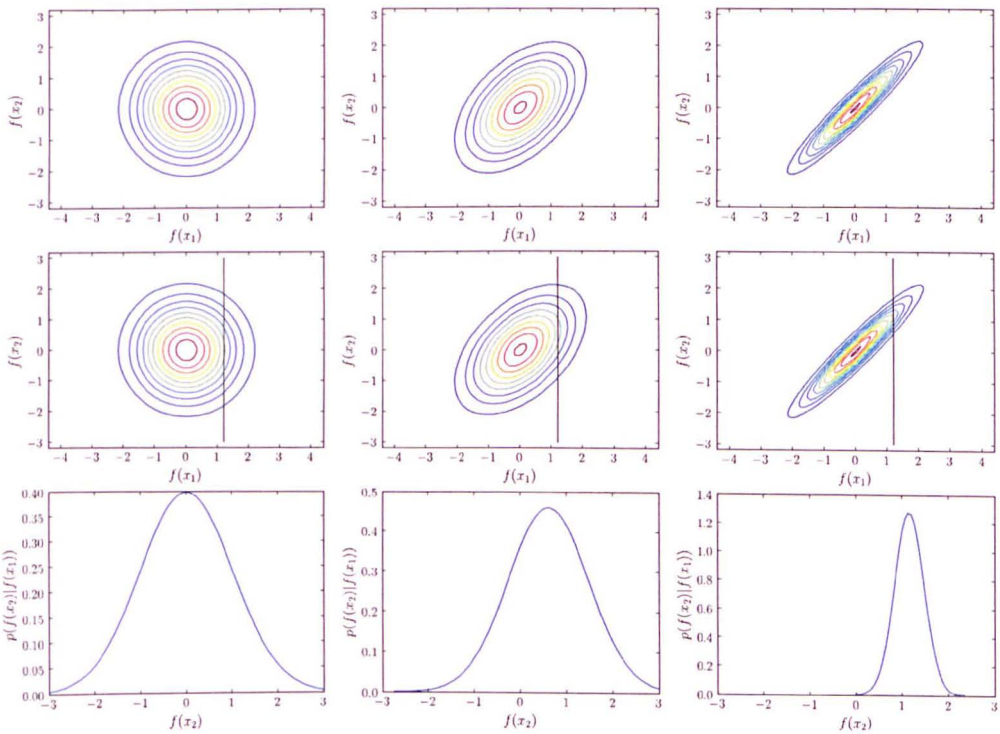


Figure 6.9: Illustration of jointly distributed variables for differing covariance values. A full discussion is given in the text.

distribution across *functions*, known as a Gaussian Process. Whilst this distribution is infinite-dimensional, in practise only the values of a limited number of points $f(X_*)$ are of interest. The prior and posterior distributions of the variables can be written in a more general form to encompass a large number of observed and test points:

$$p\left(\begin{bmatrix} \mathbf{f} \\ \mathbf{f}_* \end{bmatrix}\right) = \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K(X,X) & K(X,X_*) \\ K(X,X_*)^T & K(X_*,X_*) \end{bmatrix}\right) \tag{6.10}$$

$$p\left(\mathbf{f}_*|X_*,X,\mathbf{f}\right) = \mathcal{N}\left(K(X_*,X)K(X,X)^{-1}\mathbf{f}, K(X_*,X_*) - K(X_*,X)K(X,X)^{-1}K(X,X_*)\right) \tag{6.11}$$

where $K(X,X)$ is a matrix where the i,jth element is given by $k(x_i,x_j)$. The star notation (following [108]) denotes test points.

The limiting factor on the performance of a GP is the computational expense of inverting the matrix $K(X,X)$, which is a $n \times n$ matrix, where n is the number of observed data. Much recent work in the field of GPs has considered methods for sparsification of the result, see e.g. [110], [111].

6.6.3 Learning Methodology

It is desired to use a GP to map a vector of ΔT for each pair of sensors to (x,y), a position on the structure. In some (most) real world cases, an AE event will only activate a small number of sensors, whilst in the controlled conditions of the laboratory (where the training data was collected) all sensors were activated. This is due to the higher thresholds required during real tests (due to higher levels of background noise) and system saturation. The vector of ΔT information will therefore be incomplete for many of the test cases.

One approach could be to treat the prediction of a position using only a limited amount of ΔT information as a missing data problem (see Smola et al. [112] for discussion of how to treat missing data in machine learning). In this implementation however, it was decided to create a GP for every conceivable group of sensors. This is preferred for two reasons: 1) we are not concerned with the values of the missing data, and 2) for large tests the data is not so much missing as simply not present. For large tests with many sensors, acoustic emissions will occur a large distance from some of the sensors, such as to be undetectable, and so attempting to use

these distant sensors to perform a location makes no sense.

We define therefore a series of groups (or 'arrays') of sensors. In the case of the small geometry of the plate described above, this was every set of three or more sensors; one Gaussian Process is then created for each group. Upon the arrival of a new vector of ΔT information, there may be one or more GPs which can be used to make a prediction for (x, y).

Combining Predictions

The posterior distribution of each of the Gaussian Processes is a Gaussian with centre μ and variance Σ. In many circumstances, there will be many arrays which are capable of locating an event, and hence there will be multiple predictive distributions. In order to create a single distribution from these distributions we take a Product of Experts approach [113], whereby the distributions are multiplied together and re-normalised.

Since each of the distributions is Gaussian, the product of them is also Gaussian, with mean:

$$\hat{\mu} = \frac{\sum_i \frac{\mu_i}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2}} \tag{6.12}$$

Since the product of many Gaussians is not normalised (i.e does not integrate to 1), a simple method for creating a sensible covariance for the combined distribution is to take the mean of the variance for every individual distributions:

$$\hat{\sigma} = \frac{1}{n} \sum_i \sigma_i \tag{6.13}$$

Hyperparameter selection

In training the series of GPs, one must select some *hyperparameters*. These are the noise model σ_n and kernel length-scale l in equation 6.8.

The Bayesian setting in which Gaussian Processes are formed provides a natural way of selecting the hyperparameters through marginal likelihood [48]. In the case of a series of GPs however, it is unclear as to how one might implement this: should one maximise the *average* marginal likelihood for all the GPs? What about differ-

entiating between those acting on different dimensional data (i.e arrays of 3 sensors vs arrays of 6 or 7)?

Considering these intricacies and the abundance of data, it was decided that cross validation procedures would suffice for the selection of hyperparameters. For simplicity, all arrays of the same size were considered to have the same hyperparameters. A grid-search across the range of suitable values was performed in order to select hyperparameter values.

6.7 Results

The data was split into training, validation and testing sets. The hyperparameters which minimised the mean-squared-error (MSE) on the validation set were selected to train a final set of GPs to use on the test set. The result of this is shown in Figure 6.10; red crosses represent the known positions of the sources, and solid circles show the positions estimated by the algorithm. The colours of the solid circles indicate the average variance of each array. The MSE of the test set was 3.63mm.

In order to emphasise the benefits of the machine learning approach, as opposed to the trigonometric approach, Figure 6.11 shows the location of 100 randomly selected testing events located using the latter. Many of the test cases were located far from the plate (for reasons already discussed), and these are marked with arrows pointing towards this solution. Figure 6.11 can be compared directly to Figure 6.10, although the samples are not identical (due to a different random number seed).

6.8 Cross-Plate generalisation

The comparatively high cost of producing a GP locator (in generating training data and actual training) has been justified on the basis that the locator will be able to generalise to nominally identical structures without further training. It therefore remains to establish how well the algorithm does in fact perform when trained on data from one plate, and is required to locate a source on another (nominally identical) plate. This is possible here as several plates were investigated, differing only in respect to manufacturing tolerances. A complete matrix of results was produced (table 6.1) by training a GP with data from one plate and testing with each

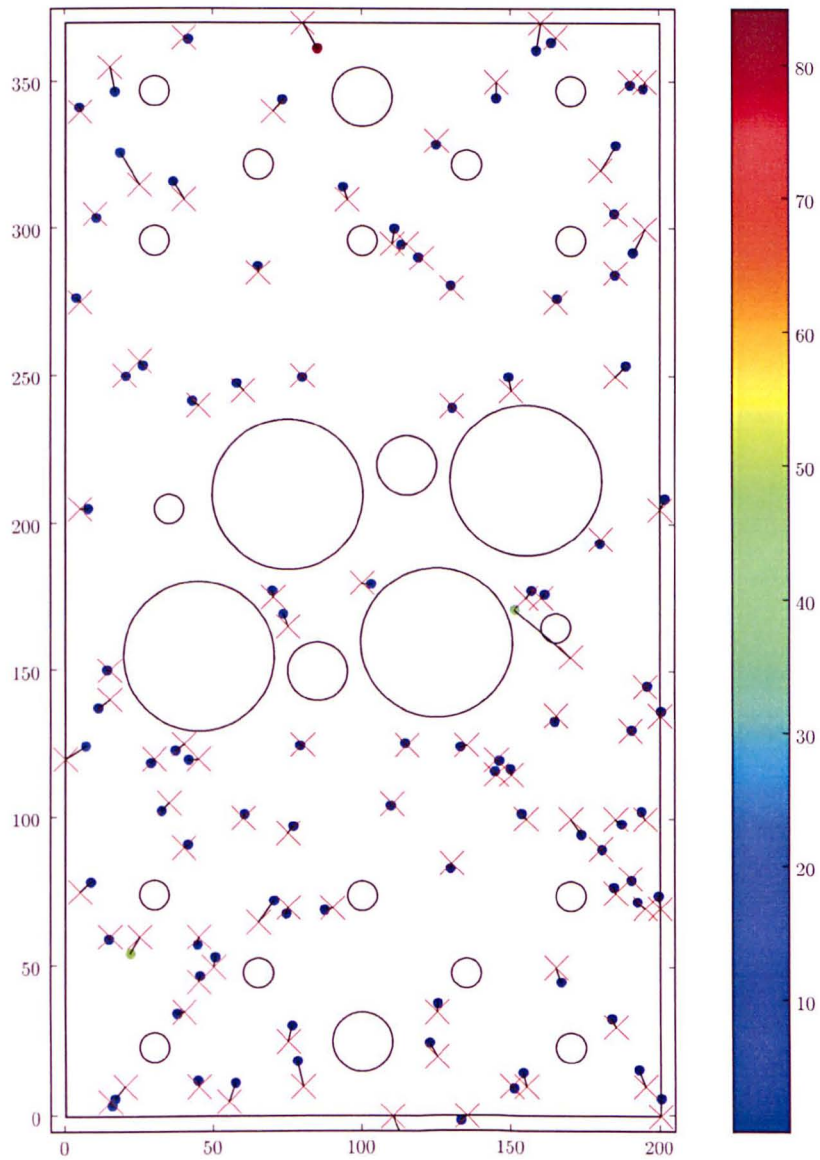


Figure 6.10: Results for the 100 selected testing events. Red crosses show the known positions of the events, whilst solid dots show the positions as estimated by the algorithm. The colour of the dots represents the posterior variance - a measure of our confidence of the location

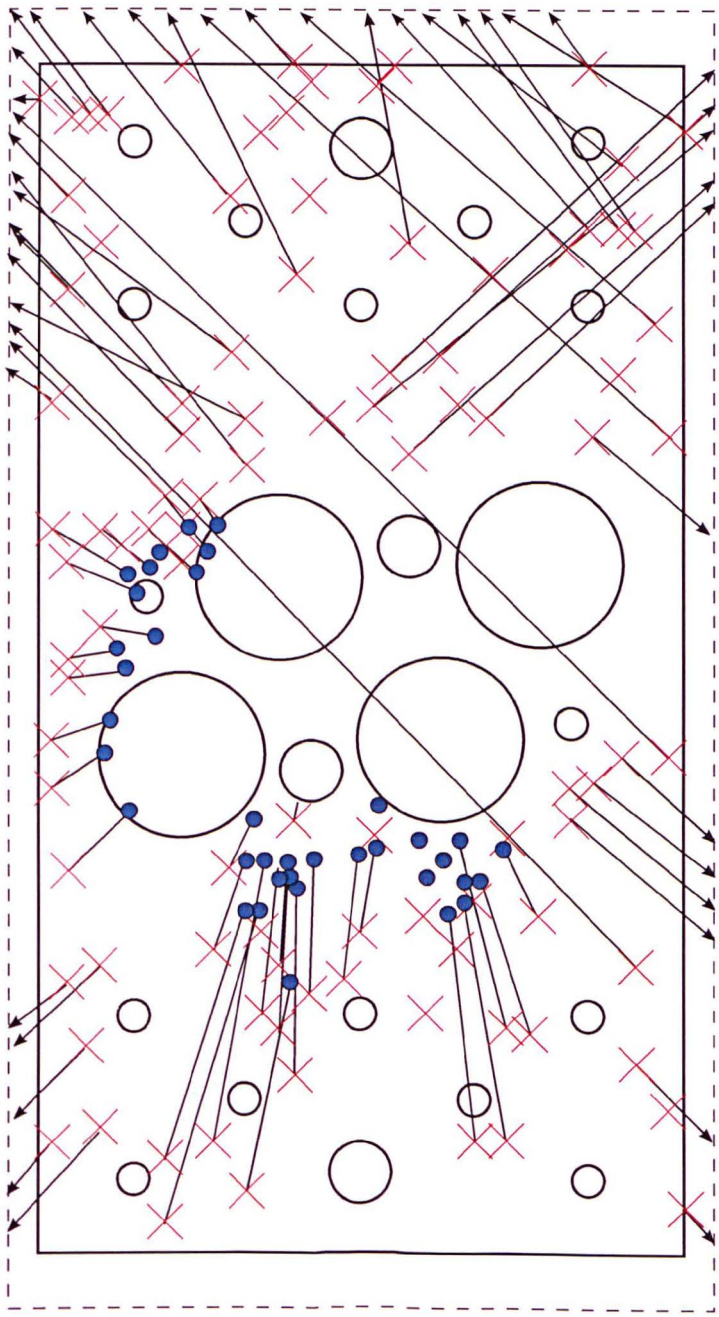


Figure 6.11: Results for 100 selected testing events using the trigonometric approach. Red crosses show the known positions of the events, whilst solid dots show the positions as estimated by the algorithm. Events which located a long way from the plate are represented by arrows pointing in the appropriate direction.

Table 6.1: Table showing the Root Mean Squared Error for the GP location algorithm trained and tested on combinations of plates.

		Training Plate			
		Plate A	Plate B	Plate C	Plate D
Testing Plate	A	3.63	6.77	8.50	5.37
	B	7.42	5.60	7.44	5.45
	C	6.03	6.17	4.56	5.71
	D	7.04	6.48	8.09	5.40

of the others. The results were uniformly excellent with the highest generalisation error being 8.5mm. The diagonals of Table 6.1 show the location errors when the GPs were tested on their training plates, and these have a maximum of 5.6mm.

6.9 Landing Gear Data

The machine learning location technique was employed during testing of a main landing gear, as detailed in section 7.2. Figure 6.12 compares the performance of the GP location method (with array selection) with the simple location method described in section 6.1. The test specimen is the main landing gear fitting of an Airbus A320, and is cylindrical, with step thickness changes, lugs, arms and other irregularities. In Figure 6.12 the horizontal axis is along the length of the fitting, and the vertical axis is circumferentially around the fitting.

A GP was trained for each possible array of the six used sensors, using data collected on a 50mm spaced grid using Hsu-Nielsen sources. The GP did not require all the data, in fact, for computational reasons, only 1000 pieces of data were used for each array, randomly selected from the pool.

The GP method and the trigonometric method were compared on fifty data randomly selected from the pool before GP training. The GP method shows a clear advantage over the trigonometric method, having a smaller error in all cases, with a RMSE of 19mm, compared to the RMSE of 113mm for the trigonometric case.

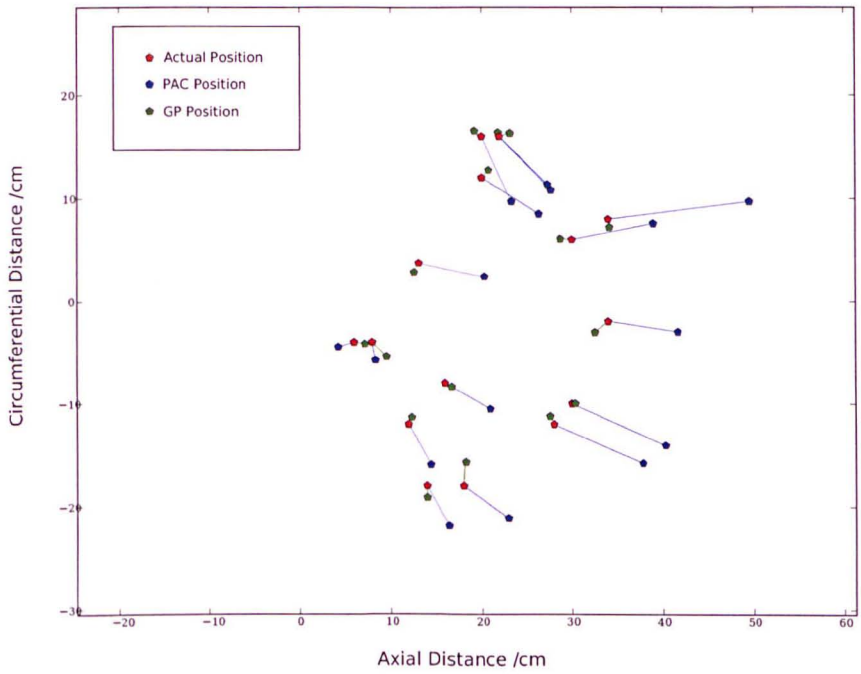


Figure 6.12: Performance of GP location algorithm, showing 15 typical results from the 50 test data used.

6.10 Summary

This chapter has discussed a GP methodology for locating Acoustic Emissions in complex structures. Extensive comparisons have been made with trigonometric methods, and the GP process has been shown to perform well. The GP process is also superior to other proposed machine learning schemes for AE source location, requiring less data to train, and producing a probabilistic output. This chapter completes the jigsaw of pieces required to implement the methodology proposed in Chapter 3. The next chapter demonstrates the successful implementation of the ideas proposed herein.

Chapter 7

Test Cases

This Chapter describes the application of many of the methods described in this thesis to real AE data. Two tests are described: in the first, an artificial source is used to represent a fatigue fracture during the testing of a major landing gear component; in the second, the main tube of a landing gear was tested to partial destruction, and an AE system was employed to monitor the event.

The first test is of particular significance because the landing gear was under load at the Messier-Dowty test facilities in Gloucestershire, where it was undergoing routine long-term testing. The positive results of the test demonstrate the applicability of the AE system to this test.

The second test was performed in the laboratories at Cardiff University. Measures were taken in order to reproduce the long-term test environment as experienced in the first test, whilst selectively fatiguing part of the gear to destruction. The success of the AE system here clearly demonstrates the usefulness of an AE system in detecting a real growing fracture in a landing gear component.

7.1 Artificial AE in Bogie Beam Test

This test case is provided in order to demonstrate the applicability of the methodology proposed in this thesis to a full scale fatigue environment. An artificial source was applied to the *bogie beam* section of an A340 landing gear during full scale

testing as described in section 1.3.1.

7.1.1 Methodology

The methodology for the AE identification strategy evaluated here is carried-out in three main phases, and so this section of the chapter is divided into three parts: The first part details the methods used to ensure that the artificial AE source was representative of a fatigue fracture, this is followed by details of the fatigue test machinery and AE recording equipment and the third section describes the signal processing techniques used to deal with the data.

Design of the Input Signal

The detection of an artificial AE source amongst 'everyday' AE activity would be relatively straightforward if the artificial source were not in some way representative of a fracture. If the artificial source was at a much higher amplitude than the other sources, or was exited in a different frequency band, the task of identifying it would be trivial. In order to make the task of identifying the artificial source non-trivial and representative of reality, the signal input to the transmitting transducer was carefully selected to faithfully represent a fatigue fracture.

The transmitting transducer chosen was a Physical Acoustic Pico, with a resonant frequency of 500kHz, and operating frequency range of 200-750kHz. This transducer was selected due largely to its convenient small size - the transducer was to be placed in awkward positions on the landing gear. The transducer was attached to a PAL ARB-1410 Arbitrary Waveform Generator board, which periodically excited the Pico transducer with a pre-defined signal, generating artificial acoustic emissions within the structure.

To ensure that identification task was non-trivial and that it faithfully represented the detection of a fatigue fracture event, preliminary tests were carried out on a small compact tension (CT) coupon of the same material as the landing gear. One CT coupon was fatigue tested to failure, and a typical waveform relating to crack growth was identified. The selected transmitting transducer was placed on a second CT coupon, and the settings of the signal generator were adjusted until the fatigue fracture signal was suitably simulated. Several signal types were tried including

	Fracture Channel 1	Fracture Channel 2	Artificial Channel 1	Artificial Channel 2
Amplitude (dB_{AE})	71	70	67	67
Rise time (μs)	270	290	240	202
Energy (aJ)	3.9e6	3.1e6	3.5e6	7.1e6
Duration (μs)	1000	1900	1990	1880
Init. Freq. (kHz)	31	42	230	181

Table 7.1: Comparison of signals identified during fatigue testing and stimulation of an artificial AE source.

sine, square and saw-tooth waves; the frequencies, amplitudes and burst-count were adjusted by trial and error until the received signals were a reasonable match to those captured during crack growth. Table 7.1 shows the characteristics of the signals which were matched, including the features from a typical fracture waveform, and those from the designed input.

The signal selected to excite the transducer was a two-cycle square wave, amplitude $0.5V$, period $2.5\mu s$

Test Details

The component designated for inspection was the *bogie beam* section of the landing gear from an Airbus A340. This section of the gear attaches the main tube to the wheels; it is approximately 2.4m long and 0.3m in diameter. The full gear was under test, designed to simulate the working conditions of the gear during take-off and landing. This involved the actuation of 16 hydraulic devices, activated in cycles so as to represent taxiing, take-of, flight and the landing stages of an aircraft's working life. One cycle (representing one flight) is completed in approximately 180 seconds, and contains a series of 'transitions' of frequency 0.3-2 Hz. The bogie beam, with AE sensors attached is shown in the photographs in Figure 7.1.

A Physical Acoustics DISP 20 channel system was used to record AE activity during the test. The sensors were distributed according to Figure 7.2. The sensors used were Physical Acoustics resonant *Nano 30* AE sensors, which have a bandwidth of 125 - 750 kHz, and a resonance frequency of 300 kHz. Each sensor was connected to an in-line pre-amplifier of gain 40dB, which also contained anti-aliasing filters. The sample rate used to record the AE signals was 40MHz.

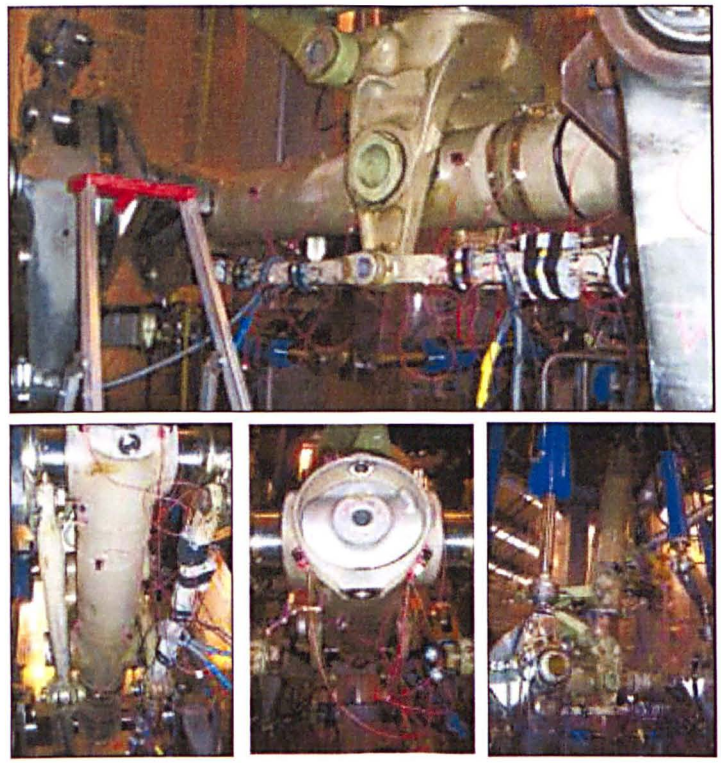


Figure 7.1: Photographs of the test setup. The bogie beam component is visible in the top photograph; subsequent photographs show the AE sensors attached to the beam by magnetic clamps.

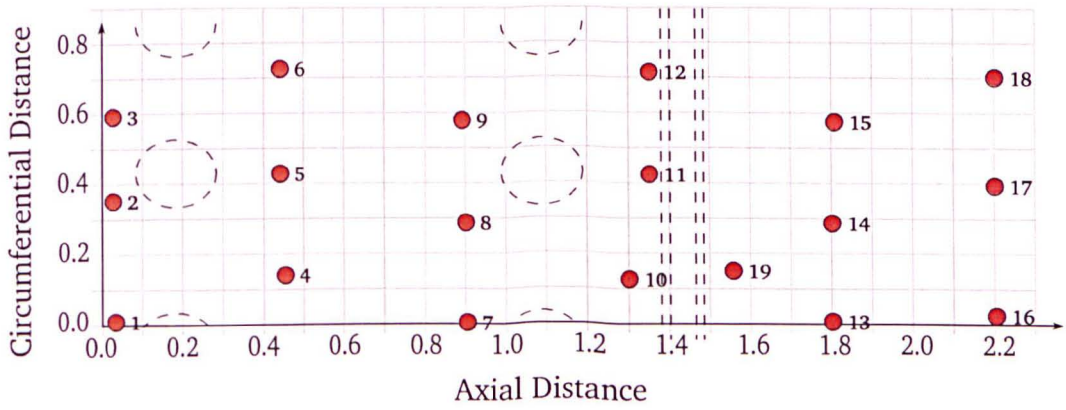


Figure 7.2: Schematic showing the setup of the AE equipment. The 19 sensors are shown (filled circles), placed on a grid along the bogie beam structure. Structural elements of the beam are identified as dashed lines - vertical lines show a thickness change and dashed circles show connections to other parts.

A total of three tests were performed over a period of three days; the duration of each test ranged from 1.5 to 3.5 hours, corresponding to approximately 30-70 load cycles. At the beginning of each test, the artificial AE source was moved to a new position by Messier-Dowty technical staff. These three tests are henceforth denoted Experiment A, B and C. The technicians did not reveal the positions of the sources until the analysis had been carried out.

For the duration of each test, the 'burst rate' of the artificial AE source was gradually increased. This is in line with the behaviour of a true fatigue fracture source (see section 3.4). The rate of acoustic emissions can be related to the crack length by an equation similar to the standard Paris law for crack propagation (7.1),

$$\frac{d\eta}{dN} = A\left(\frac{da}{dN}\right)^{\alpha} \tag{7.1}$$

where η is the acoustic emission count rate, N is the number of load cycles, a is the crack length, and A and α are constants whose exact values are not important here.

In order to properly correspond to the behaviour of a growing fatigue crack, the rate of input of the actuation signals was gradually increased over the duration of the test. This information was not used in the detection or identification of the source, but was used later to verify results and in particular, to explain an anomaly which occurred.

Signal Processing

The methodology described in Chapter 3 was implemented using the Python scripting language and scientific python modules SciPy [114]. The AE capture software *'AE Win'* was used to perform basic feature extraction and location of the data.

The recorded AE events were located using the simple assumption of a plate as described in section 6.1, and then clustered using the ORACAL algorithm described in section 3.2.4. The clusters were then ranked according to the feature space criterion as described in section 3.3.

Since the energy rate of the artificial source was controlled by the operators, the related energy-rate criterion (section 3.4) was not used in detecting the source. The energy rate of the cluster was compared to the energy rate of the input simply to

confirm the correct identification of the source.

7.1.2 Results

Three experiments were performed, each one starting with the removal and replacement of the transmitting device to a different position on the structure. The goal of each experiment was to positively identify the position of the source amongst the AE data.

For each test the cluster diameter required by the ORACAL algorithm was set to 10mm. This generated hundreds of clusters for each test, nonetheless the tightness parameter (see equation 3.1) proved to be a strong factor in selecting the cluster relating to the artificial source.

Experiment A

Figure 7.3 shows the positions of all of the significant clusters (those with more than 50 events) detected during the first test. The different symbols represent the different rankings of the clusters - those with the highest ranking (containing the most similar signals) are shown as circles. The highest ranked cluster is shown with a vertical cross, at position (1.48,0.62). The true position of the source is marked with a diagonal cross at (1.2,0.54), some 300 millimetres away.

In order to determine whether this discrepancy was due to the top ranking being obtained by the correct cluster (but with an error in the location) or whether it was due to identification of the incorrect cluster (i.e. the top ranking cluster pertains to some other event which is not the artificial source), the rate of the event in this cluster was plotted against the known rate of injection of the artificial signals (Figure 7.4).

Since Figure 7.4 clearly shows that the rate of acoustic emission events detected in the top-ranked cluster does not correspond with the rate of events input at the artificial AE source, it was assumed that the top-ranked cluster in this case does not represent the artificial source. This discrepancy is further investigated in section 7.1.2.

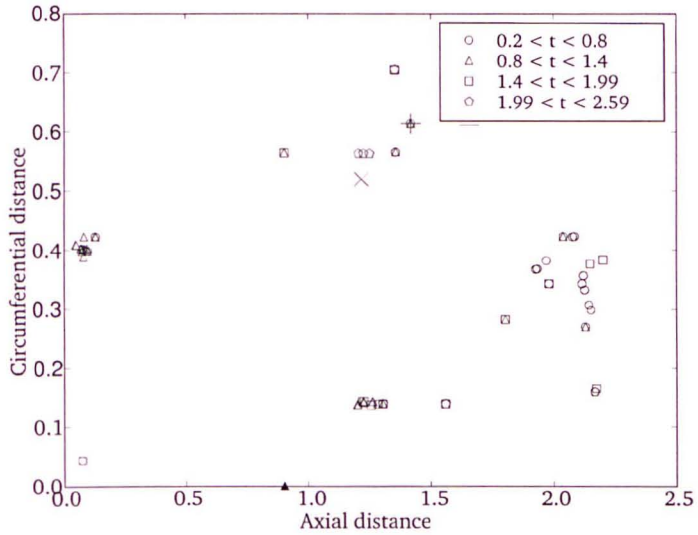


Figure 7.3: Scatter plot of the clusters detected during experiment A; the symbols show the relative 'tightness' of each cluster. The tightest cluster is shown at position (1.48,0.62) by a vertical cross, and the true position of the artificial source is show at position (1.2,0.54) by a diagonal cross

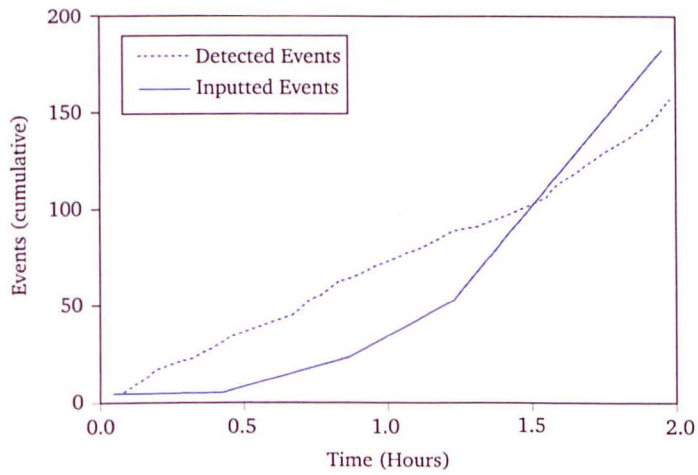


Figure 7.4: Event rate for experiment A, showing the rate of events detected for the most prominent (i.e. 'tightest') cluster. Also shown is the injected event rate: The two do not correlate well at all - it would seem that the cluster does not represent the input source.

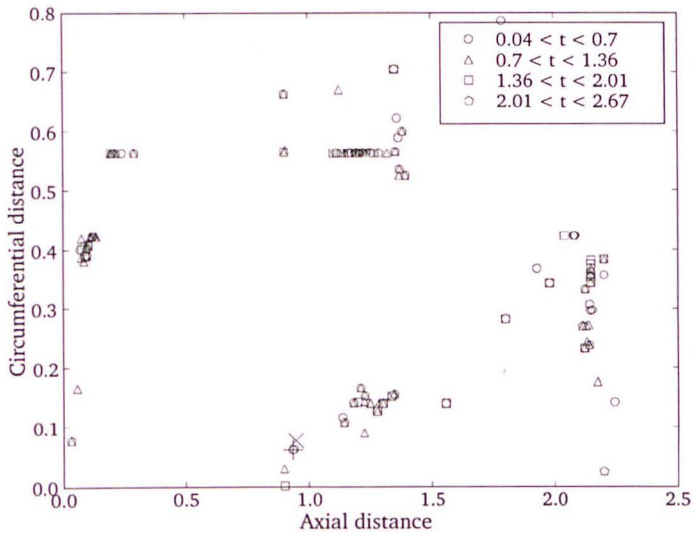


Figure 7.5: Scatter plot of the clusters detected during experiment B; the symbols show the relative 'tightness' of each cluster. The tightest cluster is shown at position (0.93,0.087), and the true position of the artificial source is shown at position (0.95,0.09)

Experiment B

Figure 7.5 shows a scatter plot of the significant clusters found during the second test. The primary cluster is marked with a vertical cross at (0.93,0.087); the true position (diagonal cross) of the artificial AE source was recorded to be (0.95,0.09), 30mm away. It can be noted that the top ranked cluster had a tightness score of 0.04 - this is an order of magnitude smaller than the top ranked cluster in experiment A, which did not represent the artificial source. This small value and the verification of the position of the source gives some degree of confidence that the detected cluster was the correct one.

In order to verify this, the event rate for the cluster was plotted against the input signal rate in Figure 7.6. There is a clear correspondence between the two rates, though the detected rate is slightly lower than the known input rate. This is likely due to the small cluster diameter used - using a broader cluster diameter would ensnare signals from a slightly wider area, and would likely correct this minor discrepancy. However, expanding the cluster size would open up the possibility of engulfing events in the cluster which did not relate to the artificial source (making the detected event rate higher than the true one).

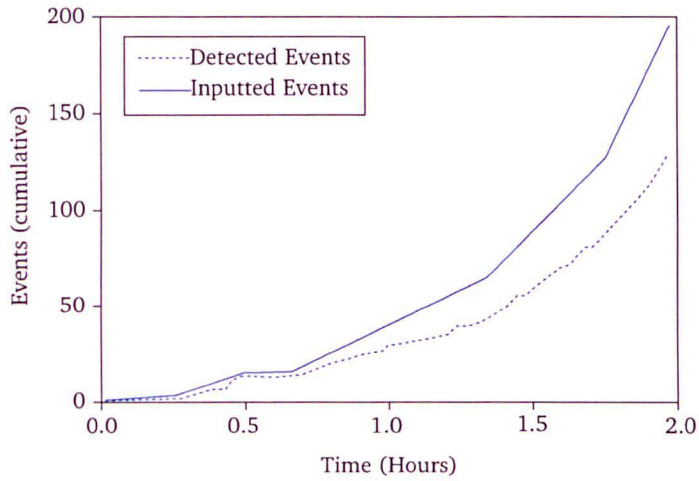


Figure 7.6: Event rate plot for experiment B, showing the input event rate and the event rate seen in the tightest cluster. The two rates show good correlation, indicating that this cluster does represent data from the artificial source.

The good correlation of the event rates along with the very small tightness value and correct identification of the position are good indicators that the artificial source was correctly identified in this test.

Experiment C

The artificial source was moved again and a third set of data were recorded. The tightest, top-ranked cluster was located at (1.97,0.34), which is 30mm from the known position of the source at (1.95,0.34). Again, the tightness score for the top cluster is relatively small – 0.16, giving a high degree of confidence that the artificial source is represented. Verification comes from Figure 7.8, where the rate of input artificial AE events is compared to the rate of detected events; an almost perfect correlation is observed.

The near-perfect correlation shown in Figure 7.8 gives a high degree of confidence that the artificial source was detected. Further, it would seem that these signals were detected with only small loss (96% of transmitted events were detected), and with no confusion with other signals.

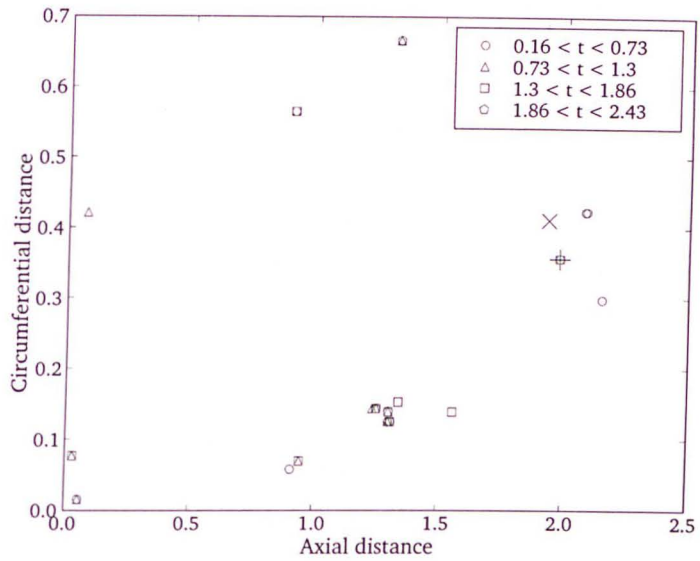


Figure 7.7: Scatter plot of the clusters detected during experiment C; the symbols show the relative 'tightness' of each cluster. The tightest cluster is shown at position (1.97,0.34) as a vertical cross, and the true position of the artificial source is shown as a diagonal cross at position (1.95,0.42)

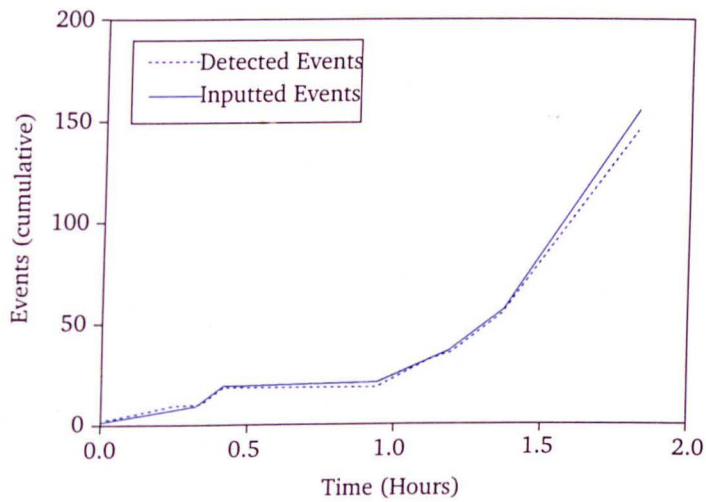


Figure 7.8: Event rate plot for experiment C, showing the input event rate and the event rate seen in the tightest cluster. The two rates show good correlation, indicating that this cluster does represent data from the artificial source.

Revisiting the data from Experiment A

Figures 7.6 and 7.8 give a high level of confidence that the artificial source was correctly detected in two of the three experiments conducted; in these cases the error in location (i.e. the distance between the known location and the one obtained from the algorithm) was small. Experiment A, on the other hand, seems to have failed; the error in location was larger (300mm) and Figure 7.4 does not show the expected correlation between the input and detected event rates. Further work was therefore conducted to investigate the cause of this.

The first consideration was that the ranking system was not effective for experiment A. However, a manual investigation of the clusters which were ranked lower than the top one did not reveal any clusters with event rates corresponding to the known input rate.

Since the known position (1.2, 0.54) was in the proximity of sensors 10,11 and 12 (see Figure 7.2) all the data arriving at these sensors was located linearly, around the circumference of the gear. A two dimensional plot was produced (Figure 7.9) showing this linear location plotted against amplitude. The number of events around at any discretised point is indicated by colour. By matching absolute times of the events, the data corresponding to the top-ranked cluster in experiment A was identified (indicated on figure 7.9 as 'detected cluster'). A nearby cluster, of lower amplitude was also evident (labelled 'potential cluster'). Plotting the event rate of this latter data against the known input event rate gave a very good correlation and this is shown in Figure 7.10.

It can be concluded then that the input signals were of lower than expected amplitude. They were therefore not detected by the minimum of three transducers, and were discarded in the initial analysis before the clustering stage. The likely cause of this lower amplitude is incorrect mounting of the transmitting transducer.

Further investigation of the data is necessary: should the amplitude of the inputted source have been sufficiently high to have been detected by three sensors (and therefore located and clustered in the original analysis), would the resulting cluster have had a high ranking? Alternatively: do the signals detected on channels 10 and 11 fulfil the hypothesis that signals from the artificial source will be similar, more so than surrounding signals? To answer this question, graphical methods were used.

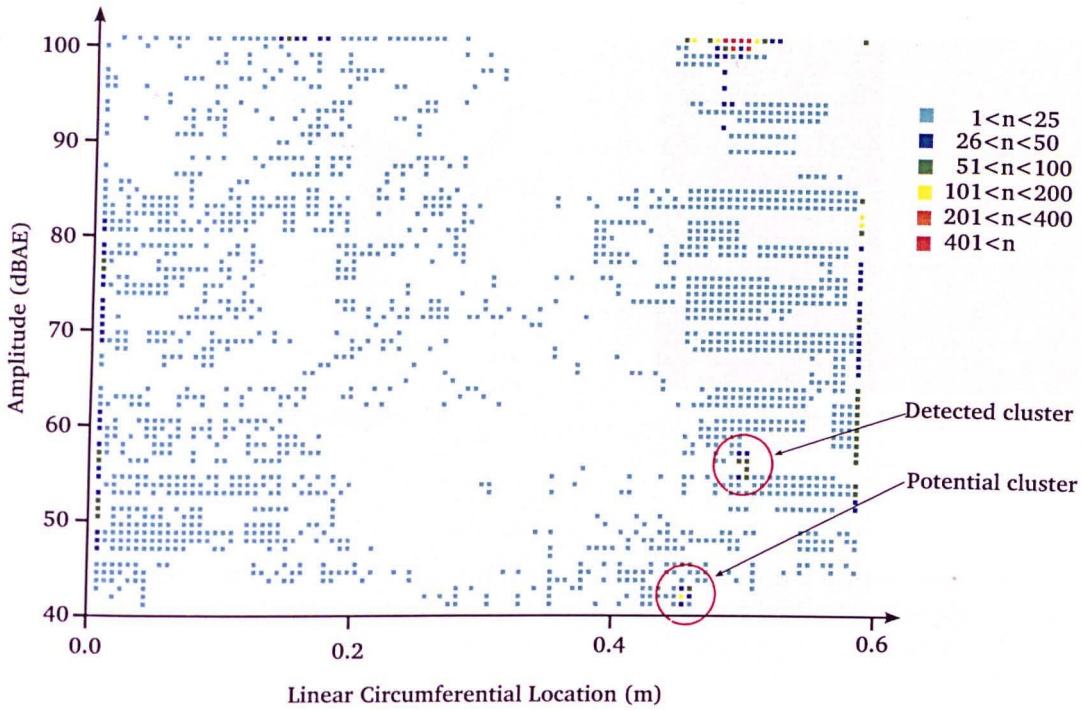


Figure 7.9: Linear location plot of the data from the region of sensors 10, 11 and 12. The vertical axis shows the amplitude of the events. The events relating to the detected cluster are highlighted, as a a series of lower amplitude events.

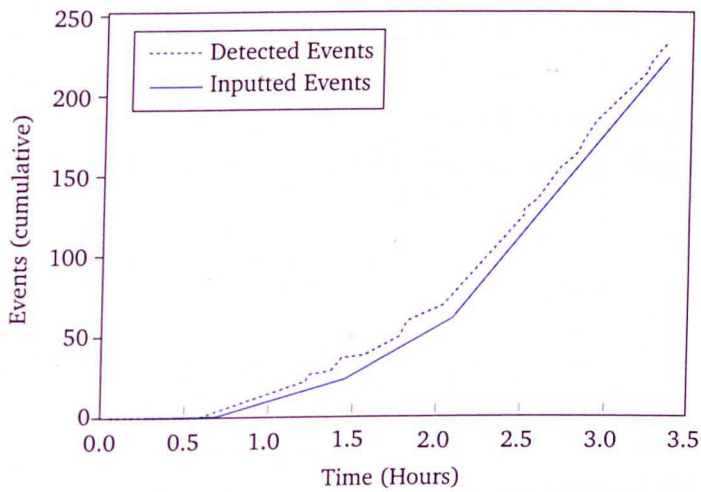


Figure 7.10: Input artificial AE rate (solid line) and detected AE rate (dashed line) for the group of events detected using linear location from the data in experiment A.

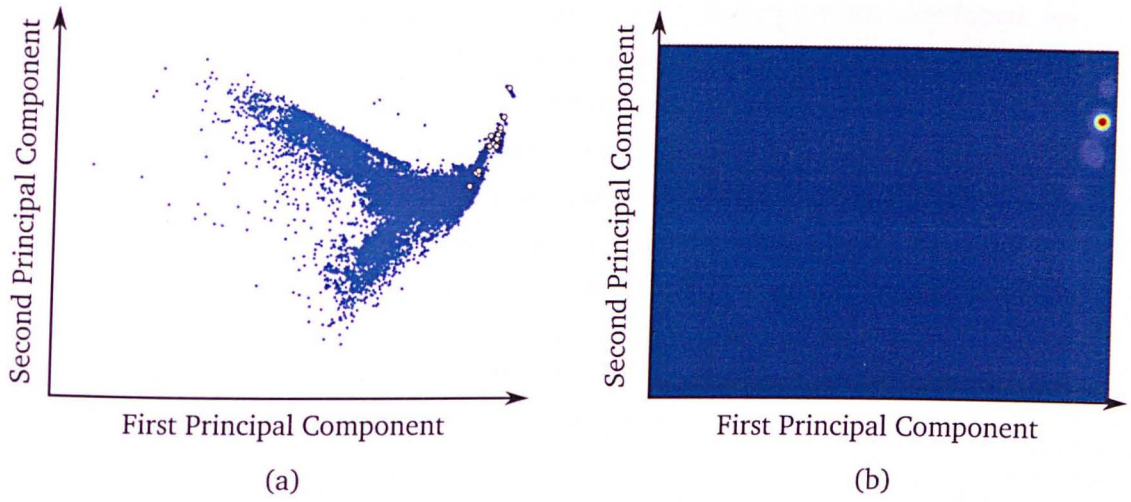


Figure 7.11: (a) PCA plot of the signal features detected on channels 10 and 11. Each dot corresponds to one signal – highlighted signals show the 'potential cluster'. (b) Kernel density estimation of the signals highlighted in (a)

Principal Component Analysis (PCA) is a useful tool for visualising high-dimensional data , see section 5.2. Figure 7.11(a) shows a PCA plot of all the feature data from the signals detected on channels 10 and 11. Each dot corresponds to (the features from) one signal. The signals identified in the secondary analysis as being a potential result (due to the match of the rates plot) are highlighted as white circles. Figure 7.11(b) shows the spatial density of these signals (on the same axes) as estimated by kernel probability density estimation [115]. Clearly most of the signals in this group are concentrated around one point, and it can be said with some certainty that these signals, if included in the original analysis, would have formed a tight, high-ranking cluster.

7.1.3 Discussion and Conclusions

This work shows clearly that the developed methodology allows fracture-like signals to be detected in the extremely noisy environment of a structural fatigue test. This is potentially very useful in the sense that it will allow suspension of the test in order to carry out an NDE inspection of the structure in order to evaluate the damage *in situ* before complete failure. The location information provided by the methodology means that the NDE inspection can be immediately directed to the area of interest.

The methodology differs in one important respect from previous developed approaches. The method does not attempt to recognise the individual burst waveforms characteristic of a crack-propagation event. Rather, the method exploits the fact that crack-related events (propagation at least - not crack-face rubbing) will be very similar to each other and this similarity can be detected by assessing the tightness of the event data when considered in the feature space. The major advantage of this approach is that it is expected to transfer between different materials without difficulty. In contrast, a system that has been carefully trained to recognise crack propagation in Aluminium, will probably not transfer successfully between Aluminium alloys, let alone between different metals.

7.2 Fatigue Testing of Main Landing Gear Component

A test was devised in order to further evaluate the algorithms presented in this thesis, in particular the methodology outlined in Chapter 3. An Airbus A340 main landing gear tube was placed in a fatigue rig with two actuators. The first actuator moved the sliding tube in and out of the main gear at a rate of 0.5 Hz, whilst the second applied a bending moment to a lug at a rate of 2Hz. The sliding tube presented no resistance and was effectively zero load, whilst the lug was loaded incrementally up to 11.1kN. The intention was to simulate the failure of the gear during certification testing, where a large number of signals are generated throughout the gear due to the various applied loads, and the challenge is to 'filter' these signals and identify the ones relating to fatigue activity. Due to hardware constraints, the data was recorded and post-processed off-line, though the software produced is capable of running faster than real-time.

The following sections describe the set-up of the test, the software implementation and the results of the testing.

7.2.1 Test setup

The test involved applying a large bending load to a lug on the landing gear, whilst simulating test conditions by applying a small load to the main tube. This proce-

dure was designed to create multiple sources of 'background AE', which need to be separated from the true fracture AE. The sliding tube was moved back and forth at a rate of 0.5 Hz, 40mm stroke, with very little load. The friction between the tube and the main gear structure created enormous quantities of background AE events (evident by the huge number of cluster shown in e.g. Figure 7.12). Simultaneously, a lug was loaded at a rate of 2Hz, initially with load 1kN. Load was gradually increased to 11.1 kN in order to grow a fatigue crack in a reasonable time-frame. Failure occurred at 630000 cycles.

7.2.2 Software Methodology

The methods described in this thesis, as outlined in Chapter 3 were implemented in software and used to post-process the collected data.

A grid was constructed across the surface of the gear, and Hsu-Nielsen sources were used to generate training data for a GP location algorithm (Chapter 6). This data was used to compare the performance of the algorithm to traditional methods on section 6.9. The GP location method was then used to locate AE events on the structure.

The ORACAL clustering method described in section 3.2.4 was used to cluster located events into clusters. The cluster diameter was set to $10mm$.

The feature-space-distribution clue and the energy-rates clue were implemented and used to rank the clusters, and this information was presented to the user for detailed analysis via a software interface.

7.2.3 Software Interface

A software implementation of the algorithms presented in this thesis was produced in the python programming language with the use of some open source libraries including pyGTK, numpy, scipy and matplotlib [114]. A typical screen-shot from the program is shown in Figure 7.12.

The user interface to the software is divided into two parts. On the left hand side, a scatter plot shows the positions of detected clusters. In a commercially realised version of the software, clusters would ideally be shown against the background of a

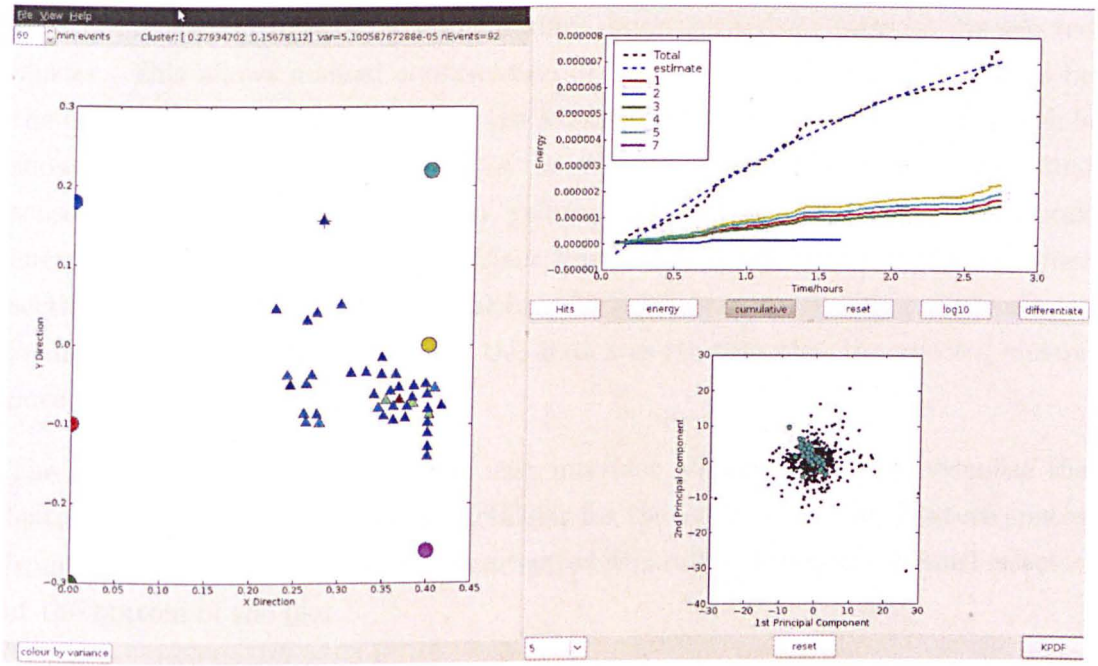


Figure 7.12: A typical screen-shot from the AE analysis software

schematic drawing of the component in question. In this prototype, the structure is represented by six of the sensors used in the test, marked by large coloured circles. Each detected cluster is marked with a coloured triangle.

The colour of the triangle is used to represent the level of a clue: the clues can be cycled between using the button in the bottom left corner (marked 'colour by variance' in Figure 7.12). Clues with high values are marked as red, clues with lower values are marked as green.

In this implementation of the software, the 'feature space distribution' and 'energy rate' clues were implemented. Novelty detection was left as a goal for further work, see section 8.2.1. A commercially realised version of the software would also include data fusion as discussed in section 3.6.1. This has also been assigned as future work, see section 8.2.2.

Clicking on a coloured triangle 'selects' the associated cluster of data. The selected cluster is marked with a vertical cross and its relevant information (position, number of events, level of clues) is displayed at the top of the graph. The selected cluster can also be explored in more detail in the two plots of the right-hand side of the user-interface.

The plot in the top-right of the user-interface shows the activity rates for the selected cluster. This allows manual confirmation of the energy-rates clue as indicated by the colour of the selected cluster on the left-hand plot. In Figure 7.12, the graph is showing the cumulative energy rate for the selected cluster. The rates for individual sensors are coloured red, blue, green, yellow, cyan and magenta, whilst the overall energy rate is shown by a dashed black line. The model of the energy rate (see section 3.4) is shown as a dotted blue line. In this instance, the trend of the cluster is almost linear, slightly downward: this indicates strongly that the selected cluster does not represent a fracture.

The bottom right-hand plot in the user interface allows the user to visualise the feature space for the test, and in particular for the selected cluster. Feature spaces from each channel in the test are investigated separately, using the channel selector at the bottom of the plot.

In this implementation, the simple PCA algorithm (see section 5.2) was used to create a 2D representation of the feature space, though in a commercial realisation of the project a more powerful algorithm (or selection of algorithms) from Chapter 5 could be used.

All signals detected during the test on the selected channel are represented by black dots. Those which are additionally within the selected cluster are highlighted. This allows the user to identify whether the selected cluster contains typical signals (near the centre of the plot) or unusual signals (near the edges of the plot) and whether the signals in the selected cluster are similar (close together on the plot) or widely differing (spread across the range of the data on the plot).

Due to the amount of data gathered in a typical AE test, it is difficult to asses the spread of the signals using the highlighted dots (as in Figure 7.12), and so a density estimation tool is included, which shows the user the density of the highlighted signals in the reduced feature space. This is demonstrated in Figure 7.13, where the bottom-right-hand plot is showing a density estimate of the signals highlighted in Figure 7.12. The Kernel Probability Density Estimation procedure was used to generate the plot.

This procedure allows the user to visually confirm the presence of the feature space distribution clue, as indicated by the colour of the cluster on the cluster on the left-hand side of the user interface.

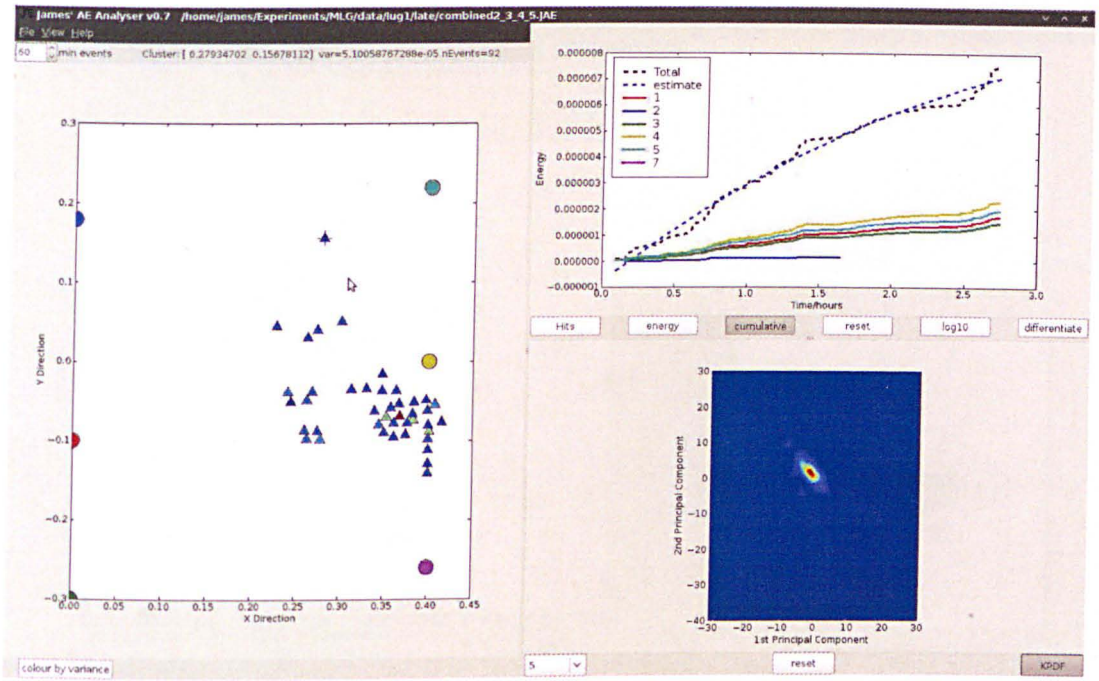


Figure 7.13: A typical screen-shot from the AE analysis software, showing the Kernel Probability Density Estimate of the reduced feature space.

7.2.4 Results

Figure 7.14 shows the AE analysis software at the point where fracture was detected. The clusters have been coloured according to the feature-space clue, and three clusters are particularly prominent (shown in red). Inspection of the energy rate for two of these clusters (near the top and bottom of the screen) showed that they were uninteresting: the energy rose linearly. Inspecting the energy trend for the remaining cluster showed that it did indeed represent a fracture under the 'energy trend' clue; in fact when colouring all of the clusters according to this clue, the selected cluster remains prominent (i.e. it has one of the strongest rising energy trends).

Although the features - based clue is not sufficient in this case to identify a fracture by itself, it is clearly a very strong indicator, and with the use of the additional energy-rates clue, the fracture was detected. This kind of 'manual' data fusion (where the engineer uses his judgement and common sense) would not be required with the use of data fusion, which will be implemented in future work (see section 8.2.2).

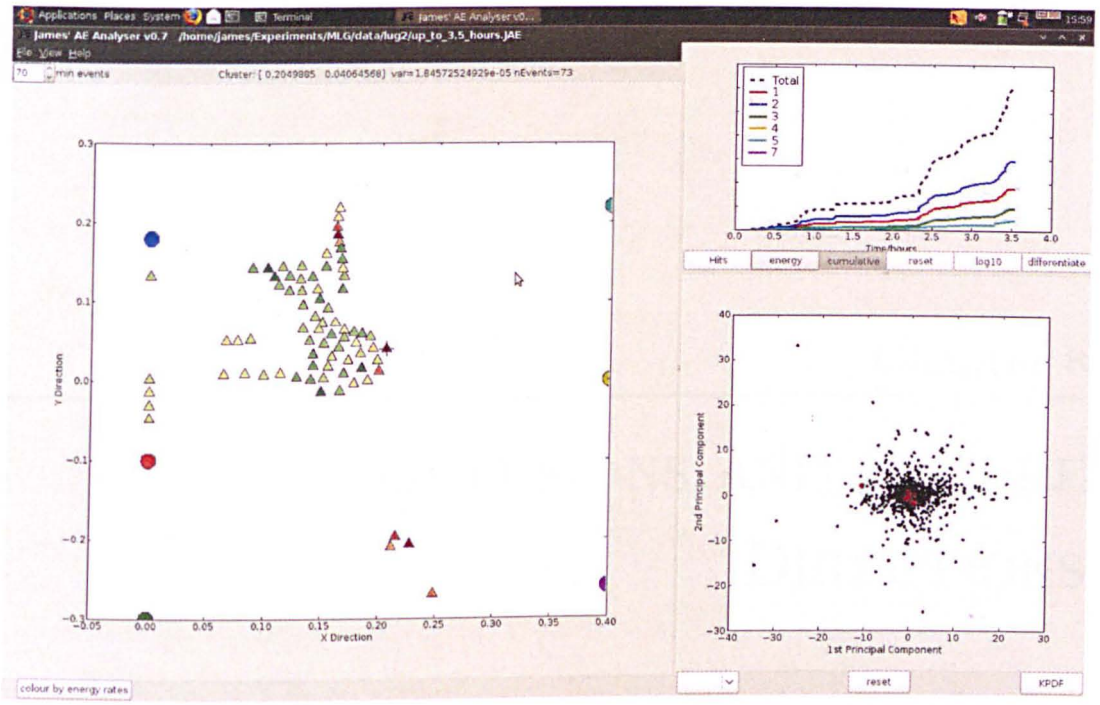


Figure 7.14: A Screenshot of the AE analysis program, with suspected fracture highlighted

The position of the detected cluster was (0.2, 0.04), which is a mere 140mm from the true failure position at approximately (0.3, -0.03). Whilst this is larger than experienced during the training of the GP location algorithm, it is still very acceptable given that the structure is 2m long!

The fracture was identified 3.5 hours into the testing. The component failed 4.5 hours later, after a total of 8 hours. This demonstrates that the AE procedure can indeed provide early warning of a potential fatigue crack failure.

Chapter 8

Conclusions and Future Directions

8.1 Conclusions

This thesis has presented a novel technique for detecting fatigue fractures in landing gear tests, which could be extended to other fatigue environments. Along with the issues discussed herein, there are a multitude of practical considerations regarding sensor placement, sensor selection, equipment setup etcetera. These issues are addressed in the methodology document produced during this project [116].

8.1.1 Aims and Objectives Revisited

In Chapter 1, some objectives were identified. They are listed again here with comments as to the extent to which they were achieved.

(i) **Study wave propagation in complex structures**
Wave propagation was studied in Chapter 2. The AE process was considered in the form of a 'forward model', from source to sensor. Different types of AE source were considered, different wave propagation modes were described. In particular, the effect of dispersion on AE measurement was studied. The conclusions arriving from this chapter were that wave propagation is very

difficult to model. Indeed, the author considered only Lamb waves arriving directly at a transducer in a simple plate, and did not consider boundary reflections or mode conversions which would complicate the understanding of AE signals further.

It was concluded that distinguishing fracture sources on a signal-to-signal basis would be extremely difficult, and two potential mechanisms were identified as potentially useful for understanding or identifying fatigue fractures: the variability of the source and the energy trend of the source.

(ii) **Overcome problems with AE location in complex structures**
In Chapter 6, a methodology for locating AE in complex structures was presented. This involved an onset-picking algorithm due to Kurz [102], which was studied, implemented and shown to be useful; as well as a Gaussian Process regression technique, using modern machine learning methods. The system was developed and extensively tested on a small but complicated component, and was shown to scale well to a large landing gear component. It was also demonstrated that the system was capable of generalising across nominally similar structures, i.e. training data collected in the lab could be used successfully in the field.

(iii) **Devise a system capable of handling large quantities of AE data**
The feature extraction methodologies presented in Chapter 4 were implemented for two reasons: firstly to allow comparison of signals in a vector space; and secondly in order to compress the data. Two feature extraction methodologies were presented, using continuous and fast wavelet transforms. It was concluded that the continuous transform based approach, whilst elegant, was of little practical use due to the amount of computational effort required. The fast wavelet transform was compared to the traditional signal features for a controlled laboratory test, where it was demonstrated that many different decomposition levels can adequately act as features. The method showed consistent but minor improvements over the traditional features.

In order to select a decomposition level to act as a suitable feature set *for fracture-related* data, one must strike a compromise between the amount of compression an the potential for information loss. In order to make a reliable decision in this regard, additional work is needed to extend the methods presented in this thesis to tests involving real fractures. This is planned (see 8.2.3).

(iv) **Employ novel signal processing techniques and pattern recognition**
In Chapter 6, the modern machine learning technique of Gaussian processes were successfully implemented to solve a problem pertaining to the location of AE signals. Wavelet transforms were used extensively in Chapter 4, and various signal processing and machine learning techniques appear throughout this work, included the k-nearest-neighbour algorithm and the particle swarm optimisation algorithm. A simple clustering algorithm was devised (after consulting the literature) in Chapter 3. This work makes successful use of a wide variety of signal processing and machine learning methodologies.

(v) **Devise a methodology of presenting AE data to the user in a simple form**
Chapter 5 reviewed various dimension reduction techniques. These varied in computational cost but all had the same goal: to find a latent 2D representation of AE data. The techniques were applied to a toy data set and to a a laboratory experiment. The PCA algorithm was applied to the identification of a real fracture in Chapter 7.

8.1.2 A Methodology for the Identification of Fractures

The ultimate goal of this project was to identify and develop method for distinguishing between fracture-related signals and non-fracture related signals during the fatigue testing of a landing gear component. Very early on in the thesis, it was realised that training a traditional classifier (such as a neural network) to identify such signals would be difficult for two main reasons:

(i) Collecting training data for such a classifier would be almost impossible: the 'sound' of a fracture in one component will appear different to that in another: indeed the sound of (the same) fracture will differ from sensor to sensor, depending on the sensor's placement.

(ii) Modelling the AE propagation (to account for signal differences due to sensor placement) is a challenging prospect, especially when one considers the modelling of mode conversion, reflection and uncertainty due to sensor coupling etc.

This, along with an understanding gained from studying the AE process and evidence obtained from AE tests, led to the development of the method presented in Chapter 3, where fractures are identified on the basis of *clusters* of signals. Three 'clues' which might help identify a fracture were identified. These clues are:

(i) Feature space distribution – fracture signals seem to be 'more similar' to each other than non-fracture signals are to each other

(ii) Energy trend – from the Paris law, we surmise that a cluster of fracture related signals will have a rising energy trend

(iii) Novelty – fracture related signals will only occur in the presence of a crack, other signals should remain constant

The effectiveness of the first two clues were assessed to a minor extent in Chapter 3, where some visual and numerical evidence was presented. During the testing phase in Chapter 7, it became clear that the most effective process was to first identify clusters with a strong $p(fracture|features)$, and then examine the energy rates for potential candidates. Indeed, the features based clue alone was sufficient to detect an artificial signal in a real environment.

The effectiveness of the clues (and hence information necessary for data fusion) should be rigorously tested in a controlled study, using fatigue data from various landing gear fatigue tests. This is beyond the resources of this project. Further, the framework is capable of accepting additional clues should they become available in future.

In order to make use of all three clues, a data fusion scheme was proposed, though the implementation of this remains an objective for the future (see 8.2.2).

8.2 Future Directions

8.2.1 Novelty detection

Section 3.5 proposed a spatial novelty as a clue in the data driven method. Time constrains have meant that this avenue of research has not been thoroughly explored, and it is left as an area of future work.

Part of the extended work being funded by Messier-Dowty includes thorough investigation of a spatial novelty detection clue. This will approached by fatigue testing small specimens (CT and four-point-bend tests) where active AE locations on the specimen can easily be attributed to either crack growth or the loading device.

One challenge in this approach is to test the specimens for a long period without fracture, in order to collect sufficient data to build a model of the 'normal' AE behaviour of the system. Ensuring that the specimen does eventually fracture under the same conditions involves selecting exactly the right loading rate and load.

Another challenge in this approach is to benchmark the method. It is impossible to asses the effectiveness of the novelty detection without knowing the 'true' point of onset of the crack. Inevitably, one is thrown the impossible task of finding a technique to benchmark the AE novelty against. Suggested methods have involved foil gauges attached to the side of the specimen and regular microscopic visual inspection.

SATSCAN

One area of research that has already developed some technologies regarding spatial novelties is epidemic detection. The SATSCAN project [117] has developed algorithms based on detecting novelty in spatial activities in order to detect the onset of disease outbreaks. The parallels with acoustic emission are uncanny. The methods can be surmised as follows:

Every time a packet of paracetamol is sold in the UK, the event is recorded along with its time and position. The rate at which paracetamol is sold obviously varies spatially, i.e. more packets are sold in more densely populated areas. The SATSCAN method works by clustering (spatially and temporally) abnormal sales of paracetamol. which identify outbreaks of influenza amongst the population.

The method has been extended to detecting various other diseases based on other spatial statistics, such as detecting West Nile disease by clustering the occurrences of dead birds [118]. It has also reached other fields including criminology [119].

Despite its clear parallels to acoustic emission monitoring, the method has not been applied to this field, and the author feels it would be well worth future investigation.

8.2.2 Data Fusion

Section 3.6.1 discussed briefly the idea of data fusion, and laid the groundwork for the fusion of the clues described in Chapter 3. In Chapter 7, when applying the methods in this thesis to real test situations, data fusion was not implemented. This is partly due to time constraints, and partly due to a lack of suitable data.

Fusing the clues onto one single value essentially involves collecting large amounts of test data and correlating the occurrences of clues with the occurances of fractures, in order to identify the 'trustworthiness' of each clue. One must be meticulous in labelling data as 'fracture' and 'non-fracture' in order to accomplish this.

The collection of such data is planned as part of the future work (see section 8.2.3), where data fusion will be attempted.

8.2.3 Further Fatigue Testing

A further phase of testing at Cardiff University has been agreed with Messier Dowty. The objectives of the testing are threefold:

(i) To develop the use if spatial novelty detection as a 'clue'.

(ii) To collect suitable data in order to learn parameters for fusing clues.

(iii) To demonstrate further the capabilities of the system on a series of fatigue tests.

The further development of novelty detection and data fusion have been discussed above. It is clear that thorough demonstration of the success of the system is necessary for further development and commercialisation.

The tests will take the form of a series of CT specimen tests and some four point bend tests in order to gather data for use in the development of the novelty detection clue and the data fusion. The tests will continue with a 'pseudo landing gear' structure, designed to replicate the challenge of locating fractures on a landing gear at a cost which allows the test to be repeated several times. Finally, further tests on will take place on real landing gear specimens.

8.2.4 Commercial Exploitation

The commercial exploitation of this work depends on the cooperation of various companies including Physical Acoustics, Messier-Dowty and of course The University of Sheffield. Negotiations have been underway for some time, and currently preparations are being made for funding the realisation of the work in this thesis in conjunction with work at Cardiff University into a product which can be exploited by Messier-Dowty.

Work is also underway to apply for a patent for some of the ideas described in this work, incorporated into a larger vision with work from Cardiff University.

Communications have begun between the author and software engineers at Physical Acoustics. This took the form of a two day meeting in June 2008. Further communications are planned.

Other Domains

This thesis has been focussed toward the application of AE monitoring to landing gear, however the work herein is applicable in a much wider context. The framework for detecting fatigue fractures is not domain-specific, and one could envisage a similar system being employed in other parts of the aerospace industry, as well as parts of the automotive, offshore and construction industries.

This extensibility stems from the way that fatigue fractures were defined in Chapter 3, which is *completely independent of the structure.* This is in contrast to other successful AE applications such as the BALRUE project [120] which rely on learning the 'sound' of a fatigue crack in a particular metal, under particular loading and propagating through a particular geometry. Such a system can then identify fatigue cracks on a hit-by-hit basis, but the work cannot be extended to other materials or structures.

APPENDIX

A.1 Elastic Constants

The Material constants E (Young's Modulus) and v (Poisson's Ratio) along with ρ (density) are sufficient information to calculate wave-speeds for a material. The Lamé constants used in the wave equations are also related. Lamé's first constant is given by

$$\lambda = \frac{Ev}{(1+v)(1-2v)}$$

The second Lamé constant is also known as the shear modulus and is given by

$$\mu = \frac{E}{2(1+v)}$$

The longitudinal wave (a.k.a. p-wave) speed is related to the Lamé constants by

$$c_p = \sqrt{\frac{\lambda + 2\mu}{\rho}}$$

The transverse wave (s-wave) speed is given by:

$$c_t = \sqrt{\frac{\mu}{\rho}}$$

A.2 Particle Swarm Optimisation

The particle swarm is a stochastic search algorithm, not dissimilar to genetic algorithms[121]. The algorithm is based on the idea of a swarm of creatures such as fish or bees, which communicate in a simple way in order to perform an effective search [71].

The advantage of the algorithm is that there are very few parameters which control the performance of the optimisation. These are

- **N**, the number of particles in the swarm.
- $\mathbf{C_1}$, the tendency of the swarm toward local optima.
- $\mathbf{C_2}$, the tendency of the swarm toward global optima.
- **LF**, (loss factor), a measure of the momentum of each particle.

Theory

The algorithm proceeds as follows. Each member of the swarm is initialised with a random starting position, as well as a random *velocity.*

The objective function is then evaluated at the position of each particle. The algorithm stores the best position achieved by each particle (the *local* best, p_{lb}), as well as the best position achieved by any particle (the *global* best, p_{gb}).

The velocity of each particle is then updated according to equation A.1

$$v_{new} = LFv_{old} + C_1R(p - p_{lb}) + C_2R(p - p_{gb}) \tag{A.1}$$

where R represents a random number in the range $(0, 1)$, and p is the current position of the particle.

The position of each particle is updated according to equation A.2

$$p_{new} = p_{old} + v_{new} \tag{A.2}$$

This procedure then iterates until either some convergence criterion is reached, or a pre-defined maximum number of iterations is met.

Example

Figure A.1 shows a particle swarm optimisation of a 'camel back' function with six local minima. Each sub-figure shows the position of the particles after zero, five, ten, fifteen, twenty and twenty-five iterations (reading across the page): particles are depicted as red dots. The camel-back function used is:

$$z(x,y) = 4x^2 - 2.1x^4 + \frac{x^6}{3} + xy - 4y^2 + 4y^4 \tag{A.3}$$

where z is the quantity to be minimised and x and y are the free variables. The particles are initialised at random across the area of interest (top left sub-plot), initially investigating the problem on a large scale. Towards the end of the optimisation (bottom right plot) the particles have condensed on a local solution and can be considered to be performing a 'local' search.

A.3 Hyperbolae

A hyperbola is a locus of two points such that the difference between the distance to two points (the *foci*) is constant - it can be thought of as an inside-out ellipse. Consider figure A.2: the foci are shown at positions $A = (x_A, y_A)$, $B = (x_B, y_B)$. In the acoustic emission application, these foci represent the positions of the sensors, and the locus represents possible locations of an acoustic emission. Since the difference between the distances from $P = (x, y)$, a point of the hyperbola to the foci is constant, we can write:

$$\overline{PA} - \overline{PB} = 2c \tag{A.4}$$

Examining figure A.2 and noting the two right-angle triangles with hypotenuses AP, BP, we can write the following two equations:

$$(x + f)^2 + y^2 = \overline{PA}^2 \tag{A.5}$$

$$(x - f)^2 + y^2 = \overline{PB}^2 \tag{A.6}$$

Where the foci A and B lie at $-f$ and $+f$. Substituting for (A.4),

$$\left((x - f)^2 + y^2\right)^{\frac{1}{2}} - \left((x + f)^2 + y^2\right)^{\frac{1}{2}} = 2c \tag{A.7}$$

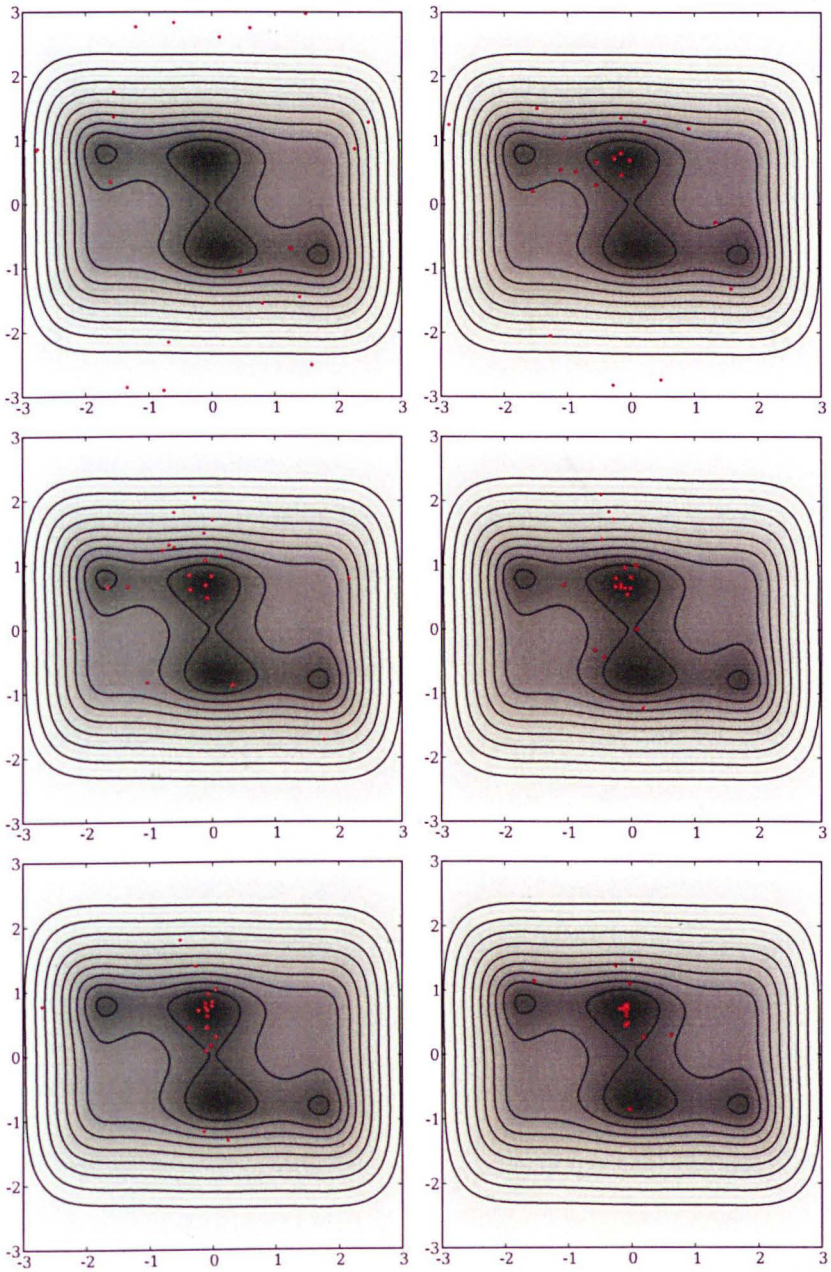


Figure A.1: Particle swarm optimisation of a 'camel back' function with six local minima. Each successive sub-figure shows the position of the particles (red dots) after 5 further iterations. Note that the contours are logarithmically spaced.

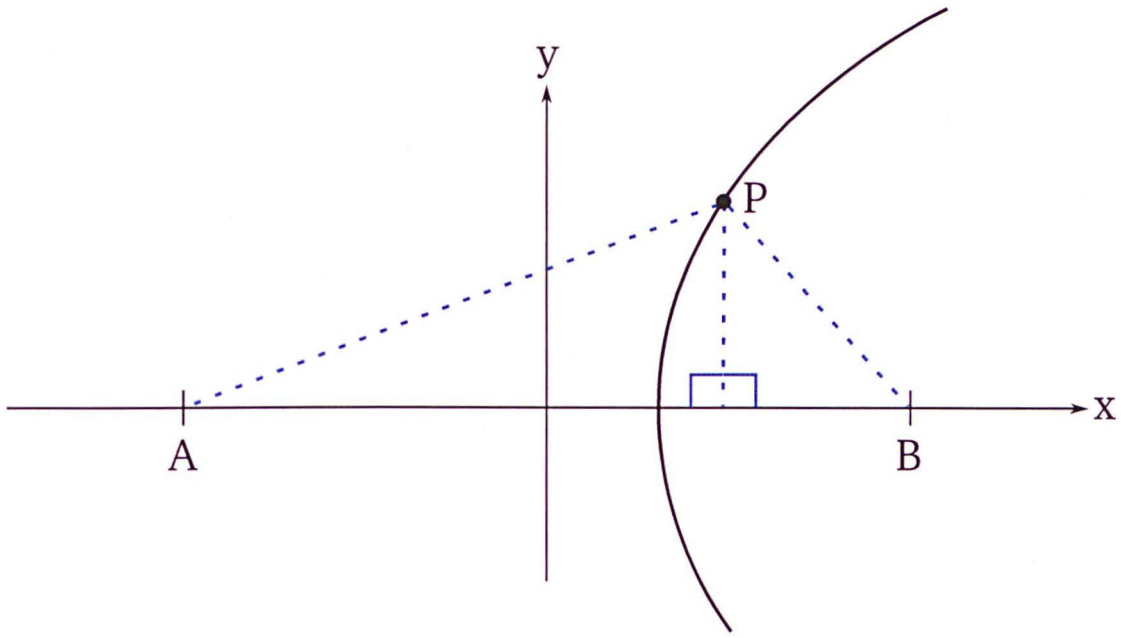


Figure A.2: A Hyperbola about two foci A and B on the x axis.

Re-arranging and completing the square gives

$$x^2(f^2 + c^2) - c^2y^2 = c^2(f^2 - c^2) \tag{A.8}$$

dividing through by the right hand side leaves

$$\frac{x^2}{c^2} - \frac{y^2}{(f^2 - c^2)} = 1 \tag{A.9}$$

BIBLIOGRAPHY

[1] B.R.A. Wood and RW Harris. Structural integrity and remnant life evaluation of pressure equipment from acoustic emission monitoring. *International Journal of Pressure Vessels and Piping*, 77(2-3):125–132, 2000.

[2] J.H. Williams Jr and S.S. Lee. Acoustic Emission Monitoring of Fiber Composite Materials and Structures. *Journal of Composite Materials*, 12(4):348, 1978.

[3] M. Giordano, A. Calabro, C. Esposito, A. D'Amore, and L. Nicolais. An acoustic-emission characterization of the failure modes in polymer-composite materials. *Composites Science and Technology*, 58(12):1923–1928, 1998.

[4] NF Casey and PAA Laura. A review of the acoustic-emission monitoring of wire rope. *Ocean Engineering*, 24(10):935–947, 1997.

[5] DO Harris and HL Dunegan. Acoustic Emission Testing of Wire Rope. 1974.

[6] A. Choudhury and N. Tandon. Application of acoustic emission technique for the detection of defects in rolling element bearings. *Tribology International*, 33(1):39–45, 2000.

[7] MW Hawman, WS Galinaitis, U.T.R. Center, and E. Hartford. Acoustic emission monitoring of rolling element bearings. *Ultrasonics Symposium, 1988. Proceedings., IEEE 1988*, pages 885–889, 1988.

[8] J Kaiser. *Investigation of Acoustic Emission in Tensile Testing.* PhD thesis, PhD Dissertation, Technische Hochschule, Munich, Germany.

[9] Wikipedia page on landing gear, accessed 15 June 2009. URL http://en.wikipedia.org/wiki/Landinggear.

[10] R. Kyle Schmidt and Pia Sartor. Landing gear. In C. Boller, F. Chang, and Y. Fujino, editors, *Encyclopedia of Structural Health Monitoring*, chapter 114, pages 1983–1994. J. Wiley & Sons, Chichester, UK, 2009.

[11] S.G. Allison. Ultrasonic measurement of aircraft strut hydraulic fluid level, 2002. URL http://techreports.larc.nasa.gov/ltrs/PDF/2002/mtg/NASA-2002-wacd-sga.pdf.

[12] C. Seror. Method of measuring the compression of a shock absorber, and an airplane undercarriage, constituting an application thereof, 2005. U.S¿ Patent 2005/0230200.

[13] Messier dowty public webpage, accessed 15 June 2009. URL http://www.messier-dowty.com/ (and pages therein).

[14] P.D. Wilcox, C.K. Lee, J.J. Scholey, M.I. Friswell, and M.R Wisnom. Progress Towards a Forward Model of the Complete Acoustic Emission Process. *Advanced Materials Research*, 13–14:68–79, 2006.

[15] D. Chetwynd, K. Worden, A. Spencer, and J.Hensman. The Effects of Uncertainties within Acoutic Emission Modelling. *Proceeding of IMAC 2008*, 2008.

[16] MA Hamstad, SL Quarles, and RL Lemaster. Experimental Far-field Wideband Acoustic Waves in Wood Rods and Plates. *9th International Symposium on Nondestructive Testing of Wood, Madison Wisconsin*, 1994.

[17] HL Dunegan. Modal analysis of acoustic emission signals. *Journal of Acoustic Emission*, 15(1):1–4, 1998.

[18] M.R. Gorman and W.H. Prosser. AE Source Orientation by Plate Wave Analysis. *Journal of Acoustic Emission(USA)*, 9(4):282–288, 1990.

[19] R. Pullin, KM Holford, and MG Baxter. Modal Analysis of Acoustic Emission Signals from Artificial and Fatigue Crack Sources in Aerospace Grade Steel. *Key Engineering Materials*, 293:217–224, 2005.

[20] Spencer et al . Source Identification of Laser Generated Acoustic Emission using LISA and differential Evoluition. *in press*, 2008.

[21] J.L. Rose. *Ultrasonic Waves in Solid Media.* Cambridge University Press, 1999.

[22] P Theobald and F Dar. Ae sensor calibration for out-of-plane and in-plane displacement sensitivity. *Advanced Materials Research*, 13-14:91–98, 2006.

[23] F. Hamel, JP Bailon, and MN Bassim. Acoustic Emission Mechanisms During High-Cycle Fatigue. *Engineering Fracture Mechanics*, 14(4):853–860, 1981.

[24] A. Berkovits and D. Fang. Study of fatigue crack characteristics by acoustic emission. *Engineering Fracture Mechanics*, 51(3):401–416, 1995.

[25] M. Klesnil and P. Lukas. Fatigue of Metallic Materials. *Elsevier Science Publishers, 1992,*, page 270, 1992.

[26] PC Paris and F. Erdogan. A critical analysis of crack propagation laws. J. Basic Engng. *Trans. ASME*, 85:528–534, 1963.

[27] TC Lindley, IG Palmer, and CE Richards. Acoustic Emission Monitoring of Fatigue Crack Growth. *Mater. Sci. Eng.*, 32(1):1–15, 1978.

[28] MA Hamstad. Contrasts between the Acoustic Emission Signals Generated by Monopole versus Dipole Sources. *Advanced Materials Research*, 13–14, 2006.

[29] CK Lee, JJ Scholey, PD Wilcox, MR Wisnom, MI Friswell, and BW Drinkwater. Guided Wave Acoustic Emission from Fatigue Crack Growth in Aluminium Plate. *Advanced Materials Research*, 13–14:23–28, 2006.

[30] RJ Boness, SL McBride, and M. Sobczyk. Wear Studies Using Acoustic Emission Techniques. *Tribology International*, 23(5):291–295, 1990.

[31] J. Shiroishi, Y. Li, S. Liang, T. Kurfess, and S. Danyluk. Bearing Condition Diagnostics via Vibration and Acoustic Emission Measurements. *Mechanical Systems and Signal Processing*, 11(5):693–705, 1997.

[32] N. Tandon and B.C. Nakra. Defect detection in rolling element bearings by acoustic emission method. *Journal of Acoustic Emission*, 9:25–28, 1990.

[33] J. Miettinen and P. Andersson. Acoustic emission of rolling bearings lubricated with contaminated grease. *Tribology International*, 33(11):777–787, 2000.

[34] JA Williams. *Engineering Tribology.* Oxford University Press, Oxford, 1994.

[35] IM Hutchings. Tribology: friction and wear of engineering materials. *Butterworth Heinemann*, 1994.

[36] N.N. Hsu. Acoustic emissions simulator, April 19 1977. US Patent 4,018,084.

[37] A. Nielsen. Acoustic emission source based on pencil lead breaking. *The Danish Welding Institute Publication*, 80:15, 1980.

[38] *Standard guide for determining the reproducibility of acoustic emission sensor response.* ASTM E97694.

[39] L. Wu, J.C. Cheng, and S.Y. Zhang. Mechanisms of laser-generated ultrasound in plates. *Journal of Physics D: Applied Physics*, 28(5):957–964, 1995.

[40] B. Lee. *LISA*. PhD thesis, Dept. of Mechanical Engineering, the University of Sheffield, 2006.

[41] H. Lamb. On Waves in an Elastic Plate. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character (1905-1934)*, 93(648):114–128, 1917.

[42] K.F. Graff. *Wave Motion in Elastic Solids.* Dover Publications, 1991.

[43] I.A. Viktorov. *Rayleigh and Lamb waves.* Plenum Press New York, 1967.

[44] R. Pullin, P. Theobald, KM Holford, and SL Evans. Experimental Validation of Dispersion Curves in Plates for Acoustic Emission. *Key Engineering Materials*, 2005.

[45] M. Hamstad. A wavelet transform applied to acoustic emission signals. II. Source location. *Journal of Acoustic Emission*, 20:62–82, 2003.

[46] WH Prosser, M.D. Seale, and B.T. Smith. Time-frequency analysis of the dispersion of Lamb modes. *The Journal of the Acoustical Society of America*, 105:2669, 1999.

[47] J.A. Hartigan. *Clustering Algorithms.* John Wiley & Sons, Inc. New York, NY, USA, 1975.

[48] C.M. Bishop. *Pattern recognition and machine learning.* Springer, 2006.

[49] I. Nabney. *NETLAB: Algorithms for Pattern Recognition.* Springer, 2002.

[50] H. Frigui and M.B.H. Rhouma. Clustering based on synchronization of pulse-coupled oscillators. *Fuzzy Information Processing Society, 2000. NAFIPS. 19th International Conference of the North American*, pages 128–132, 2000. doi: 10.1109/NAFIPS.2000.877403.

[51] L.J. Heyer, S. Kruglyak, and S. Yooseph. Exploring Expression Data: Identification and Analysis of Coexpressed Genes. *Genome Research*, 9(11):1106–1115, 1999.

[52] R. Pullin, J. Hensman, K.M. Holford, and Worden K. A Principal Componenet Analysis of Acoustic Emission Signals from a landing Gear Component. *Key Engineering Materials*, 347:139–144, 2007.

[53] MN Bassim. Detection of the onset of fatigue crack growth in rail steels using acoustic emission. *International Journal of Fatigue*, 17(4):307–307, 1995.

[54] TM Roberts and M. Talebzadeh. Acoustic emission monitoring of fatigue crack propagation. *Journal of Constructional Steel Research*, 59(6):695–712, 2003.

[55] C.M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, USA, 1995.

[56] K. Worden, G. Manson, and N.R.J Fieller. Damage Detection Using Outlier Analysis. *Journal of Sound and Vibration*, 229(3):647–667, 2000.

[57] L. Tarassenko, P. Hayton, N. Cerneaz, and M. Brady. Novelty detection for the identification of masses in mammograms. *Artificial Neural Networks, 1995., Fourth International Conference on*, pages 442–447, 1995.

[58] K. Worden. Structural Fault Detection Using a Novelty Measure. *Journal of Sound and Vibration*, 201(1):85–101, 1997.

[59] S. Singh and M. Markou. An approach to novelty detection applied to the classification of image regions. *Knowledge and Data Engineering, IEEE Transactions on*, 16(4):396–407, 2004.

[60] SJ Roberts. Extreme value statistics for novelty detection in biomedical dataprocessing. *Science, Measurement and Technology, IEE Proceedings-*, 147(6): 363–367, 2000.

[61] DL Hall and J. Llinas. An introduction to multisensor data fusion. *Proceedings of the IEEE*, 85(1):6–23, 1997.

[62] M. Friedrich, SG Pierce, W. Galbraith, and G. Hayward. Data fusion in automated robotic inspection systems. *Insight-Non-Destructive Testing and Condition Monitoring*, 50(2):88–94, 2008.

[63] K. Worden and WJ Staszewski. Data fusion–the role of signal processing for smart structures and systems. *Smart Technologies*, pages 71–100, 2003.

[64] P. Welch. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *Audio and Electroacoustics, IEEE Transactions on*, 15(2):70–73, 1967.

[65] HV Sorensen and CS Burrus. Efficient computation of the short-time fast Fourier transform. *Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on*, pages 1894–1897, 1988.

[66] P.S. Addison. *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance.* Institute of Physics Publishing, 2002.

[67] H. Suzuki, T. Kinjo, Y. Hayashi, M. Takemoto, K. Ono, and Y. Hayashi. Wavelet transform of acoustic emission signals. *Journal of acoustic emission*, 14(2):69–84, 1996.

[68] M.A. Hamstad, A. OGallagher, and J. Gary. A wavelet transform applied to acoustic emission signals. Part 1: source identification. *J. Acoust. Emission*, 20:39–59, 2002.

[69] H. Jeong and Y.S. Jang. Fracture source location in thin plates using the wavelet transformof dispersive waves. *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, 47(3):612–619, 2000.

[70] A. Grossmann and J. Morlet. Decomposition of functions into wavelets of constant shape, and related transforms. *Mathematics and Physics, Lect. Recent Results, Bielefeld/FRG*, 84:135–165, 1983.

[71] J. Kennedy and R. Eberhart. Particle swarm optimization. *Neural Networks, 1995. Proceedings., IEEE International Conference on*, 4, 1995.

[72] S. Kirkpatrick, CD Gelatt, and MP Vecchi. Optimization by Simulated Annealing. *Science*, 220(4598):671–680, 1983.

[73] R. Storn and K. Price. Differential Evolution–A Simple and Efficient Heuristic for global Optimization over Continuous Spaces. *Journal of Global Optimization*, 11(4):341–359, 1997.

[74] WJ Staszewski and KM Holford. Wavelet signal processing of acoustic emission data. *Key Engineering Materials*, 204:351–358, 2001.

[75] J. Hensman and K. Worden. Wavelet Based Feature Extraction for Acoustic Emission. *Proceedings of the 13th International Conference on Experimental Mechanics, Alexandroupolis, Greece, July 1-6, 2007.*

[76] DE Newland. *An introduction to random vibrations, spectral and wavelet analysis.* Halsted Press, 1994.

[77] J. Hensman, K. Worden, and S. Mohd-Saleh. Feature Extration from Acoustic Emission Waveforms using a Discrete Wavelet Transform. *proceedings of EWSHM 2008*, 2008.

[78] S. Rippengill, K. Worden, KM Holford, and R. Pullin. Automatic Classification of Acoustic Emission Patterns. *Strain*, 39(1):31–41, 2003.

[79] G. Manson, K. Worden, K. Holford, and R. Pullin. Visualisation and Dimension Reduction of Acoustic Emission Data for Damage Detection. *Journal of Intelligent Material Systems and Structures*, 12(8):529, 2001.

[80] G. Shakhna-Rovich, T. Darrell, and P. Indyk. Nearest-neighbor methods in learning and vision. *IEEE Trans Neural Netw*, 19(2):377, 2008.

[81] R. Shalkoff. *Pattern Recognition.* John Wiley, 1992.

[82] S. Roweis and Z. Ghahramani. A Unifying Review of Linear Gaussian Models. *Neural Computation*, 11(2):305–345, 1999.

[83] PF Baldi and K. Hornik. Learning in linear neural networks: a survey. *Neural Networks, IEEE Transactions on*, 6(4):837–858, 1995.

[84] T. Kohonen. *Self-Organizing Maps.* Springer, 2001.

[85] N.D. Lawrence. Gaussian Process Latent Variable Models for Visualisation of High Dimensional Data. *Advances in Neural Information Processing Systems 16: Proceedings of the 2003 Conference*, 2004.

[86] C.M. Bishop, M. Svensen, and C.K.I. Williams. GTM: The Generative Topographic Mapping. *Neural Computation*, 10(1):215–234, 1998.

[87] C.M. Bishop, M. Svensen, and C.K.I. Williams. Developments of the generative topographic mapping. *NEUROCOMPUT*, 21(1):203–224, 1998.

[88] C. Orphanidou, I.M. Moroz, and S.J. Roberts. Voice Morphing Using the Generative Topographic Mapping. *Proceedings of CCCT 03*, 1:222–225.

[89] D. Merkl. Exploration of text collections with hierarchical feature maps. *Proceedings of the 20th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 186–195, 1997.

[90] J. Peltonen. *Data Exploration with Learning Metrics.* Helsinki University of Technology, 2004.

[91] M.E. Tipping and C.M. Bishop. Probabilistic Principal Component Analysis. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 61 (3):611–622, 1999.

[92] J.W. Sammon. A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers*, 18(5):401–409, 1969.

[93] D. Lowe and M.E. Tipping. Neuroscale: novel topographic feature extraction using RBF networks. *Advances in Neural Information Processing Systems*, 9: 543–9, 1997.

[94] B. Scholkopf, C.J.C. Burges, A.J. Smola, et al. Advances in Kernel Methods: Support Vector Learning, 1999.

[95] K.Q. Weinberger, F. Sha, and L.K. Saul. Learning a kernel matrix for nonlinear dimensionality reduction. *ACM International Conference Proceeding Series*, 2004.

[96] N. Cristianini and J. Shawe-Taylor. An introduction to Support Vector Machines. 2000.

[97] S.T. Roweis and L.K. Saul. Nonlinear Dimensionality Reduction by Locally Linear Embedding, 2000.

[98] J.B. Tenenbaum, V. Silva, and J.C. Langford. A Global Geometric Framework for Nonlinear Dimensionality Reduction, 2000.

[99] Physical Acoustics Corporation. *PCI-2 Based AE System User's Manual Rev 1a.* 2003.

[100] J.A. Nelder and R. Mead. A simplex method for function minimization. *Computer Journal*, 7(4):308–313, 1965.

[101] KM Holford, AW Davies, R. Pullin, and DC Carter. Damage Location in Steel Bridges by Acoustic Emission. *Journal of Intelligent Material Systems and Structures*, 12(8):567, 2001.

[102] J.H. Kurz, C.U. Grosse, and H.W. Reinhardt. Strategies for reliable automatic onset time picking of acoustic emissions and of ultrasound signals in concrete. *Ultrasonics*, 43(7):538–546, 2005.

[103] PT Coverley and WJ Staszewski. Impact damage location in composite structures using optimized sensor triangulation procedure. *Smart Materials and Structures*, 12(5):795–803, 2003.

[104] F. Schubert. Basic Principles of Acoustic Emission Tomography. *Journal of Acoustic Emission*, 22:147–158, 2004.

[105] M.G. Baxter, R. Pullin, K.M. Holford, and S.L. Evans. Delta T source location for acoustic emission. *Mechanical Systems and Signal Processing*, 21(3):1512–1520, 2007.

[106] R.M. Neal. *Bayesian Learning for Neural Networks.* Springer, 1996.

[107] D.J.C. MacKay. Information theory, inference, and learning algorithms. *Neural Computation*, 10(1):215–234, 2003.

[108] C.E. Rasmussen and C.K.I. Williams. *Gaussian Processes for Machine Learning.* Springer, 2006.

[109] K.B. Petersen and M.S. Pedersen. The Matrix Cookbook, 2006.

[110] N. Lawrence, M. Seeger, and R. Herbrich. Fast Sparse Gaussian Process Methods: The Informative Vector Machine. *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS*, pages 625–632, 2003.

[111] E. Snelson and Z. Ghahramani. Sparse Gaussian Processes using Pseudo-inputs. *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS*, 18:1257, 2006.

[112] A.J. Smola, SVN Vishwanathan, and T. Hofmann. Kernel methods for missing variables. *Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics*, 2005.

[113] GE Hinton. Products of experts. *Artificial Neural Networks, 1999. ICANN 99. Ninth International Conference on (Conf. Publ. No. 470)*, 1, 1999.

[114] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–. URL `http://www.scipy.org/`.

[115] BW Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC, 1986.

[116] Holford K.M., Evans S., Pullin R., Eaton M., Worden K., Hensman J., Baker T., Bond D., Cole P., Bradshaw T., and Theobald P. Characterisation of damage in complex structures by acoustic emission during fatigue testing (methodology document), 2008. Technical report submitted to Messier-Dowty.

[117] M. Kulldorff. A spatial scan statistic. *Communications in Statistics-Theory and Methods*, 26(6):1481–1496, 1997.

[118] F. Mostashari, M. Kulldorff, J.J. Hartman, J.R. Miller, and V. Kulasekera. Dead Bird Clusters as an Early Warning System for West Nile Virus Activity. *EMERGING INFECTIOUS DISEASES*, 9(6):641–646, 2003.

[119] E.S. Jefferis. A multi-method exploration of crime hot spots: SaTScan resultsi. *National Institute of Justice*, 1998.

[120] C.A. Paget. Damage Assessment in a Full-Scale Aircraft Wing by Modified Acoustic Emission. *Proceedings of 2nd European Workshop on Structural Health Monitoring*, 2003.

[121] M. Mitchell. *An Introduction to Genetic Algorithms*. Bradford Books, 1996.