THE INCREMENTAL MOTION ENCODER: A SENSOR FOR THE INTEGRATED CONDITION MONITORING OF ROTATING MACHINERY

O. K. Ayandokun

A thesis submitted in partial fulfilment of the requirements of The Nottingham Trent University for a degree of Doctor of Philosophy

December 1997

Department of Computing, The Nottingham Trent University,

Burton Street, Nottingham NG1 4BU, United Kingdom

Abstract

In highly automated modern process and manufacturing industries predictive maintenance has become the prime method of avoiding costly unscheduled stoppages due to machine breakdown. A successful predictive maintenance strategy is based on the condition monitoring of equipment through the selection, measurement, and trending of critical machine parameters. Monitoring these parameters allows detection of the gradual deterioration of key components without physical examination. This ensures that only necessary repairs and replacements are conducted. Maintenance resources are optimally employed, downtime of equipment is reduced and maximum economic use is gained from components with a certain wear lifetime.

The key components in many rotating machines, from large turbines to small scale machine tools, are the rolling element bearings that support shafts and spindles. As a bearing begins to fail the frequency at which its rotation excites its supporting structure changes. The machine vibration caused by this excitation is currently the main parameter used to monitor bearing condition. An investigation of the conventional techniques for vibration based bearing monitoring using analogue electromechanical sensors showed that despite good performance in ideal conditions there were many common circumstances where the vibration at the machine surface might not reflect the state of a machine's bearings accurately. The vibration signal from a bearing commonly suffers attenuation in transmission through the machine structure to the machine casing. As the bearing may be only one of several vibration sources, adequate diagnosis of an individual fault can be difficult, particularly in the early stages of bearing failure. The most commonly used vibration sensor, the accelerometer, has a limited ability to measure the low level, low frequency vibrations from bearings in machines operating at low rpm.

Proximity measurement of shaft relative displacement has been suggested to overcome the problems inherent in monitoring bearing condition via transmitted vibration. Previous attempts to use this technique were limited by the lack of a sensor capable of measuring the extremely small clearances between a shaft and its supporting bearing. It was suggested that the principle of a new sensor concept, the Incremental Motion Encoder, offered the ability to sense the mechanical motion of a shaft due to defects in its supporting bearings.

The hardware and software of a novel bearing condition monitoring system were successfully developed around a prototype Incremental Motion Encoder sensor. Signal processing algorithms were developed which enabled the experimental system to detect a variety of induced bearing defect conditions. The system was able to identify individual defects on bearing elements and pre-defect bearing conditions of lubricant contamination and corrosion. The principal claim to novelty made is the method by which an IME signal is interpreted to yield an accurate indication of bearing condition. The successful application of the IME principle to rolling element bearing condition monitoring represents the first completely new technique for a decade in this area.

On the basis of the results of experimentation with this system a completely new method of monitoring bearings is proposed in which an Incremental Motion Encoder sensor is integrated with the bearing as a single unit.

The work described in this thesis is the author's own, unless otherwise stated, and is, as far as he is aware, original.

Dedication

This work is dedicated to all those throughout human history who have struggled to make humanity something more than what it was.

Acknowledgements

My mother, for all her hard work and for teaching me to work hard and never give up. *"I shall never rest until my good is better and my better, best."*

Colour Sergeant A. Pickford, Royal Military Academy, Sandhurst, for reinforcing my mother's lessons. *"Harder, harder yet!"*

My supervisors, Dr. N. Sherkat, Dr. P. A. Orton, and Dr P.D. Thomas, for their support, encouragement, and readily available help.

Mr. Tim Sellears, for his help with the hardware element of the project.

All my friends and colleagues who enlivened my time as a researcher.

Mr. Sanjay Bhubuta, of Hewlett-Packard Components Limited, for providing the kit encoder parts for the experimental incremental motion encoder.

Muirhead Vactric Components Limited, for supporting the project by producing the prototype dual read head and fibre optic encoders.

Contents

List of Appendices

Glossary of Terms

ABS	Anti-lock Braking System
ADC	Analogue to Digital Converter
ANN	Artificial Neural Network
CBM	Condition Based Maintenance
CIM	Computer Integrated Manufacturing
CM	Condition Monitoring
DCU	Data Collection Unit
DMA	Direct Memory Access
DSP	Digital Signal Processing
DTI	Department of Trade and Industry
FFT	Fast Fourier Transform
FIFO	First In First Out
GUI	Graphical User Interface
HFRT	The High Frequency Resonance Technique
HP	Hewlett Packard
IME	Incremental Motion Encoder
IOC	Iron ore company Of Canada
LAN	Local Area Network
MES	Maintenance Expert System
MMS	Maintenance Management System
MMU	Memory Management Unit
PC	Personal Computer

PI Parallel Interface

PM Predictive Maintenance

RCM Reliability Centred Maintenance

REB Rolling Element Bearings

REBAM Rolling Element Bearing Activity Monitor

VSM Vital Signs Monitoring

Chapter 1

Introduction

1. Introduction

This chapter introduces the work conducted in this PhD research project. The sensor concept of the Incremental Motion Encoder is described from its basis in previous work. From this point, the research objectives set at the start of the project are outlined. The third section summarises the principal findings made as a result of this project.

1.1 Previous work

Work conducted by Orton into the instrumentation and control of precision grinding machines had involved the use of optical shaft encoders for the precision measurement of angular velocity during the grinding process [1]. The study focused on an existing grinding machine that used open loop manual control. The development of a comprehensive control system for this machine depended on the measurement of several critical process variables. Orton concluded that a single precision digital sensor could be developed from the standard optical incremental shaft encoder. This, it was suggested, would partially rationalise the instrumentation requirements of the grinding process. It was predicted that the sensor would be able to measure angular position, angular velocity, angular vibration and shaft centre position in two dimensions.

The proposed sensor was christened the Incremental Motion Encoder (IME). An outline description of the IME was contained in an appendix to Orton's work [1]. The appendix described the device's principle of operation and contained a schematic illustration, reproduced in figure 1. The appendix also contained the derivation of the mathematical formula required to calculate shaft displacement from the geometrical relationship of two opposed encoder readheads. The actual physical design of the sensor and its signal processing requirements were identified as directions for future research work.

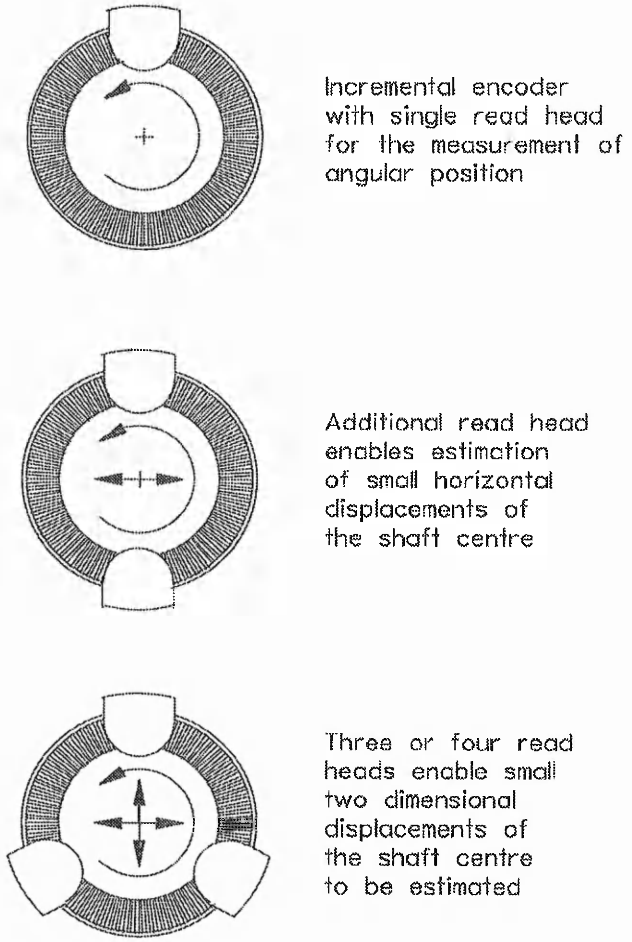


Figure 1. Original IME schematics [1].

Based on the parameters that the IME was expected to be able to measure three possible applications were suggested:

1. Measurement of shaft motion and loading relative to automatically controlled hydrostatic oil or magnetic bearings.
2. Measurement of vibrations in the structure of a machine.
3. Gauging of cylindrical components.

After the completion of his initial research Orton went on to patent the idea of the IME [2]. The patent application consisted of the basic idea of the IME, as outlined in [1], with the addition of an algorithm for the calculation of shaft displacement from three readheads and a schematic design for a computer interface circuit for data collection. A reproduction of this patent application is contained in appendix A.

The IME patent application formed the basis for an MSc project [3] undertaken by the author in the six months prior to the start of the PhD research presented in this thesis. At the start of the MSc project the basic requirements of an IME system were analysed. From the results of this analysis a slightly modified version of the interface circuit described in the patent application was built. A computer system and software were also developed to allow partial testing of the IME principle. A prototype IME device was created from a standard optical incremental shaft encoder. This had been modified by the manufacturer so that it had two readheads in opposition. This device was not used to experimentally monitor any external event or process. It was simply driven by a motor and the data collected by the system analysed to see if the normal motion of the encoder disk could be resolved. Technical problems with the implementation of the system meant that only a small amount of data was collected. This data could be displayed graphically, but the software written at this time was not sophisticated enough to allow any data analysis. At the end of this project it was concluded from this data that the system was returning the type of data related to shaft motion expected.

A limited literature survey was undertaken as part of the MSc project. This was aimed at providing background knowledge of the IME's proposed application area. Initially the search was based on the idea, contained in the previous work of Orton, that the IME would be a precision sensor for process measurement [1,2]. However examination of the literature showed that the condition monitoring of rotating machines was the main field of interest for shaft displacement measurement. In particular the work of Collacott [4] describing the advantages of shaft proximity measurements over remote casing mounted sensors was very influential on the approach that was followed in the succeeding PhD project. Chapter 2 details this and its context in subsequent literature surveys. Based on this initial research the conclusions of the MSc thesis placed the emphasis on the potential application of the IME principle to condition monitoring. This set the direction for future work.

1.2 Research objectives

The objectives for this PhD research project were decided partly by the state of the previous work outlined above and partly by accounts of other condition monitoring research projects found in the literature. At its start the project aim was stated as:

> To develop a new type of shaft displacement transducer system based on an incremental motion encoder. This system will require signal processing computer hardware and software implementing novel signal processing algorithms. The

investigation will examine the ability of this system to measure sub-micron displacement motion of a rotating shaft for the purpose of real-time condition monitoring. Other applications for this transducer technology will also be investigated.

To achieve this aim a research plan was drawn up and followed during the course of the project. The structure of this thesis generally follows the outline of that plan. Chapter 2 is drawn from the literature survey conducted during the project and describes the condition monitoring of rolling element bearings. The main problems encountered with current techniques are discussed. Chapter 3 describes the prototype IME sensor and the construction of the experimental condition monitoring system. Chapter 4 describes the results of the initial experimentation undertaken with the sensor system. Chapter 5 explains the model that was developed to interpret the IME signal based on this initial experimentation. Chapter 5 also contains a description of the main experimental work done to validate the new model. This involved the use of the sensor to detect various types of bearing defect conditions. Chapter 6 contains the conclusions drawn at the end of this work and outlines the areas envisaged for future work.

1.3 Principal research findings

A prototype incremental motion encoder was successfully constructed based on the design of Orton with some modifications to the computer architecture [2]. This was used as the basis of a novel experimental rolling element bearing condition monitoring system which was designed and developed during this project. With this system it was possible to apply the IME principle in a series of experiments designed to investigate the sensor's ability to detect and measure different types of rolling element bearing condition. These experiments proved that the novel, and previously untested, IME concept is valid and can be successfully applied in this area.

The principal claim to novelty which will be made by this project is in the manner of interpreting the IME signal to yield an accurate indication of the condition of a rolling element bearing from measurements of shaft motion. This successful application of the IME principle to the condition monitoring of rolling element bearings represents the first completely new technique in this area for a decade.

A model has been developed to predict the response of the IME signal to defects on individual bearing elements. This model has been initially confirmed by experimentation which has shown the IME based system can clearly detect large, but still realistically sized, defects on the inner and outer races of a bearing. A method of discriminating between the impulses in the IME signal due to these defects and those due to other machine components has been developed based on the angular separation of the impulses and their relative position between successive rotations. This potentially gives the IME based system the important attribute of being able to monitor individual bearings in a complex machine with many sources of impulsive vibration.

The main conclusion drawn at the end of the project is that there is a great potential for a new generation of bearings incorporating the IME principle. These could use the existing technology available for speed sensing bearings which have an internal encoder. Such bearings could monitor machine parameters such as angular velocity, angular vibration, torque and shaft loading together with their own condition. A number of conventional sensors and their respective data analysis systems could be replaced with a single unit. This could be integrated into existing machine designs with only minimal modifications for cabling. A bearing of this type could be used for process control and monitoring during normal machine operation as well as for detecting fault conditions. The term 'smart bearing' has been used by the author to introduce this novel concept into the literature. Figure 2 shows an artist's impression of a smart bearing based on a drawing of an existing speed sensing bearing. As a system component, a smart bearing would be a significant step forward in the direction of the type of intelligent machine required for use in computer integrated manufacturing systems [5].

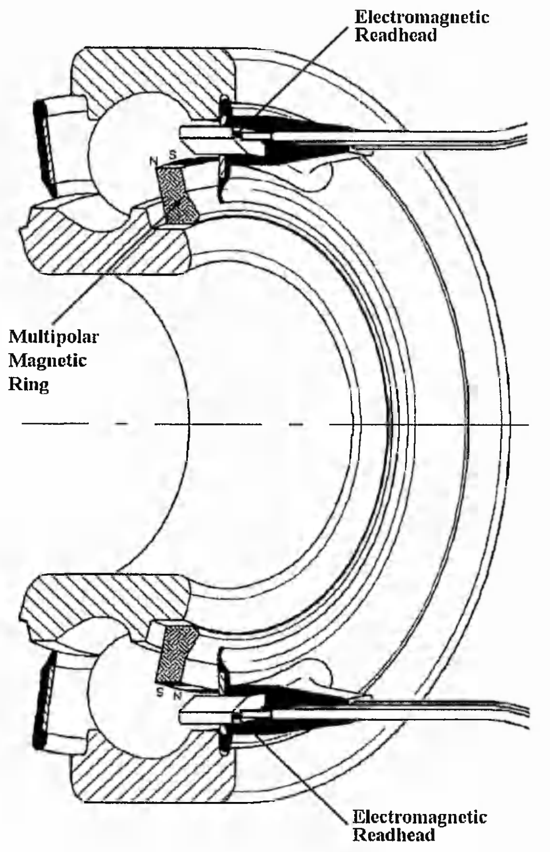


Figure 2. An intelligent or "smart" bearing could be created by the addition of a second readhead to a speed sensing bearing.

Chapter 2

Condition monitoring for rolling element bearings

2. Condition monitoring for rolling element bearings

This thesis introduces research contributing to the industrially important field of condition monitoring. This chapter begins by explaining the reason for the increased emphasis on effective maintenance in modern industry. The different strategies used for maintenance management are briefly reviewed as an introduction to condition based maintenance, the strategy increasingly being adopted for modern machinery. A summary of condition monitoring methods for rotating machines focusing on rolling element bearings (REB) is then presented. This is based upon the literature survey conducted throughout the project. The survey aimed to cover current commercial and experimental methods based upon vibration analysis. These techniques and the methodologies employed are directly relevant to the IME. No attempt has made to comprehensively cover all available condition monitoring methods. Other possible techniques have only been examined for comparison.

2.1 The increasing importance of maintenance in modern industry

Over the past four decades the general industrial scene has been transformed by sustained economic pressure and rapid technological change. In the process and manufacturing industries the result has been a large increase in the size of production units in which the continuous process has largely replaced the batch process. Many of these units are now integrated to form large process plant complexes with minimal, if any, standby capacity. These units are highly inter-dependent. As operating cycles have become faster and more demanding, complex control and management systems have been introduced to replace manually controlled operations.

To meet the objectives of maximum output combined with minimum production and maintenance costs, plant stoppages must be avoided and their disruptive impact minimised. Maintenance has become an essential requirement to get optimum usage from today's increasingly complex industrial machinery. The high capital cost of the sophisticated machine tools in the computer integrated manufacturing (CIM) systems in modern automated factories requires them to be operated at peak efficiency to generate maximum return. Effective maintenance is essential to guarantee the high reliability that successful operation of a CIM system requires. Machine condition is also directly related to production quality.

Within industry, maintenance management has sought to devise new strategies and use new techniques to avoid unscheduled plant stoppages and gain maximum service with absolute safety. Three generally applied approaches to maintenance management have evolved with the changes in industrial conditions [6]. The following sections summarise each in turn.

2.1.1 Breakdown maintenance

Breakdown maintenance is the oldest approach to maintenance management. Using this strategy machines are simply run until they fail or their performance degrades to the point where parts must be scrapped. Failure is random and unexpected. It can be catastrophic at worst and untimely at best. If production is not to be seriously delayed a full inventory of replacement parts must be carried which ties up capital. Personnel must also be constantly ready to effect timely repairs. In modern industry this type of maintenance is restricted to non-critical or easily repaired machinery whose failure will not affect production [7].

Where machinery is critical to production, or technically complex, maintenance prior to any breakdown or reduction in product quality is required.

2.1.2 Preventative maintenance

Preventative, or scheduled, maintenance aims to shut down critical machines after specified periods of operation based on statistical mean-time-to-failure data. The machine is partially or completely disassembled for inspection and worn parts, if any, are replaced. This method has several disadvantages. Dismantling a critical machine is time-consuming, expensive and may well be unnecessary. A healthy machine may also be degraded by incorrect reassembly leading to an increased likelihood of failure.

The life of an industrial machine can be summarised by the classic 'bath-tub' curve shown in figure 3 [6]. During an initial run-in period, when new and after repair, the potential for failure is high. The machine then goes into a period of continuous stable operation with a reduced risk of failure. As the machine's components begin to wear, the risk of failure increases and the machine goes into the third and final phase of wear out. A preventative maintenance strategy increases the number of run-in phases and thus the overall potential for machine failure is increased.

Correctly estimating the period of stable machine operation is difficult. This is especially true for REB. Bearing service life is calculated from load analysis of the bearing materials and models describing the kinetics in the bearing [8]. A review by Tallian of eleven published bearing life models shows the development of these models is a large area of research in itself [9,10]. An ISO standard has been defined which provides the baseline for current work in this field [11]. In all cases however the models only allow a calculation of bearing life in ideal operating conditions. Various commonly encountered operating factors such as lubricant contamination, shaft misalignment or improper loading can cause bearing life to be significantly shorter than an estimate derived from a model [8]. The same problem is encountered with other machine components. The result is that unpredictable machine failure cannot be entirely avoided by a preventative maintenance approach.

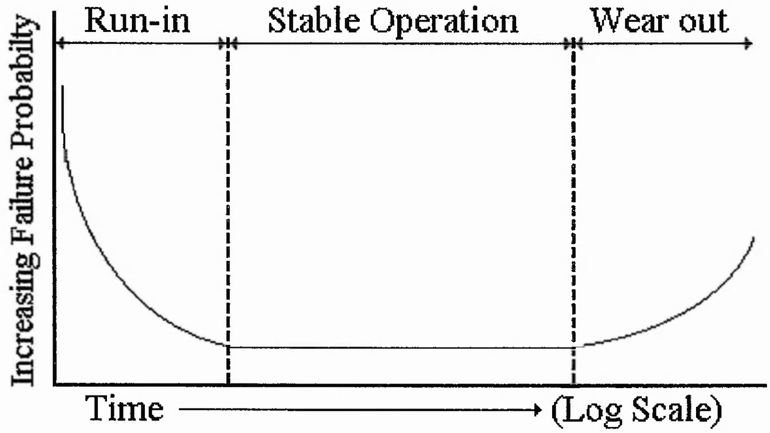


Figure 3. The machine life "bath tub" curve.

The difficulty in accurately estimating the start of machine failure due to wear means a significant margin of error will be allowed in the scheduling of maintenance. As a result components will not be allowed to reach the end of their full useful operating lives before being scrapped. This safety factor will obviously result in a significant unnecessary cost equivalent to the remaining useful life of the scrapped component. For systems employing preventative maintenance this cost is usually offset against the cost avoided in lost production or safety terms of machine failure.

Despite its disadvantages preventative maintenance remains a widely accepted approach to maintenance management. This is particularly true in those industries where safety is a paramount requirement, such as airlines and nuclear power generation. In the airline industry preventative maintenance has been developed into a sophisticated maintenance strategy known as reliability centred maintenance (RCM). RCM relies on a rigorous logical analysis of the maintenance needs of each aircraft system to determine servicing intervals and maintenance tasks [12]. The application of RCM in this industry is helped by the fact that aircraft designs feature systems that are clearly specified and standardised to operate in a known manner. Production and process industries conversely almost always feature unique designs to meet a wide range of output requirements. This has prevented a widespread application of RCM in these areas.

2.1.3 Condition based maintenance

Nickerson and Hall, reviewing research in this area [13], define condition based maintenance (CBM), also known as predictive maintenance (PM), as:

> *A maintenance philosophy in which equipment is maintained only when there is objective evidence of an impending failure in that piece of equipment.*

Using this approach machine condition is monitored, continuously or at intervals, using some form of sensor(s). This condition monitoring (CM) allows early detection of developing faults and is used to pinpoint where and when specific repairs are actually required. Machine shutdowns can then be scheduled for convenient times, well before damage becomes too severe. Repair times, and thus production losses, are minimised, as are personnel and replacement part requirements.

Condition based maintenance is now applied widely in industries such as metalworking, electricity generation, paper making, mining, and chemical processing. The wide range of equipment types monitored include machine tools, turbines, electric motors, conveyor belt drives, rolling mills, and gear drives. By eliminating catastrophic failures, reducing machine damage and increasing machine availability, lost production is minimised. Production quality is also maintained at a high level. This leads to increased revenues. Considering the usually high costs of lost production, there is a very short payback for investment in the right condition based maintenance system [14]. Hyde *et al.* report the results of a recent extensive survey of CBM in the UK paper and food industries supported by the department of trade and industry (DTI) [15]. This estimated that £150 million per annum of lost production could be saved in these sectors by the wider application of CBM.

There are four basic areas of machine condition which are commonly monitored [5]:

1. Visual inspection - This may involve simple manual techniques such as the use of boroscopes, or more complex automatic systems, using computer based vision technology in a variety of spectra. Infra-red thermal imaging is frequently used to monitor the external temperature of insulated tanks.

2. Performance assessment - This can involve can a variety of different techniques and criteria against which an assessment of performance is based. Measuring the variations in,

and / or absolute values of, machinery output in terms of quantity and / or quality is one technique area. Others involve measuring the system input / output relationship, or simultaneously comparing output parameters with sets of standard operating conditions. Faults in machine operation can be detected by monitoring the feedback signals to, for example, a servo motor current controller, during machine operation and comparing the signal to that obtained when the machine was known to be in a healthy state [16]. For machine tools which contain rotating parts, shaft speed measurements made during drilling and milling operations for process control have been shown to be useful for condition monitoring [17,18]. The changes in speed and torque during cutting operations can be used to detect abnormal conditions such as excessively worn or broken tools.

3. Wear debris analysis - This is essentially the examination of samples of lubricant from some form of closed loop system for evidence of unusual or excessive wear. For rotating machines this frequently involves the drawing off of lubricating oil from bearings or gearboxes. A diagnosis is then made as to their condition based on the number, size, shape, colour and constituent material of wear particles present [19,20]. The techniques used to perform this analysis have been expensive to perform, difficult to automate and required the expertise of trained laboratory personnel [6]. In recent years however computer technology combined with new types of particle detector have permitted automation of some forms of this technique [21,22].

4. Vibration monitoring - This is an area that is synonymous with the monitoring of rotating machinery and is the most common method for monitoring REB. Machine condition is diagnosed from the analysis of vibration signals from the machinery under test. The

techniques involved in the diagnosis lend themselves to automation via computer technology. Section 2.2 discusses vibration based condition monitoring and the relevance of the IME principle to this area in detail.

The areas outlined above are not mutually exclusive. The first stage in the implementation of a successful condition based monitoring strategy is deciding which machine parameter(s) to monitor. A detailed analytical study of the machinery has to be carried out which may identify several possible fault conditions each of which may require a different detection technique. Modern comprehensive condition monitoring systems do not simply rely on the measurement of a single parameter. They allow diagnostics to be based on such variables as temperature, noise, ferrography, infrared spectrum and surface roughness as well as vibration [23]. This allows the integrated monitoring of such parameters as vibration, on-line wear, stress and leak detection. Recent work by Bonga and Robichaud illustrates the significant benefits of an integrated approach [24]. They describe an example where trends in vibration failed to show the development of faults in a DC motor gearbox until wear was advanced. The addition of oil analysis allowed earlier detection of faults. This added an extra factor of confidence into the monitoring of the gearbox. The oil trends could be used as a screening tool reducing the effort needed for the analysis of vibration data.

The emergence of the microprocessor based personal computer (PC) and microcontroller have allowed the creation of low cost integrated CM systems. These are able to provide plant wide monitoring and diagnostics based on many different parameters. An example of such a system is Mpulse™, a modular, fully networked hardware and software condition based maintenance data collection and analysis system produced by IRD Mechanalysis Inc.

It provides real-time untended acquisition of vibration data and other parameters from machinery throughout a plant. Figure 4 illustrates the system configuration of Mpulse™.

Data is collected by a host PC from a layered set of hardware modules which provide data acquisition and signal conditioning specific to the individual sensor type. These sensor personality cards allow the system to be tailored to the monitoring requirements of individual plant units. The connection of the host PC to a local area network allows operator displays wherever they are needed.

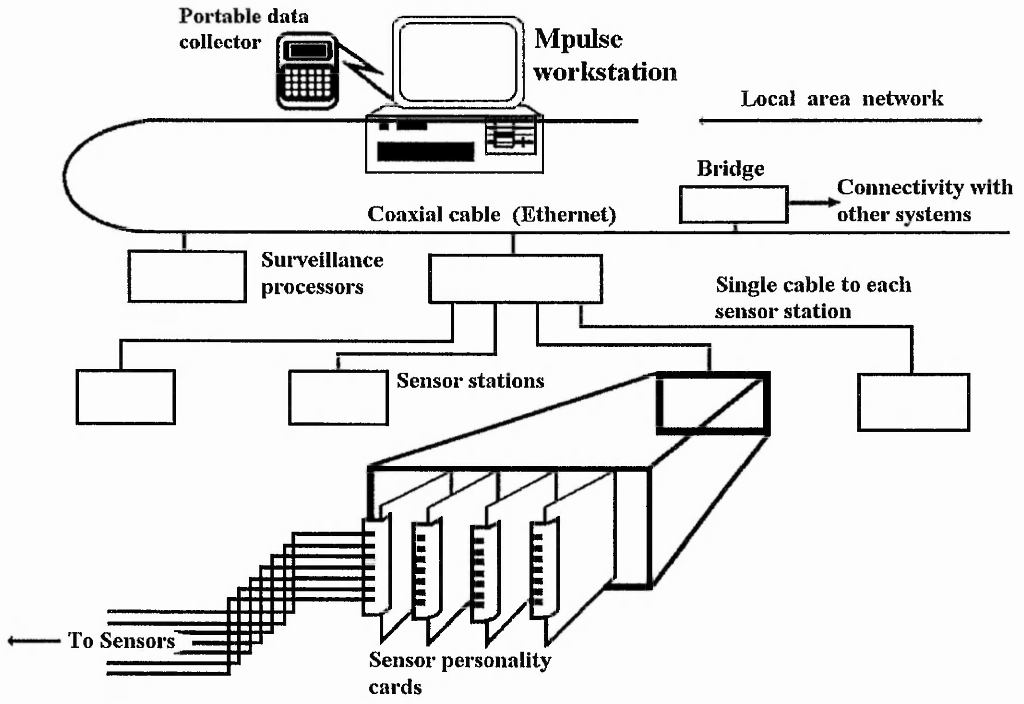


Figure 4. The system configuration of Mpulse^{TM}.

Recent studies have shown that the effectiveness of condition based maintenance depends more on the programme that is established to apply the monitoring techniques than the monitoring equipment itself [25]. At the heart of the Mpulse™ system is a database management system running under MS-DOS which is used to store the data from the network. Trend analysis continually updates the database for display and on-going analysis by expert software. This provides the system with predictive functions, automatically searching for machine problems using vibrational analysis and bearing problem expertise, for example, to pinpoint upcoming maintenance requirements.

A number of studies have proven that the pattern recognition, classification and generalisation ability of artificial neural networks (ANN) can be used for machine fault diagnosis [26,27]. The experimental work by Liu *et al.* is an example which has shown that ANN can detect bearing states and make diagnoses with a success rate between 97 - 100% [28,29,30]. Condition based maintenance systems are increasingly making use of artificial intelligence and knowledge based techniques [31,32,33]. These techniques are used to supplement, and may eventually replace, the highly trained human expert for routine operation.

The overall aim of an integrated condition monitoring system, such as Mpulse™, within a plant would be to allow the system to not only detect faults, but also to automatically order replacement parts and schedule personnel to effect repairs at the best time for optimum plant operation. A state of the art industrial application of such a system has been recently been described by Doucet and Goddard [34]. The Iron ore company of Canada (IOC) operates a large open cast mine in Newfoundland with a variety of heavy plant equipment. Maintenance costs account for 40% of the mine's budget at a time of increased competition and falling output prices. The mine has embarked on a "total maintenance integration plan" with the singular aim of increasing profitability. As is the case in most areas of industry, optimising maintenance management is seen as one of the key areas where savings can be made. Real time vital signs monitoring (VSM) systems are installed on haulage trucks, drills, shovels and the locomotives on the mine railway. Each VSM system uses a variety of sensors to monitor variables, such as vibration, oil debris or electric motor current, specific to the equipment it is monitoring. All the VSM systems throughout the mine are linked by radio transmitters to a central database which stores all digital, analogue and serial data from the sensors. Interfaced to this database are a maintenance expert system (MES) and a maintenance management system (MMS). The MES gathers uses from the database to evaluate the condition of each piece of monitored equipment. The MES itself consists of a number of constituent expert systems and trained neural networks. Each of these is designed to interpret a different set of sensor signals. When a problem is detected the MES can issue a work order via the MMS. The MMS schedules maintenance according to a rule base which takes into account the equipment condition and effect on production. It tracks the cost of materials and personnel, and compiles component histories. The whole system is interfaced to a Novell based local area network and is TCP/IP compatible. This allows unlimited access for technicians and managers across the mine site, and remotely, to the database and the MMS.

2.2 Vibration based condition monitoring

The basis of any effective condition based maintenance strategy is the accurate monitoring of the condition of the machines concerned. This requires identification of a parameter that can be reliably measured by a sensor to reveal the condition of critical components and signal developing faults. In many rotating machines the critical components are the rolling element bearings that support shafts and spindles. Machine vibration is the variable most often used to monitor these components.

All rotating equipment vibrates and the basis of vibration based condition monitoring is that the degree of vibration is a good indicator of equipment condition. Vibration monitoring is commonly used because of its ability to give warnings for a wider variety of rotating machinery faults than other detection techniques [35]. As well as bearing defects it can reveal component imbalance, seal wear, misalignment, gear damage and bent shafts. Other monitoring methods, if used in isolation, limit the variety of faults that can be detected. This leads to the possibility of failure from a fault whose symptoms are not covered.

The vibration measured on the casing of a rotating machine results from irregular motions of the machine's shaft, whether due to imbalance, misalignment, worn bearings, or other causes related to equipment being driven by the shaft. These irregular motions result in forces being transmitted via the shaft's bearings to the machine casing. It would be best to measure these forces directly, but since this has been impossible in practice, the resultant movement, the vibration, is measured. The mobility of the machine structure along the transmission path of the vibration determines the level seen at the machine casing as illustrated by figure 5.

A case study is quoted by Serridge to show the effect of mobility on vibration measurement [35]. Vibration levels measured on 17 industrial ethylene compressors of similar power, size and age, were found to vary by a factor of nearly 1000:1. This variation was due to differences in mobility.

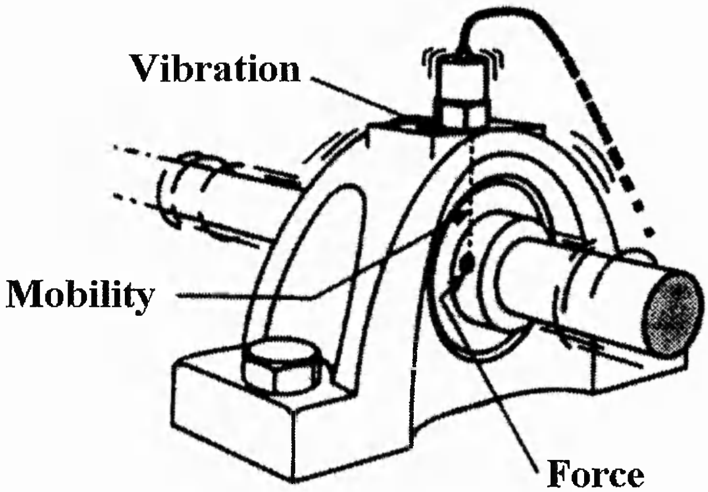


Figure 5. The vibration level measured on the surface of a machine depends on the condition *and* on the mobility of the machine [35].

Sensor placement on the machine casing is important to achieve maximum signal strength. The raw vibration signal at a single point on a machine is usually insufficient monitoring information on its own. Ideally a vibration should be measured at several carefully selected points and in several directions so that the combined signals can be analysed into the basic components of which make up the complex raw waveform.

The length of the transmission path is also a factor which determines the quality of the vibration signal received by a surface mounted sensor. The further away from the vibration source a sensor is mounted, the more the vibration signal will be attenuated. Joints between the machine parts will further attenuate the signal. Figure 6 illustrates the signal decay vs. transmission distance curve. There is sharp attenuation of the vibration signal at each joint. Documentation from the Bentley Nevada Corporation, a leading CM equipment company, reports that an interface such as an air gap or a water jacket can attenuate a vibration signal by a factor of 10 (20dB) [36].

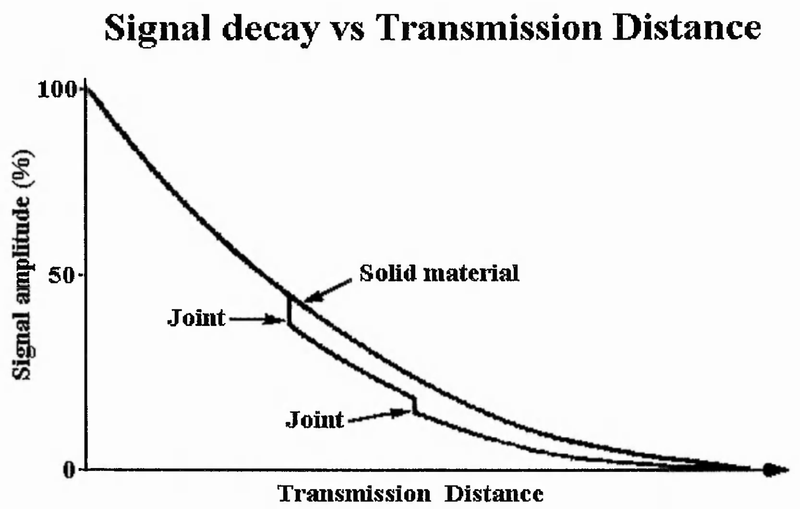


Figure 6. The effect of distance on vibration signal strength showing the sharp attenuation which occurs at joints between machine parts [36].

Ideal conditions for REB monitoring have been defined by Ray as those where the sensor can be mounted with a direct path between itself and the bearing, in the centre of the load zone and with the nominally loaded, properly lubricated bearing rotating above a minimum speed of 1000 rpm [37]. These conditions are not met in most complex machines, particularly those with a high density of vibration sources. In these cases proper diagnosis of individual components may be extremely difficult.

At its simplest vibration based condition monitoring relies on the fact that as machine condition degrades vibration level generally increases. The increasing vibration is a result of increasing tolerances in the machine due to wear [38]. Originally measurement of the overall root mean square level of the vibration was the main parameter monitored, the value measured being compared to recommended values for the equipment concerned [39]. This was then generally replaced by the trending of the overall vibration whereby the vibration level of the machine at installation was compared to that measured at subsequent times. An increase was taken to indicate a fault condition. Today almost all vibration monitoring is based upon the principle that the different components in a machine contribute certain specific frequency components to the overall vibration level [35,40]. These are examined in the frequency spectrum of the overall vibration, sometimes filtered, and are trended individually to give warning of developing faults in the relevant components.

Figure 7 shows a simple machine and its vibration frequency spectrum. The contributions of each machine element to the overall vibration amplitude are concentrated at distinct frequencies in the overall frequency spectrum. These frequencies can be calculated from the design of the machine. The frequency range expected, from the lowest to the highest useful part of the spectrum, determines which vibration sensor is selected to monitor the machine vibration.

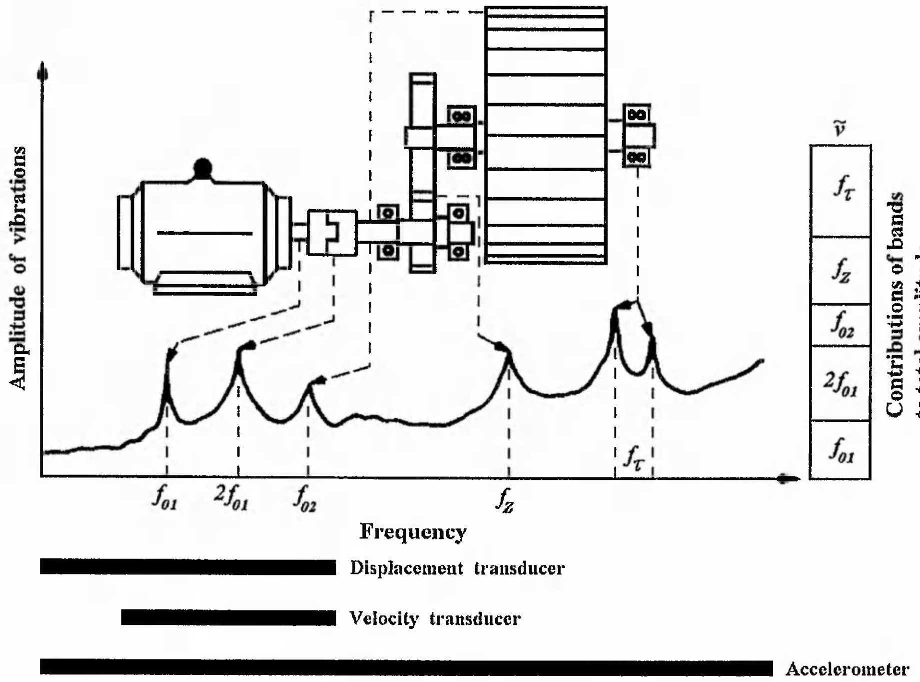


Figure 7. Different elements of a rotating machine contribute to the overall machine vibration spectrum at different frequencies [40].

2.2.1 Vibration monitoring sensors and selection criteria

Selection of the sensor type, its location and direction of measurement are the most critical factors in the successful use of vibration for monitoring of machinery condition [4]. If the raw sensor signal does not contain the discrete frequency components that are representative of a fault condition, no amount of signal processing or spectral analysis will reveal that condition.

The three dynamic parameters by which vibration is measured are velocity, acceleration and displacement. A different type of sensor is used for each parameter. The frequency at which measurement is important is the factor which determines which parameter is chosen, and thus which sensor type is used. A high signal-to-noise ratio is required to offset the effect of electrical 'noise' or interference due to stray currents. Each type of measurement records strong signals in different frequency ranges [4]. Table 1 below shows which parameter is preferred for the frequency ranges over which vibration is commonly monitored. This frequency response difference is illustrated graphically in figure 8.

Table 1

Frequency range (Hz)	*Preferred vibration parameter*
1 - 50	Displacement
50 - 1000	Velocity
1000 - 10,000	Acceleration

The dynamic range of each sensor type is also different. Figure 9 presents a comparison of the dynamic ranges of transducers commonly used in condition monitoring [35].

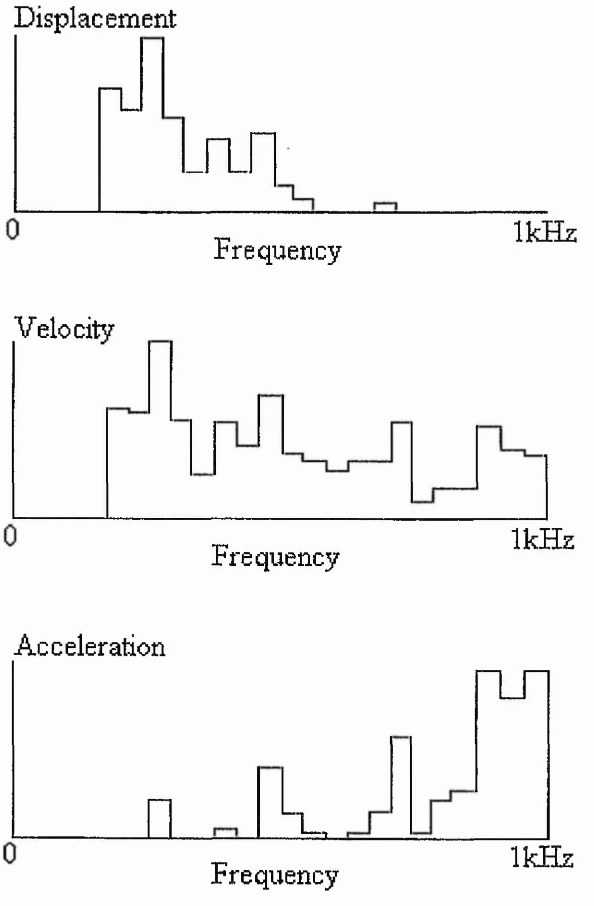


Figure 8. Octave frequency spectra resulting from the measurement of displacement, velocity and acceleration from the same rotating machine in the frequency range 0 - 1 kHz [4].

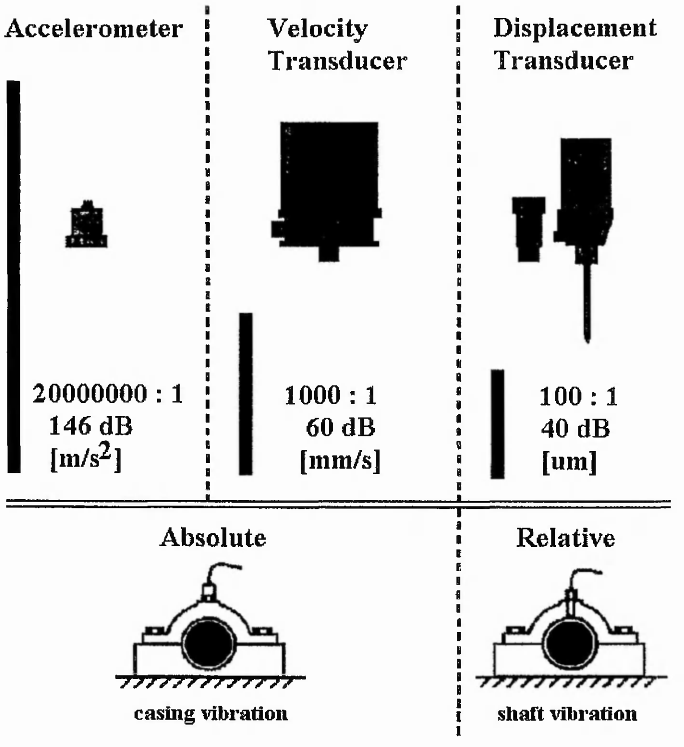


Figure 9. A comparison of the dynamic ranges of the transducers commonly used in vibration based condition monitoring [35].

2.2.1.1 Accelerometers

The casing mounted accelerometer is the current industry standard for implementing vibration monitoring [24]. Most accelerometers used today utilise either piezoelectric, piezoresistive or variable capacitance technology. All of these types are based on a similar principle. A mass is supported on a material, a piezoelectric crystal for example, whose electrical properties change with pressure. Vibration from the machine is passed to the mass causing deformation of the material. This produces an electrical signal proportional to the deforming force. Since $F = ma$ the output signal is a measure of the acceleration amplitude of the vibration.

There are many different accelerometer designs which use different piezoelectric materials. As well as machinery monitoring, various types of accelerometers are used for automotive crash studies, structural dynamics, aircraft flutter and whole body motion studies. The application determines which type of accelerometer is chosen. The piezoresistive and variable capacitance types offer the advantage of dc response. This makes them suitable for the measurement of long duration pulses. For the industrial measurement of machine vibration the compression and delta shear designs of piezoelectric accelerometer are widely used [35,4]. Figure 10 illustrates three different accelerometer designs from the catalogue of the Endevco corporation, an accelerometer manufacturer.

Low-g accelerometer with integral amplifier

General purpose industrial accelerometer

High-g shock testing accelerometer utilising silicon micromachined sensing element

Figure 10. Different accelerometer designs.

Recent work on crash sensing accelerometers for use in air bag deployment systems has produced an IC device to replace the electromechanical sensor outlined above [41,42,43]. This sensor is based on the relatively new technology of silicon surface micro-machining [44]. The sensor measures the capacitive effect between two microengineered silicon torsion bars. The sensor's output is a stream of pulses whose frequency is proportional to the acceleration measured. The development of this mass produced low cost sensor IC device has created the potential for machines with embedded accelerometers. This would allow the facilities for vibration monitoring based predictive maintenance to be designed into equipment. Chuey *et al.* have presented a design for an intelligent, embedded, bearing vibration monitor based on a micromachined accelerometer [45].

The industrial environment presents many threats to good, consistent vibration measurements. Some of these are illustrated in figure 11. When correctly used the industrial accelerometer offers a degree of immunity to these factors.

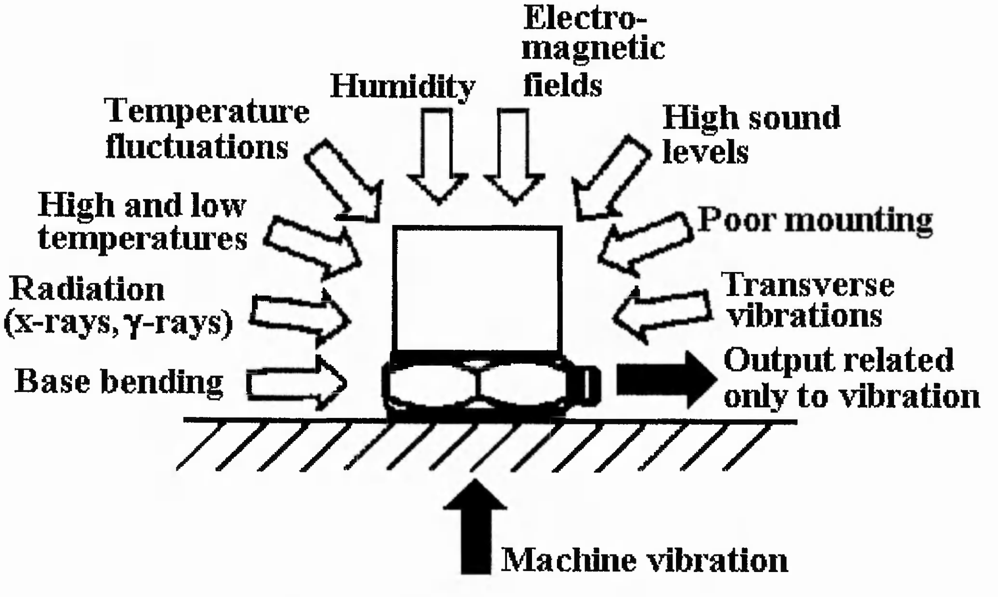


Figure 11. Factors in the industrial environment which can adversely affect accelerometer vibration measurements.

Generally for best results the sensor should be surface mounted at the point of measurement as the frequency response of the accelerometer is greatly affected by the stiffness, K, between its base and the machine [35]. The minimum requirement for repeatable readings is a smooth mounting surface to which the sensor is attached with a strong magnet or adhesive. For machines which require regular frequent monitoring permanent attachment of the sensor via a screw stud is the normal option. Mounting points on bearing housings are usually selected as this is where the basic dynamic loads and forces of the machine are applied and the bearings themselves are critical components. For a complete vibration signature triaxial measurements need to be taken at each location.

Despite their widespread use in industry handheld vibration probes, such as that shown on figure 12, often provide inconsistent vibration measurements. This is due to the variable stiffness coupled with variable location and contact angle whenever measurements are made, figure 13.

Figure 12. Handheld vibration probe and meter.

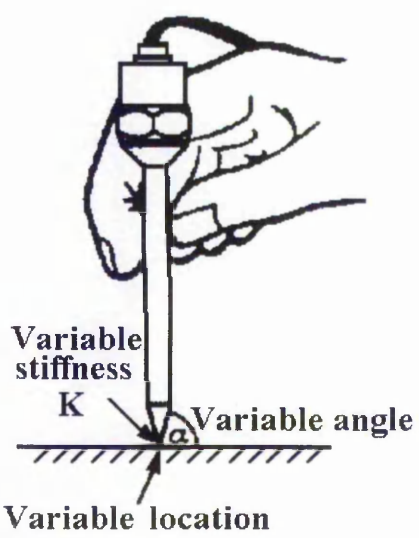


Figure 13. Vibration measurements made with a handheld probe are inconsistent due to the variables involved.

Van Ek *et al.* have presented research into the use and design of portable vibration data collectors [46]. Whilst they are cheaper and more flexible than permanently mounted sensors, there were a number of common problems with their application. Data collection points are normally arranged along a set route followed on each inspection tour. It was found that measurements could be missed out, taken at the wrong locations, or taken in the wrong order. Knowing how many measurement points had been missed out was generally difficult or impossible. These problems combined to produce unreliable data in many of the maintenance programs surveyed. Van Ek *et al.* concluded these problems could be countered by enhancing the functionality of the data collector with a graphical user interface which would provide more guidance to maintenance personnel.

The output signal of a piezoelectric accelerometer is of low power and has to be amplified before being processed. By building integration networks into the amplifier either velocity or displacement, if desired, can be measured from the acceleration signal. It should be understood however that accelerometers make a different type of measurement to the proximity probes commonly used to measure displacement directly. An accelerometer measures the bearing housing vibration at the machinery casing, whilst proximity probes measure the relative vibration between shaft and bearing.

Since the accelerometer has the wider frequency and dynamic range, integration is often used to produce velocity measurements outside the range of velocity sensors. Due to its relatively flat frequency spectrum velocity is considered by some to be the best parameter for evaluating bearing problems [35,47].

Despite its superior performance over most of the frequency spectrum of interest for condition monitoring, using the accelerometer to monitor the frequencies in the range commonly produced by low speed machines is difficult. Low RPM combined with massive structure means that little energy is transmitted to the machine casing. This results in low levels of vibration [48]. The signal lengths required for analysis from a machine operating at low RPM are typically several minutes in duration which makes them difficult to analyse. The measurement of the low level, low frequency vibration from a low speed machine is limited by the electromechanical efficiency of the piezoelectrical material used in the accelerometer [49]. The noise level produced by the accelerometer and its preamplifier overwhelm the low signal generated by the actual vibration.

Low RPM and massive machine structure are just two examples of the adverse conditions in which REB often have to be monitored. Ultra high speeds, difficulty of access and especially the presence of more powerful vibration sources are three other adverse conditions commonly met. In these conditions the ability of current vibration monitoring techniques drops dramatically [37]. The signal to noise ratio can become too low for effective monitoring.

Above the frequency range of accelerometers, stress waves propagate through machine structure at frequencies in excess of 100 kHz. These can provide a clearer indicator of bearing defect formation in environments with a high level of low frequency vibrational noise due to imbalance or structural resonance. These acoustic emissions are monitored by piezoelectric contact probes whose output is proportional to the passage of the stress waves through the body of the sensor [50,51]. Less commonly ultrasonic microphones may be used.

2.2.1.2 Velocity Sensors

Velocity sensors consist of a moving coil mounted inside a permanent magnet. The inductive EMF produced is directly proportional to the velocity of the coil and thus to the machine vibration. The output signal is relatively large and may be integrated to give a measure of displacement. The useful frequency range is in the region of 10 - 1000 Hz and is particularly influenced by the method used to mount the transducer to the machine. As velocity sensors have moving parts they are susceptible to changes in calibration over time making them less suitable for long term measurements. They can also be affected by high temperatures or mechanical damage.

2.2.1.3 Displacement transducers

The displacement transducer produces a signal which is directly proportional, in this context, to the displacement of a shaft relative to its mountings. The IME is a new type of sensor for measuring displacement. Although there are existing sensors based on capacitive, magnetic, and reluctance principles, the type most commonly found in use, the proximity probe, is based on the eddy current principle [4]. A high frequency oscillator sets up eddy currents in the shaft without actually contacting it. As the shaft moves relative to the sensor the eddy current changes, modulating the output voltage. The use of this type of sensor is generally restricted to fluid film bearings which have measurable clearances and allow the rotor system to move around in the journal area [14]. For these machines proximity displacement measurements provide the best vibration information. The damping effect of a fluid film bearing prevents a significant amount of vibration reaching the bearing housing in many situations.

The displacement proximity sensor measures gap and nothing else. This is an important fact to bear in mind when considering the discussion of the IME in the following chapters. In phase work, such as shaft balancing, proximity probes can only define high spots, that is the smallest gap or closest point the shaft comes to the sensor. To obtain an average gap a DC bias or offset is superimposed on an AC analog of the shaft's dynamic motion.

Displacements can be gauged with an error of 2 - 5 μm at speeds of up to 500 m/s [52]. Normally the probe is mounted on the machine housing and this is a source of inaccuracy as for a still shaft the machine housing will be moved by the machine's vibration. The sensor is also susceptible to changes in the shaft finish resulting in differences in electrical conductivity. The probe is unaffected by lubricating fluids.

Figure 14 shows how the axial and rotary movement of a rotor assembly can be monitored by displacement measuring transducers fixed to suitable surfaces. Typically two probes are set at right angles to each other to provide signals from which an X-Y plot of shaft orbital motion can be obtained. A plot of the orbit traced out by a point on the outside of a shaft for which the centre varies in a two-dimensional periodic manner is shown in figure 15.

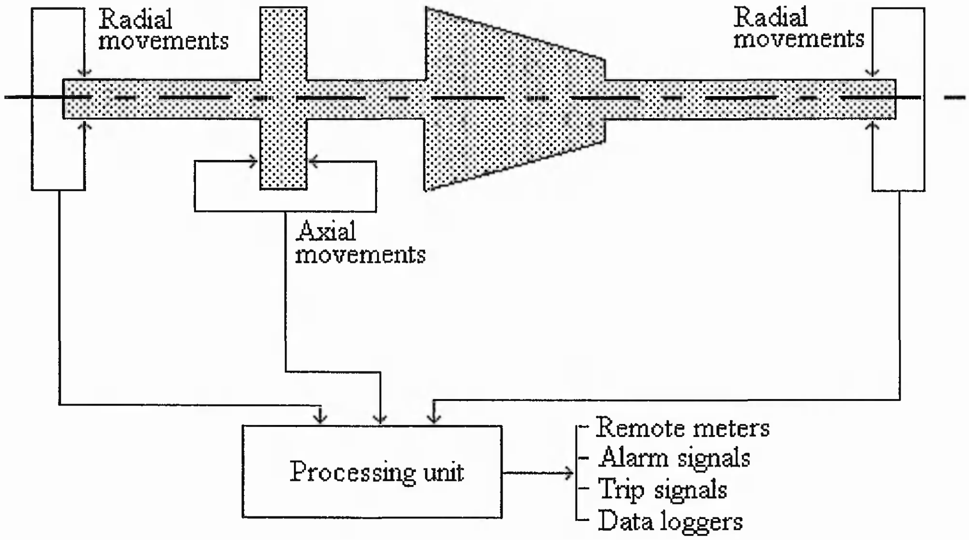


Figure 14. Proximity displacement monitoring system applied to a rotor.

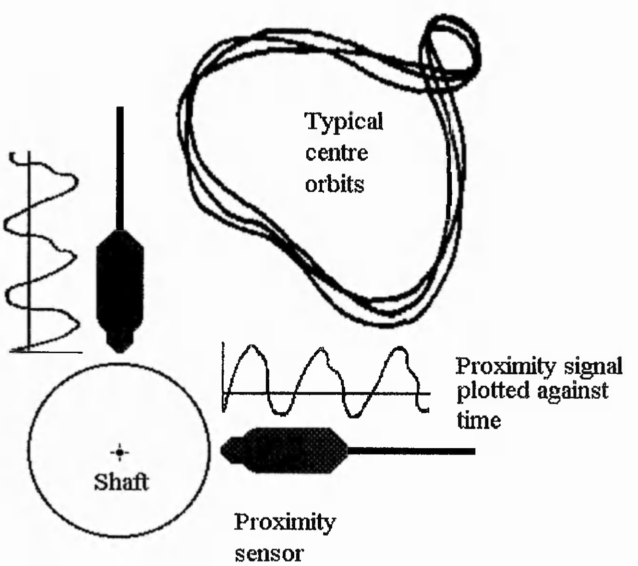


Figure 15. Shaft orbital motion (Lissajous' figure).

Orbital motion provides the basis for much malfunction monitoring of shaft-borne systems as well as providing protection against shaft-bearing friction as clearances close [53]. An extension of the displacement monitor is a facility for vibration monitoring by which the peak-to-peak displacement signal from the transducer amplifier is shifted to a reference level above which the amplitude is sensed.

Displacement measurement has been applied to REB by using the outer ring of the bearing as the target rather than the shaft. This method was first demonstrated by Philips who used a fibre optic displacement sensor to detect micron deflections of the bearing outer race [54]. The displacement sensor was mounted in the bearing housing and proved capable of detecting defects on the bearing elements. This experimental technique was developed commercially by the Bentley Nevada Corporation. As an alternative to the fibre optic sensor they developed a very high gain eddy current proximity sensor. This had the advantage of being an industrially proven technology not susceptible to contamination. The system is called REBAM, which is an acronym for rolling element bearing activity monitor. Figure 16 shows how the eddy current sensor is mounted in the bearing housing.

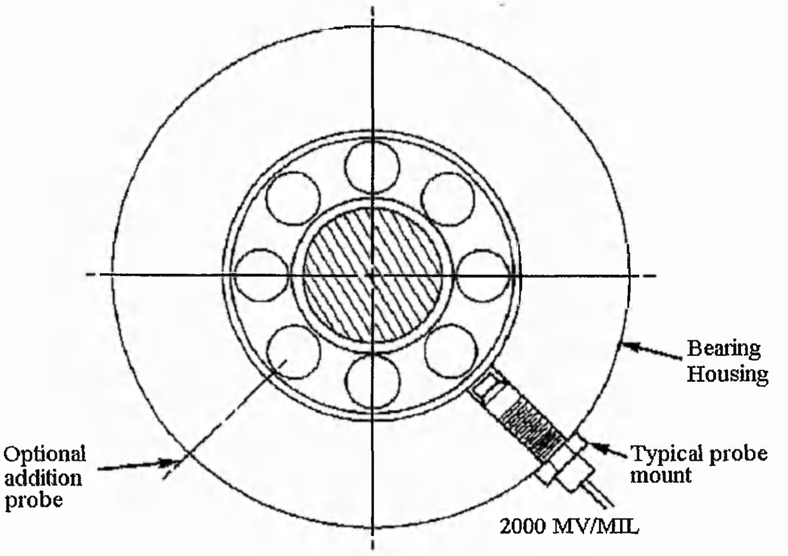


Figure 16. Typical mounting position for the REBAM eddy current proximity sensor.

The REBAM system was tested successfully in the laboratory and in a variety of industrial environments [55]. Spencer and Hansen reported that this method possessed significant advantages over other techniques used for REB monitoring [56]. They tested it against casing mounted accelerometers, acoustic emission and velocity sensors. The eddy current probe was found to be able to detect bearing damage which the other systems could not.

The main disadvantage of the REBAM system is that the sensor has to be specially mounted within the bearing housing. This means that it can only be applied to machines where the housing can be drilled and tapped in the load zone to accommodate the sensor. The bearing outer ring is also weakened by the lack of support at the sensor location. The gap between the outer ring and the sensor is an important parameter of the system, but this is subject to change due to vibration and thermal expansion [57].

The proven advantages of the REBAM system are significant in considering the application of the IME to bearing condition monitoring. The validity of the idea of using displacement to monitor the forces developed in a bearing directly is confirmed by the experimental work that has been done with the REBAM system. By looking at the shaft's response to these forces, rather than their effect on the bearing's outer race the IME avoids the problems with installation outlined above.

2.2.2 The importance of proximity monitoring

Some of the leading arguments for the preference of proximity monitoring of shaft displacement over acceleration or velocity sensing were made by Collacott as follows [4]:

1. Most harmful radial dynamic forces originate with the rotor, or with the rotor system including its restraints, compliances and damping. A bearing housing on which vibration is measured is a secondary system, vibration is transmitted to it via the bearing system and this vibration is rarely the same as the rotor vibration. The differences lie in amplitude, phase and frequency content.
2. Vibration waveforms indicative of machine performance are to be observed primarily on the rotor.
3. Experience suggests that machine cases simply do not manifest waveforms of direct significance to rotor dynamics with sufficient fidelity for dependable protection calculations, and certainly not for analysis.
4. Shaft displacement increases with bearing deterioration but vibration velocities measured on the bearing housing are not always sensitive to this increase.
5. Vibration velocity measurements depend largely on correct location of the sensor; proximity monitoring presents no such difficulties.

2.3 Vibration monitoring of rolling element bearings

2.3.1 Bearing vibrational characteristics

Central to current bearing health monitoring techniques is the idea that the relative health of a rolling element bearing can be ascertained by examining the induced vibrations at the surface of the bearing housing. These vibrations will result from changing forces inside the bearing and from the associated machinery. Ball and roller bearings usually follow a characteristic wear / failure pattern [58]. When a normal bearing is properly installed and lubricated, the main source of vibration will be the contacts between the rolling elements and the bearing races. Surface finish errors produce a roughening of the contact surfaces of the bearing. They produce vibrations of relatively low amplitude zero-mean Gaussian noise. Geometrical imperfections in bearings result from errors in manufacturing. Varying diameter rolling elements and races with a waviness of a few cycles around the circumference produce low frequency periodic vibrations. The work of Su *et al.* in studying the frequency characteristics of normal bearing vibrations has shown that geometrical errors in a normal bearing under certain conditions can produce frequency spectra indistinguishable from those of a damaged bearing [59]. These two sources of vibration combine to excite the machine structure giving rise to the normal vibration due to the running of a healthy bearing.

As a bearing ages, the vibration level increases slightly due to roughening of the contact surfaces. Hertzian contact stress is developed by a loaded rolling element in contact with the inner and outer races. The fatigue loading created by the motion of the rolling element induces the nucleation and growth of subsurface cracks. These cracks extend incrementally during successive load cycles. Eventually these subsurface cracks join and grow to reach the surface of the race. This results in a loss of material from the race which is known as a spall. This process is accelerated by improper installation, shaft imbalance, lubrication failure, excessive loading and contaminants. Following the initial spalling, wear accumulates rapidly due to the damage caused by the freed particles and due to rolling surfaces impacting the spalls.

Single spalls or point defects produce cyclic impulses each time the defect contacts another surface. The period of these impulses is a function of the defect location, bearing geometry and shaft speed. Equations have been developed which relate these factors and can be used to separately detect specific defects within the bearing on the inner and outer races or the rolling elements. Figure 17 shows the nomenclature used for REB components. Figure 18 illustrates bearing geometry.

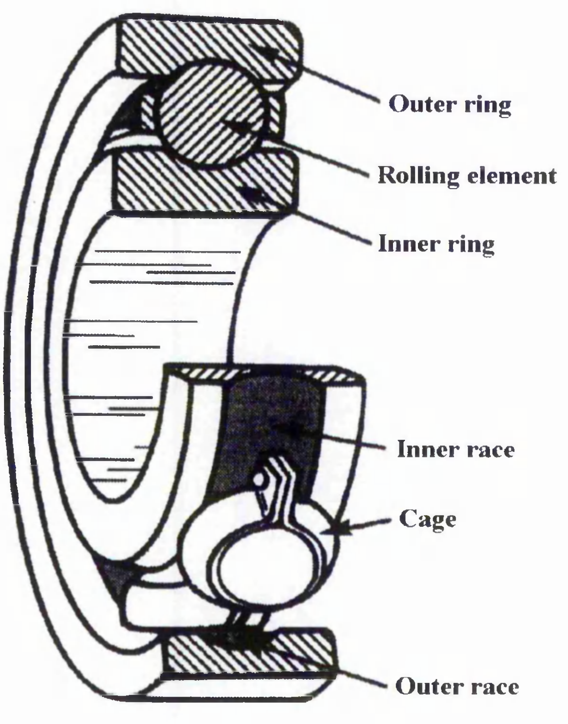


Figure 17. Bearing nomenclature.

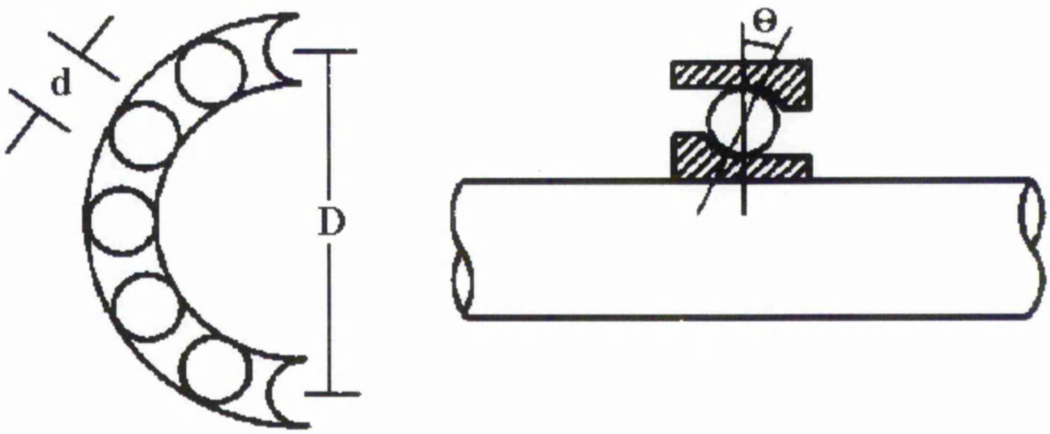


Figure 18. Geometry of a typical rolling element bearing: D = bearing pitch diameter, d = rolling element diameter, N = number of balls or rollers, θ = contact angle.

The following equations identify each of the bearing defect frequencies together with some unique relationships between them. They assume that the inner race is rotating with the shaft and the outer race is fixed [47]. These frequencies are only approximate as they will be affected by slipping of the elements within the bearings and spinning of the races on the shaft or in the housing. Slip is believed to be relatively unimportant under normal conditions, however at high speeds and under light loads slip may be as high as 50% [60].

$$\textit{Ball spin frequency} = \frac{D}{2d}\left[1 - \left(\frac{d}{D}\cos\theta\right)^2\right] x\ rpm \tag{1}$$

$$\textit{Inner race defect frequency} = \frac{N}{2}\left(1 + \frac{d}{D}\cos\theta\right) x\ rpm \tag{2}$$

$$\textit{Outer race defect frequency} = \frac{N}{2}\left(1 - \frac{d}{D}\cos\theta\right) x\ rpm = N\ x\ FTF \tag{3}$$

$$\textit{Fundemental Train Frequency} = \frac{1}{2}\left(1 - \frac{d}{D}\cos\theta\right) x\ rpm \approx 0.4\ x\ rpm \tag{4}$$

Impulses due to defects on the bearing surfaces propagate through the machine structure. The vibration measured at the machine casing or bearing housing is a result of the excitation of the machine structure by these impulses combined with that from other machine components. The signal that is detected is modified by the transmission path from the defect location to the point of measurement dependant on the natural frequencies of the bearing parts and their supports.

2.3.2 Bearing vibration analysis methods

The following sections briefly outline the main methods currently used for the analysis of vibration signals for the detection of bearing defects. This section is based on reviews by Mathew and Alfredson [61], Tranter [6] and Martin [17].

2.3.2.1 Overall level (rms) measurement

This is a simple and inexpensive measurement to take. The overall level of machine vibration is compared either to a manufacturer's recommended value or a value recorded when the machine is known to be in good health. Any increase is taken to indicate the development of a fault. Using both accelerometers and velocity probes it is possible to determine whether the increase in vibration is low or high frequency in nature. The greatest limitation of this approach is the lack of sensitivity and information in the data. Unless a bearing problem is severe, the overall level measurements may not change significantly. Conversely other sources of vibration in the machine can cause the overall vibration level to rise. For this reason this technique is rarely used for the early detection of bearing defects.

2.3.2.2 Peak level detection

As an alternative to the rms level of the vibration, the peak level of the signal can be trended. The rationale for this is that it is the impulsive component of the signal that will increase due to bearing damage. This method is not reliable on its own as, again, other vibration sources in the machine can also affect the peak level of the signal.

2.3.2.3 Crest factor

In the time domain the vibration waveform due to a healthy bearing is mostly random. As bearing defects develop the impulses they generate increase the peak level of the signal. The crest factor is the ratio of the peak level to the rms level. This is one of the dimensionless parameters that is claimed to be relatively insensitive to bearing speed and load [61].

2.3.2.4 Shock pulse

The shock pulse method detects development of a mechanical shock wave caused by increasing bearing damage. It is actually a measure of the level at the bearing resonance ~32 kHz. This method has been used widely, although concern has been expressed that the shock pulse reading can decrease in later stages of bearing damage when the impulsiveness of the signal decreases. Other conditions such as turbulence and cavitation in pumps can act as a source of noise which can lead to false readings.

2.3.2.5 Spike energy

This measurement, patented by IRD Mechanalysis, is based on the detection of high frequency peak acceleration [6]. It uses a circuit to reject vibration due to low frequency sources which highlights the peak level due to the excitation of machine resonances.

2.3.2.6 Kurtosis

This is a statistical parameter, derived from the statistical moments of the probability density function of the vibration signal. A bearing in good condition has a gaussian distribution of accelerations. With damage there is a relative increase in the high levels of acceleration, i.e. the tails of the distribution become more dominant [4]. Kurtosis has the major advantage that the calculated value remains constant for an undamaged bearing irrespective of load and speed, yet changes with damage [62]. The extent of the damage can be assessed from the distribution of this statistical parameter in selected frequency ranges. Experimental studies over bearing life durations have shown that while kurtosis can identify defective bearings, it cannot indicate incipient damage effectively [61]. This technique will be explained in more detail in chapter 5 as it has proven to be valuable in the interpretation of the IME signal.

2.3.2.7 The High Frequency Resonance Technique (HFRT)

One major problem in attempting to monitor bearing damage is that more dominant low frequency signals, from other machine elements, can effectively mask the bearing frequencies. The HFRT is based on selecting a narrow band signal, band passed filtered from the broad band overall vibration [60]. This narrow band is centred on a structural resonance which is excited by the impulses from any bearing defect at a much higher frequency than the vibration generated by other machine elements. This resonance could be the vibration of the elements between the races, the vibration of the races themselves in ring mode, the vibration of the bearing housing or machine structure, vibration in the transducer or a combination of these. The resonance frequency is amplitude modulated at the characteristic defect frequency at which the impulses occur. The resonance acts effectively as a carrier frequency. The signal centred on the resonance frequency is half wave rectified and then demodulated to reveal the presence of the defect frequency.

2.3.2.8 Spectral analysis

Once the vibration signal from a machine has been collected by a sensor those components of it due to the bearings must be isolated so an evaluation of the bearing condition can be made. A frequency spectrum is derived from the vibration waveform by performing a fast fourier transform (FFT). From this spectrum the peaks due to the calculated characteristic defect frequencies can be examined [63]. Diagnosis is, however, rarely this simple for all but the most simple machines. In practice the spectrum will contain a number of peaks which may represent bearing conditions and others which are related to other sources of vibration. As discussed previously the signal may also not accurately represent the condition of the bearing completely due to its mounting position, the signal path, etc. Use of the correct model to explain the spectrum and its peaks is central to the successful use of spectral analysis. McFadden and Smith have defined a model to explain the appearance of vibration frequency spectra produced by bearing defects [60,64,65]. This has done a great deal to contribute to the reliability of the HFRT and other vibration techniques. Parts of this model will be explained in detail in chapter 5. The work of McFadden and Smith was used as a basis for the development of the model used to interpret and explain the form of the IME signal.

Once the frequency spectrum of the vibration has been calculated there are various methods commonly used to aid further analysis. Waterfall plots, where successive spectra are represented in three dimensions with time as the third, are used to identify changes over time by eye. They can be difficult to interpret and it is hard to quantify the differences in successive spectra. The spectrum of the logarithm of the power spectrum is called the cepstrum. This is used to highlight periodicities in the power spectrum, summing all the harmonics into one peak in the cepstrum. This allows simplified trending and identification of specific fault frequencies. Difference spectra may be calculated by subtracting a baseline spectrum representing the machine in a healthy state from later spectra. Changes in level at frequencies of interest can then be easily identified. Since the frequencies at which bearing defects appear in the spectra are dependent on the shaft speed this technique does not work well when the shaft speed changes between measurements.

Chapter 3

The Experimental Incremental Motion Encoder System

3. The experimental incremental motion encoder system

The basis of this research work has been the application of an experimental incremental motion encoder system to bearing condition monitoring experiments. At the start of the project no operational IME device or bearing condition monitoring system existed. The project therefore began with the requirement to build an IME device and then design and implement the novel hardware and software of a bearing condition monitoring system based upon it. This chapter details the essential first phase of the project in which an experimental IME sensor was constructed based on the work by Orton [1,2] and the design of the experimental system was developed. The chapter begins with an explanation of the theoretical basis of the IME from Orton's work and other sources. The experimental IME system developed during the project is then described in successive sections. This system is shown as a block diagram in figure 19. It consists of the IME device which actually senses motion, the interface circuit and computers which capture the IME signal, and the computer which processes, interprets and displays the signal.

3.1 The incremental motion encoder

3.1.1 High precision angular measurement with incremental encoders

The incremental motion encoder is based on a measurement principle that is used in conventional high precision incremental encoders to correct for the eccentricity of the disk centre. For high precision angular measurement applications the measuring step given by the actual grating line spacing of an encoder disk is insufficient. It is not technically feasible to use an optical grating with the line spacing necessary to give resolutions of 0.001° or better. For these applications a process of interpolation is used to subdivide the spacing between the physical grating lines [66,67]. This may be an analogue or a digital process

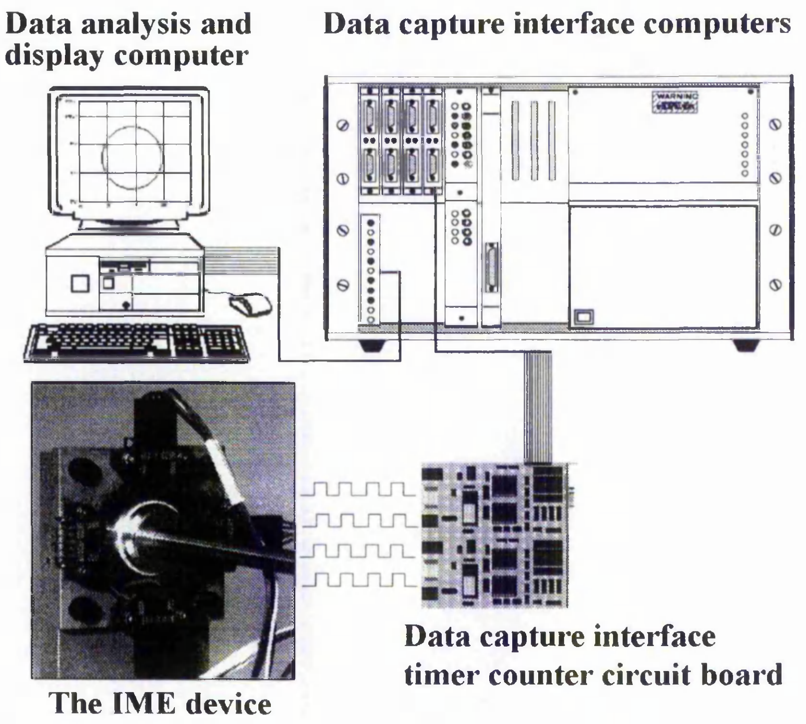


Figure 19. Block diagram of the experimental IME system developed during this project.

depending on the accuracy required, but in both cases the cycle of the voltage signal from the read head due to the passage of a single grating line is subdivided by using voltage levels. In an analogue system resistor networks are used to produce additional phase shifted signals by vectorial addition. An 18° phase shift produces a 5x interpolation of the original signal. In this case the read head output signals have five times the frequency of the input signals and twenty edges are produced from each grating line. Digital interpolation allows for the highest possible subdivision of the grating line spacing. The analogue signal from the read head is converted into digital voltage values and a microprocessor is used as an arc-tangent calculator to pick an interpolation value from a look-up table. Interpolation of 1024x on a 36000 line grating allows a measuring step of 0.00001° to be achieved.

Due to this high degree of angular accuracy, high precision angular encoders are particularly susceptible to shaft axial eccentricity as a source of error. This axial eccentricity may be due either to the motion of the encoder shaft relative to the read heads in the plane of the encoder disk or the error in mounting the disk to the encoder shaft. It is not possible to achieve absolutely perfect centring of the graduations of the encoder disk to the centre of rotation of the encoder shaft. A typical minimum value for this eccentricity given by Heidenhain, a manufacturer of high precision encoders, is 1μm when a precision mounting device is used. Figure 20 shows how this eccentricity of the encoder graduations to the centre of rotation produces an error in the angular measurement made by the encoder. The angle measured by the encoder θ' differs from the actual angle the shaft has rotated through θ, by an angular error $\Delta\theta$. This angular error is proportional to the eccentricity e.

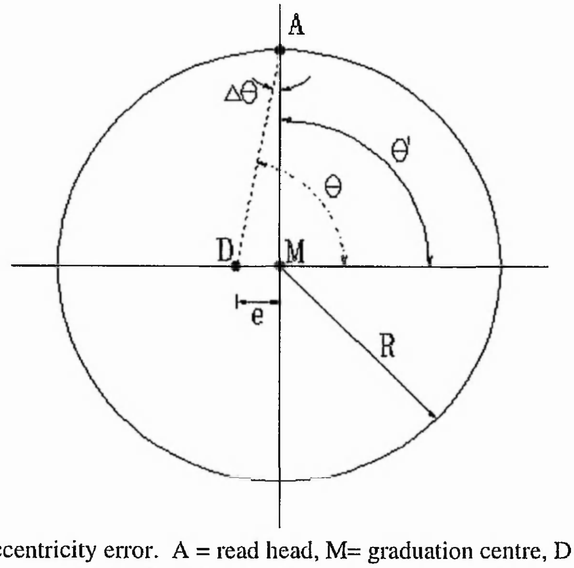


Figure 20. The eccentricity error. A = read head, M= graduation centre, D = axis of rotation or the shaft, e = eccentricity, θ = "true" angle, θ' = measured angle, $\Delta\theta$ = measurement error.

In order to correct for this error, high precision encoders use two read heads in diametrical opposition. The correct angle of rotation can be calculated from the angle measured at both read heads. Digital interpolation provides for the highest possible angular accuracy and so four read heads are used in two pairs. The corrected values from each pair are averaged to give a single output measurement. Figure 21 shows an exaggerated example of how the error correction calculation is made. The angle measured by read head A is 90° and that measured at read head B is 70°. The actual angle of rotation is the average of these two values, 80°.

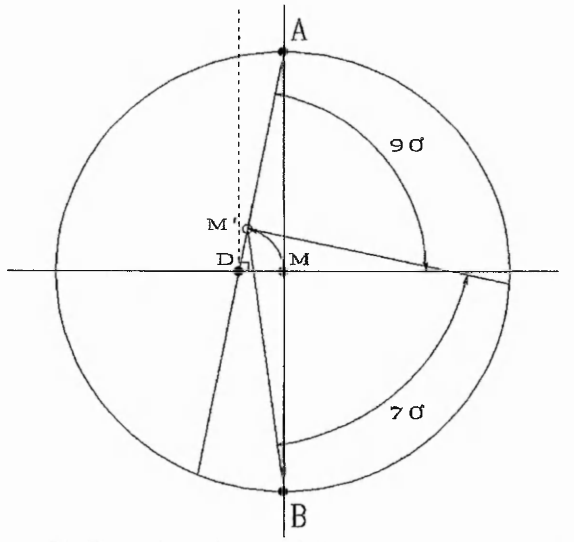


Figure 21. Correction of eccentricity error with two read heads.

The value for $\Delta\theta$ that is found using this method is discarded after the angular measurement has been corrected. The eccentricity of the disk centre could be found by applying equation 5, but since it is angular accuracy that is required this measurement is not made.

$$e = R \tan \Delta\Theta \tag{5}$$

In the above example the angular error $\Delta\theta$ is 10°. In practice the angular error of a high precision encoder with a 1μm eccentricity will range between ±1.6″ and ±5.1″ dependent on the radius of the graduations on the disk [67]. Equation 6 shows the relationship between the graduation radius R in mm, the eccentricity e in μm and the angular error $\Delta\theta$ in arc seconds.

$$\Delta\Theta = \pm 412 \times \frac{e}{2R} \tag{6}$$

The fact that such small angular differences are involved in making the angular measurement correction limits this method of ensuring angular accuracy to high precision encoders. Using interpolation these devices can make the high resolution measurements required. The low resolution encoders commonly used for machine speed control or positioning are unable to make such measurements using conventional methods. The incremental motion encoder is based on a method which allows high precision interpolation of a low resolution encoder disk. This allows measurement of shaft axial eccentricity and hence the motion of the shaft.

3.1.2 Time base interpolation

The most common approach to the use of incremental encoders for the measurement of angular velocity is for the number of encoder signal cycles to be counted in a series of equal time periods, figure 22a. The measurement of angular distance travelled by the encoder is accurate to ±1 encoder grating interval in each time period. When precision measurement of angular velocity is required a method known as reciprocal timing is used [68]. Using this method the time taken for the encoder to travel through the equal angular interval represented by the grating line spacing is measured. This is achieved by counting the pulses

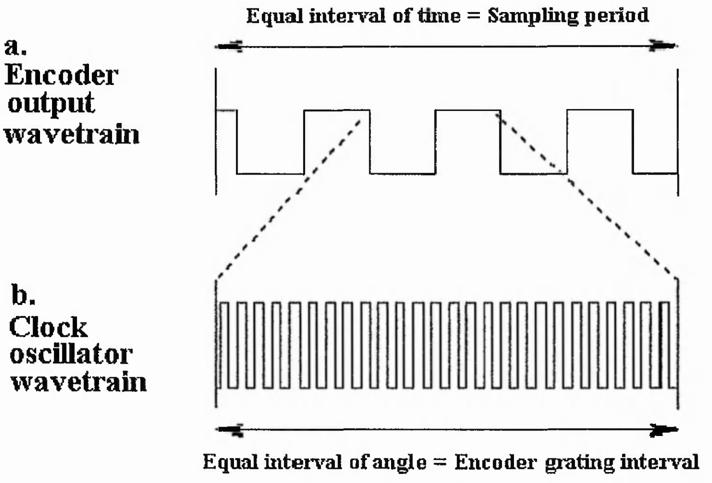


Figure 22. Conventional vs. reciprocal method of measuring angular velocity.

of a clock oscillator between the edges of each successive cycle of the square wave encoder signal, figure 22b. This measurement of time will be accurate to ±1 clock cycle.

Extensive use of the reciprocal method for the precise measurement of angular velocity during grinding processes led Orton to formulate the idea which is the basis of the measurement accuracy of the IME [1]. He proposed that by using the subdivision of the encoder grating line interval provided by a clock oscillator, an encoder with a relatively low resolution disk could make measurements of very small differences in angle. These would be equivalent to those that could be made with a disk of a much higher physical resolution. The term 'time base interpolation' is now used by the author to describe this method and its application to the IME. In theory, the size of the smallest interpolation step is limited only by the frequency of the clock oscillator. In practice the ability of electronics to count high frequency edges acts as the limiting factor. The number of interpolation steps, N, is given by equation 7.

$$N = \frac{F_{tb}}{rps \times G} \tag{7}$$

F_{tb} = frequency of the time base, rps = rotations per second of the encoder disk,

G = number of grating lines on the encoder

Angular rotation cannot be measured with any more than the actual resolution of the encoder disk using time base interpolation. The high degree of angular accuracy provided by this method can only be used to measure the angular measurement error $\Delta\theta$.

Figure 23 illustrates how the IME measures $\Delta\theta$ and uses this to find horizontal displacement. A single clock oscillator provides a common timebase for a pair of diametrically opposed encoder read heads at positions A and B, distance L apart. When the centre line of the disk coincides with the centre line joining the read heads the difference between the angular position measured at A and that at B is equal to π. In this case the time difference between successive edges of the encoder square waves will be zero. The displacement of the disk centre with respect to the centre line of the read heads by horizontal distance d leads to a difference in angular position measurement between A and B of θ radians. The angular measurement error, $\Delta\theta$, will be given by equation 8. The time difference, Δt, which will be recorded between the encoder square wave output signals from read heads A and B is proportional to $\Delta\theta$.

$$\Delta\theta = (\pi - \theta) \tag{8}$$

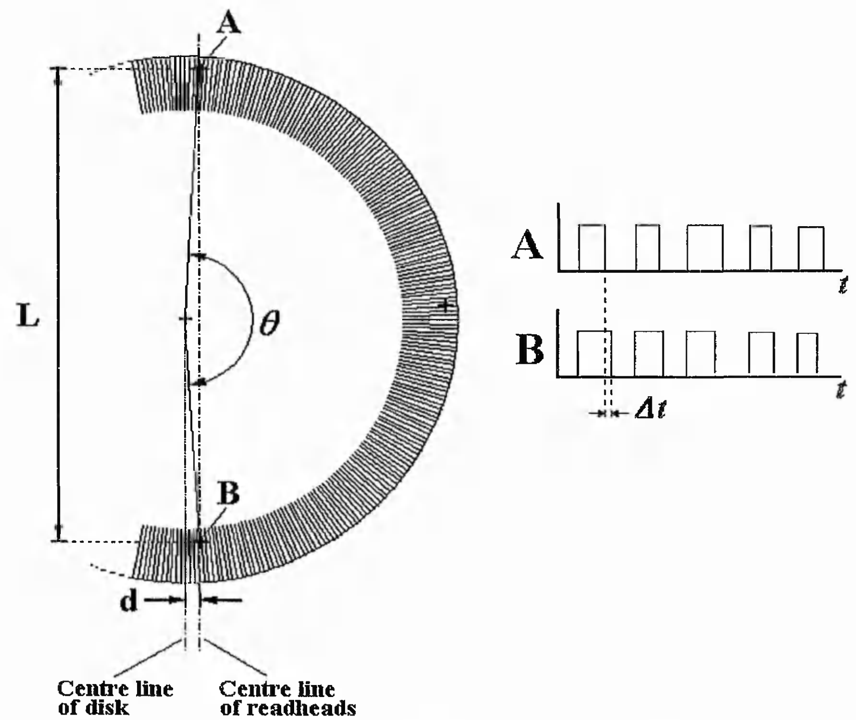


Figure 23. Measurement of horizontal displacement.

Since for small angles measured in radians the tangent of an angle is approximately equal to the angle itself, the value of *d* can be determined by substituting equation 8 into equation 5 and rearranging to account for the diameter of the disk rather than the radius. The displacement of the shaft centre from the centre line of the measuring read heads, *d*, is therefore given by equation 9.

$$d = \frac{(\pi - \theta)L}{4} \tag{9}$$

This is an approximation which can be made for centre displacements of the order of tens of microns if the disk is of the order of tens of millimetres in diameter. Each pair of diametrically opposed read heads calculates the position of the shaft centre in a single dimension. Figure 24 shows the read head configuration of an IME device constructed with four read heads, marked A - D, spaced at 90° intervals around the circumference of an optical grating disk. Together the angular measurements from the two read head pairs allow the shaft centre, shown at X, to be located in two dimensions relative to the intersection of the read head pair centre lines, shown at X'.

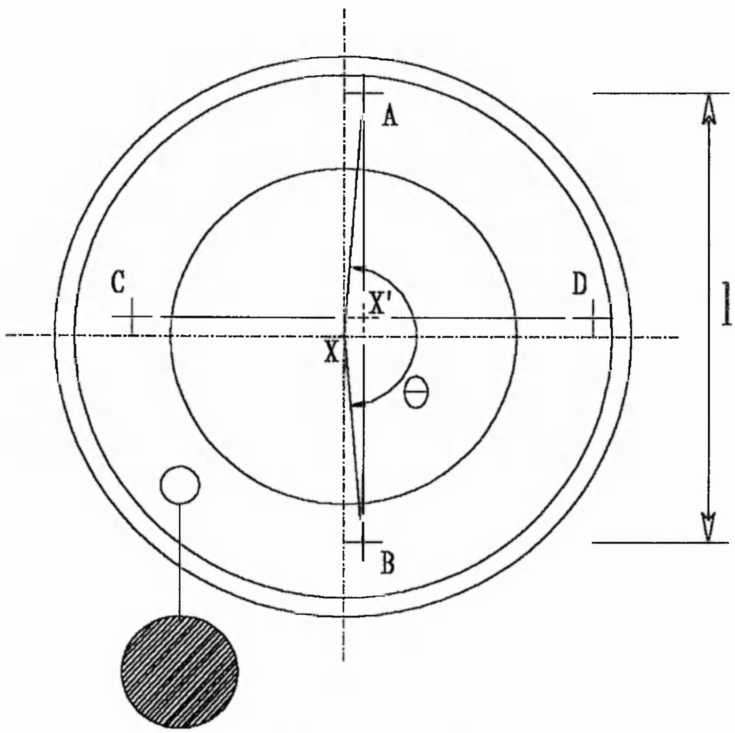


Figure 24. Shaft centre position calculation.

The actual calculation is performed within the IME's computer system based on the timing values recorded at each read head. The method by which these values are collected is described in section 3.2 which describes the hardware of the experimental IME system.

Routson *et al.* [69] and Wang *et al.* [70] have conducted research into the precise measurement of angular motion required for the analysis of torsional vibration using optical gratings. Their work was based on using a method of time base interpolation similar to that of the IME, but applied to coarse resolution reflective strips. Their interest was restricted to angular motion and although they calculated the shaft centre position, this was only used to correct the angular measurements in the same manner as that described for a precision encoder. Within their work the placing of the read heads around the drum grating was considered in detail. They give an extensive derivation of a set of equations for the calculation of shaft centre position from multiple read heads [69]. These complement the work by Orton for the IME in [2] showing how three read heads at 120° could perform the centre position calculation.

The fact that the basic principle of time base interpolation has already been validated by other researchers pursuing different applications related to rotating machines meant that within this project no attempt was made to re-examine this fundamental area.

Other research in the field of torsional vibration measurement has also made use of the reciprocal timing method for the accurate measurement of angular motion. Lenaghan [71] describes using two strips of photographic film with twenty clear and opaque zones mounted on reflective backing and wrapped around the crank shaft of an engine. A 2 MHz oscillator is used with two counters and a computer to measure the twist angle of the shaft by accurately measuring the angular difference between each strip. Citron [72] developed a similar system using electromagnetic sensors and toothed wheels. Neither of these projects used multiple read heads.

3.1.3 IME applications

At the start of this project a wide range of applications were envisaged for sensors based on the IME principle and implemented with various technologies. Figure 25 presents these in the form of a tree diagram. The IME is by no means limited to the condition monitoring of rolling element bearings, but as this project was limited as to what could realistically be done it was necessary to select one main direction for the research and direct most emphasis and all practical work into this. The aim of the approach was to prove the operating principle in one area sufficiently for it to be extrapolated into others.

3.1.4 The experimental IME device

A sensor based on the IME principle can be implemented using a variety of optical or electromagnetic technologies in such a way as to give a square wave output from a marked, rotating, surface. Within the scope of this project the concentration has been solely on the use of optical devices. This has been partly due to the extensive use of optical encoders as part of the research that the IME grew out of. There was also the practical consideration that non-optical components fabricated in such a way that they could be engineered to suit this application were not available within the project's limited resources.

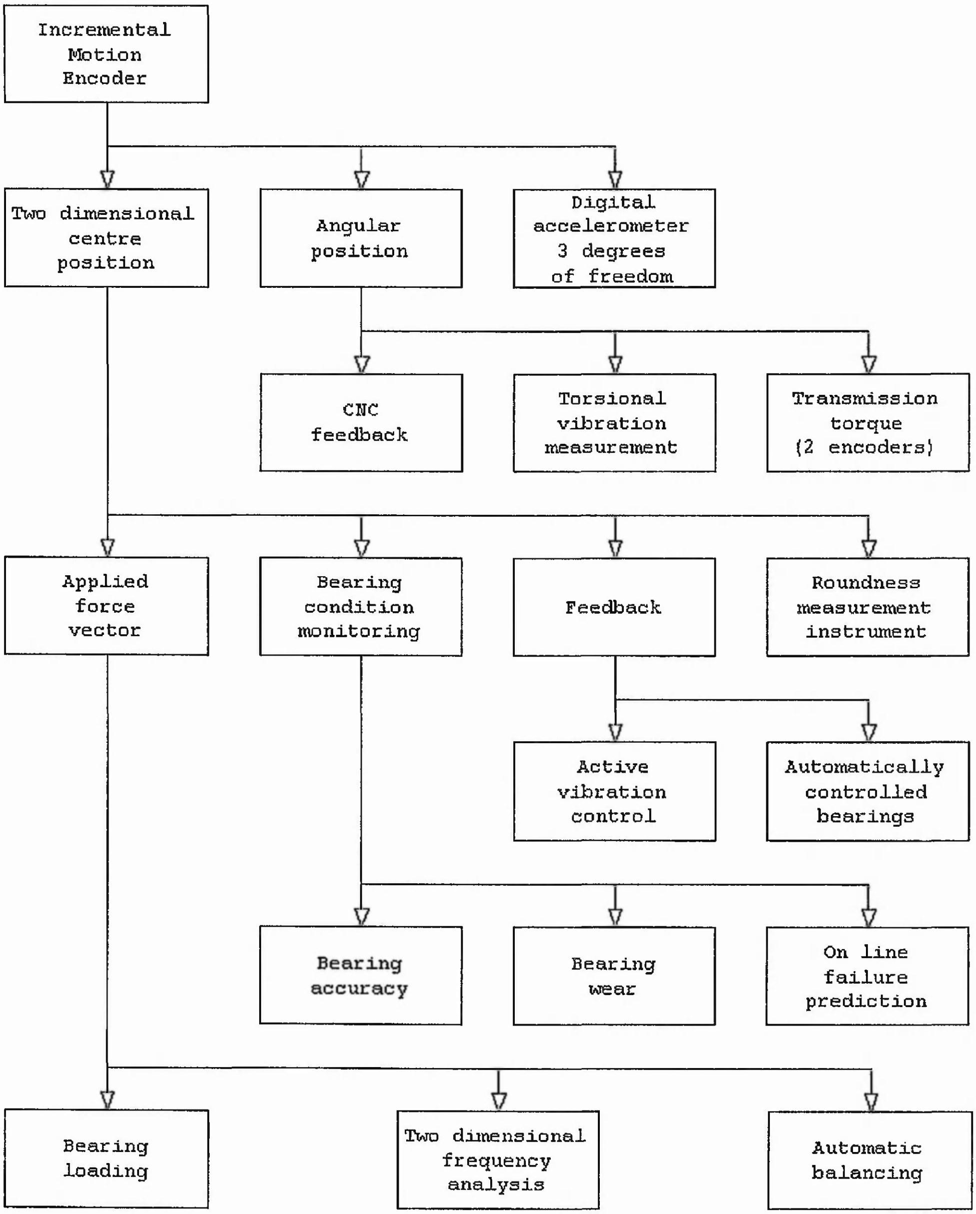


Figure 25. Application areas for IME technology.

Before the start of this research the only attempt at creating an IME device had been the encoder used in the preceding MSc [3]. This one off adaptation of a standard encoder device had been able to satisfy the limited requirements of the MSc project, but experimental work could not be undertaken with it. Its construction was both too solid and too intricate to allow the type of work that was envisaged, which would have required the disassembly of the device and the remounting of its disc and read heads around a moveable shaft. This encoder was never used for measurement experiments. Instead its output signal was used extensively during the first nine months of the project to test the hardware and software of the data capture interface as this was developed. These tests are detailed in the sections that deal with the development of the data capture interface.

Three options were considered for the creation of the experimental IME device. An existing encoder could have been purchased and modified, an IME could have been manufactured especially for the project or an IME could have been built from kit encoder parts. Existing encoders examined lacked the flexibility to be used for experimental work as the disk could not be removed from the encoder housing. Although some had two read heads in opposition, none had the full four read head configuration required. The expense of a bespoke encoder prohibited the consideration of this option in the initial phase of the project, although an encoder using fibre optics was later built as a collaboration with Muirhead Vactric Ltd. It was decided to construct the experimental IME from kit encoder parts as these offered the greatest flexibility for the design and subsequent application of the device.

The design of the fibre optic encoder built by Muirhead Vactric Ltd. is described in appendix D. This device was never used experimentally due to its low resolution and problems with its electronics.

The IME device that was built for this research project was constructed from standard optical encoder kit parts of the type commonly used for industrial motion control. These were obtained as a set of development samples from Hewlett Packard Components Limited. The components used were a HEDS-6140 OPT J00 code wheel (1024 counts per turn) and four HEDS-9040 OPT J13 photoelectric read heads, together with connectors and the associated mounting tool. Full technical specifications are given in [73]. The experimental encoder is shown in figure 26. This is the device used for all the experimental work on which this thesis is based.

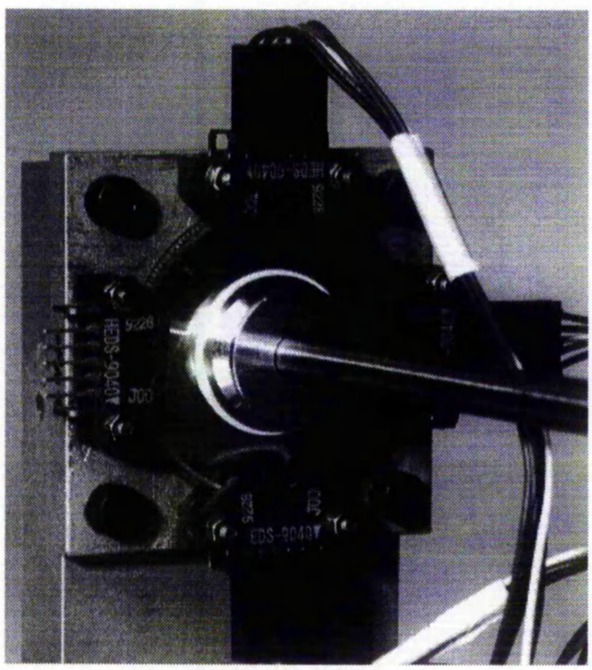


Figure 26. Experimental incremental motion encoder showing the configuration of encoder disk and read heads used to sense the motion of the shaft centre. Shaft dia. 8mm.

The experimental encoder was mounted on a test rig designed by the author specifically for this research. The design was made to satisfy two main objectives. Firstly, to allow the mounting of the four readheads and the encoder disc to embody the complete two dimensional incremental motion encoder. Secondly it had to permit the planned experimental monitoring of rolling element bearings to be conducted using the mounted heads. This meant the rig had to allow various bearings to be mounted and removed from the shaft supporting the encoder disc. The design of the rig was influenced by work on experimental bearing condition monitoring read during the literature survey [28,62,74]. In particular a paper by Ford [75], which describes in detail the design criteria for such an experimental rig, rather than simply describing its use, was influential on the design of the bearing changing mechanism. The rig was built by the university's mechanical engineering workshop from the author's design drawings.

The size of the rig was dictated by the dimensions of the Hewlett Packard (HP) components, particularly the code wheel which required an 8mm shaft size. This shaft size was not ideal for the rig. The loading that could be applied without the shaft bending significantly was much less than rated working load of the bearings used in the rig. The principal limitation of the small shaft size was the inability of the rig to test bearings to destruction under high loads. Experiments were conducted with induced bearing defects because of this.

The part of the rig shown in figure 26 is the bearing test pillar supporting the encoder mounting plate. The mounting plate shown carries the four HP readheads. The rig design uses the encoder mounting plate to allow different read head types to be mounted on the rig. Using the common dimensions of the four black corner screws, any other design of encoder can be constructed around a similar base plate. This was the case with the fibre optic based encoder produced by Muirhead Vactric.

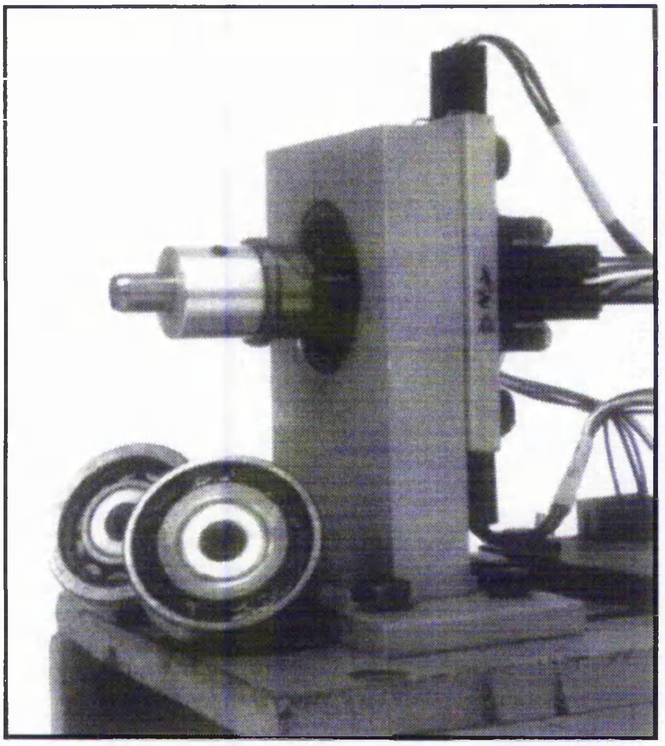


Figure 27. The bearing test pillar with the inner race of a bearing on the shaft.

The bearings chosen for experimental work were those of the roller and angular contact types. Deep groove ball bearings were not used as they did not allow the disassembly required to induce defects on the bearing surfaces. The bearings chosen had a 11mm bore diameter which allowed a collar to be used to fit the inner race to the shaft so that the bearing could be easily mounted and removed from the shaft. Figure 27 shows the reverse side of the bearing test pillar to figure 26. The inner race of a bearing is shown withdrawn on its supporting collar. The pillar houses the bearing in the recess shown secured by a single screw, not visible in the figure. A rubber washer is used to prevent this single pressure point from distorting the outer race.

The whole of the experimental rig is shown in figure 28. The bearing in the motor end pillar which supports the far end of the shaft has a 8mm bore and is therefore significantly

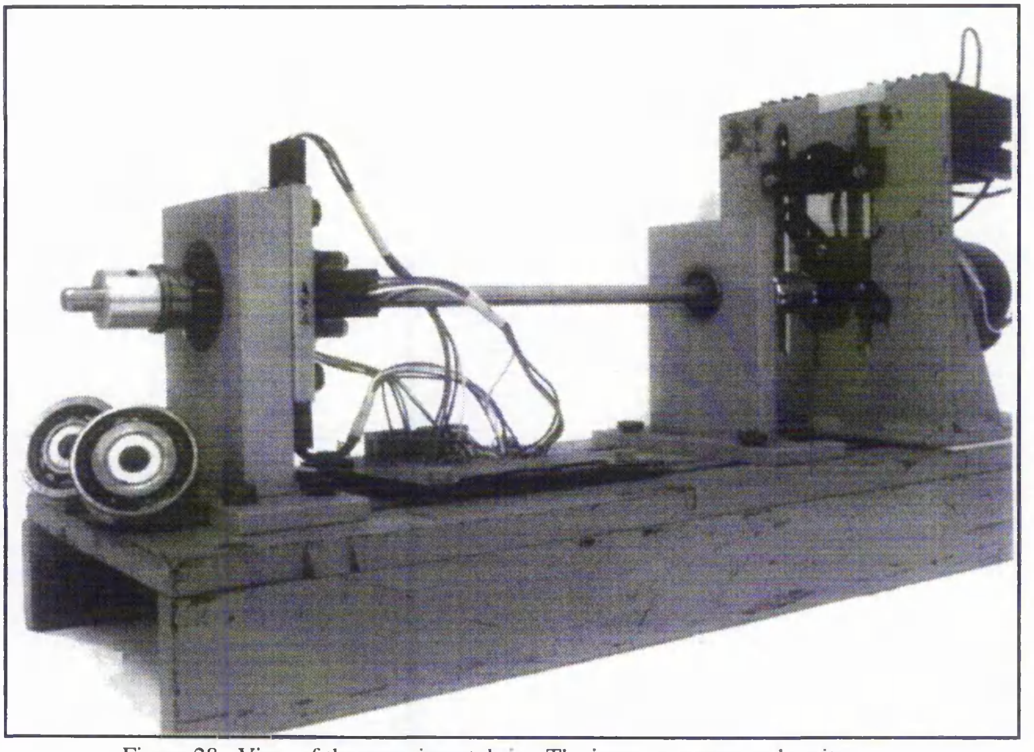


Figure 28. View of the experimental rig. The inner race mounted on its supporting collar is shown withdrawn from the bearing in the test pillar

smaller than the bearing used for testing. This difference in size is deliberate and is designed to limit the effect any vibrational frequencies from this bearing will have on the measurements taken from the sensor at the bearing test pillar end of the shaft.

The encoder read heads are shown connected to the circuit board which the holds the pull up resistors for the read head outputs. Each read head has a once per revolution channel as well as two quadrature outputs. Only one head has its once per revolution signal connected through to the system, the other three are grounded. This signal is used to synchronise the start of data collection within the monitoring system. The second quadrature output of each read head is also grounded. The encoder signals reach the data capture interface via a ribbon cable which connects to the box connector on the circuit board.

The shaft was driven by a 12v DC servo motor powered from a standard laboratory desk top power supply. This motor is able to provide speeds up to ~4000 rpm.

In the original description of the IME [1] the small variations in angular distance between the lines on the encoder disk, the interslot errors, were mentioned as a possible source of error. A special calibration of the disk was proposed to correct for this. Early in this research examination of the construction of medium to high resolution encoders revealed that the design commonly used has the property of averaging interslot distances [66]. Instead of scanning individual encoder disk slots a scanning reticle is used which carries a pattern of identical gratings to that on the grating disk. The light source is shone through this reticle, as shown in figure 29, and onto the rotating disk. The lines of the reticle coincide alternately with the lines and spaces of the disk grating as it rotates. In this way the periodic fluctuations of light which drive the photovoltaic cells and generate the encoder output signal are produced from across a number of grating lines. This averages the slight variations in interslot distances which means that these are a much less significant than in low resolution encoder devices where single angular intervals are measured.

The method of calibration for a low resolution device is described in detail by Routson [69] who used a 42 zone barcode-like printed strip on a shaft for torsional vibration measurement.

3.2 The data capture interface

The data capture interface of the experimental IME system is composed of two parts, the timer counter circuit board and the embedded computer hardware. The encoder output

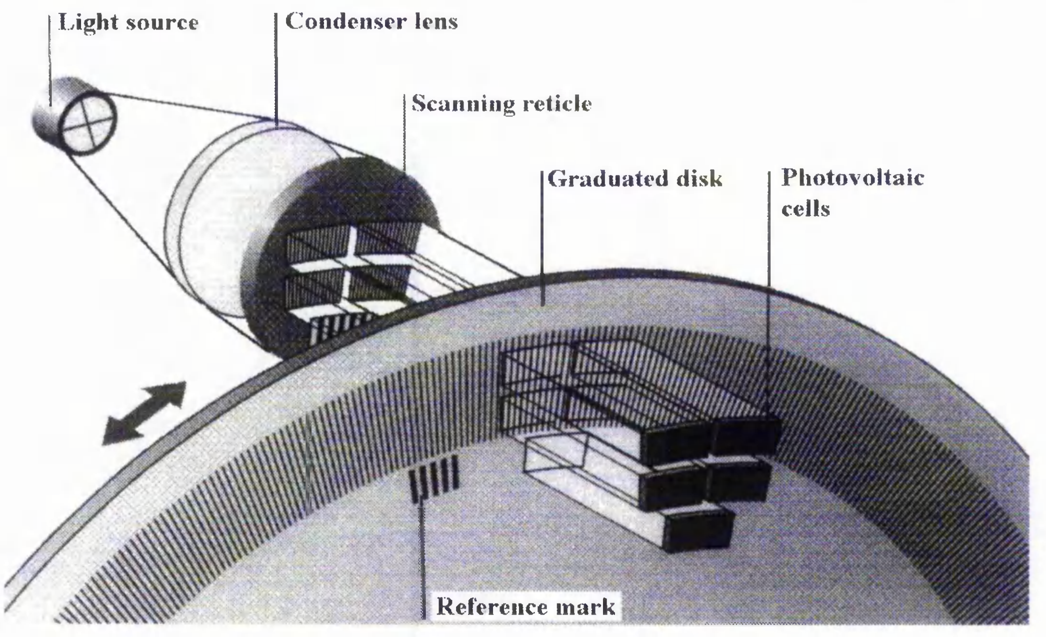


Figure 29. The method of scanning a high resolution optical encoder.

signals are connected to the timer counter circuit board which records time values for each wave edge. Embedded computers take the time values from the circuit board and store them for processing by the data processing computer. The data capture interface thus acts as a link between the outside world, represented by the encoder motion, and the internal computations of the system.

The initial design for the data capture interface was taken directly from that described in the IME patent application [2]. This initial design was based on that of the data logger previously used by Orton with a single read head encoder for precision angular velocity measurement [1]. It was expected that the components that had made up this data logger, some of which were implementation specific, would form the data capture interface of the IME. Thus the initial IME data capture interface was VME based because the data logger had been. There had been no initial attempt to make an application specific analysis of the unique requirements of the IME. The development of the data capture interface design during this project was driven by the need to analyse the requirements of the system from first principles in order to arrive at a hardware configuration which adequately met those requirements.

Appendix B describes development of the data capture interface in detail. In particular it explains the error analysis that was carried out to determine the correctness of the data produced by the system. The following two sections describe the basic function of the data capture interface's two parts and the implementations of each that were used for the project's experimental work.

3.2.1 The timer counter interface circuit

An interface circuit is required to integrate measurements of time with the square wave output of the IME device. These measurements are then processed by a computer system to produce time base interpolated angular measurements. This circuit is simple in construction and is shown schematically in figure 30. A high frequency precision clock oscillator (1) drives a free running 16-bit synchronous counter (2). Four identical 16-bit latches (3) are joined to the counter by a bus (4). The activating pin of each latch is connected to the square wave output of one encoder read head. The latch output pins are connected to the embedded computer system. The falling edge of each incoming read head's square wave latches a count value onto its respective latch. This count value represents the exact time the edge arrived at the latch relative to preceding and succeeding edges. Since the same counter provides the count values for all four latches, the latched counts also records the arrival time of the edge relative to those from the other read heads.

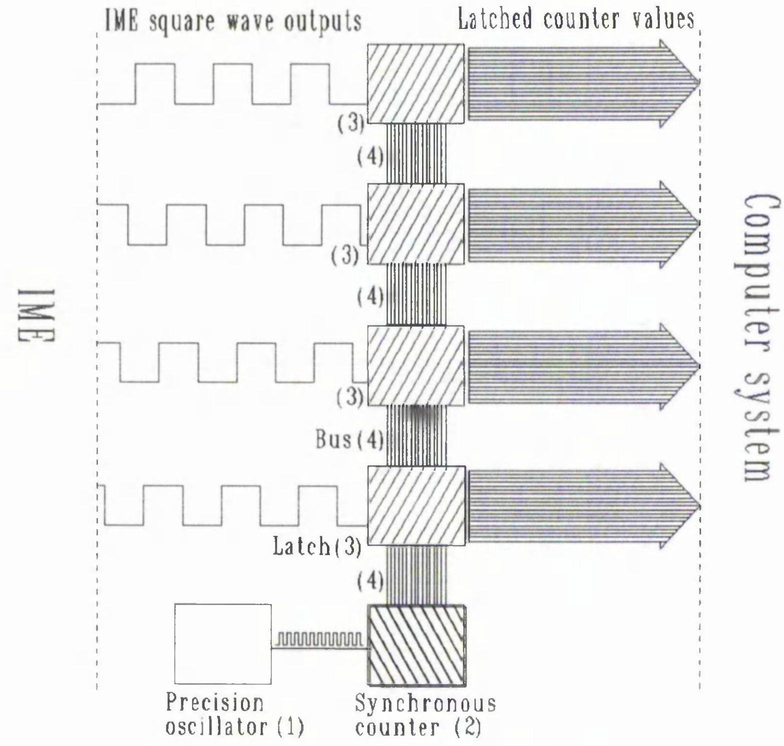


Figure 30. A schematic diagram of the IME interface circuit.

The embedded computer system is signalled by the falling edge that data is waiting on the latch. The computer can then collect the latched count for storage in a memory table. The difference between successive values gives the time base interpolation count.

Appendix C contains a block diagram and full circuit schematics of the circuit board that was implemented for the experimental system using TTL logic chips on a full height VME card.

3.2.2 The embedded computer hardware

From the overall design point of view there is no division between a data processing computer element and a data collection element in an IME system. Both these tasks could be done by a single computer and the correct set of peripherals using the current design of the experimental system. The origin of the distinction lies with Orton's original design and implementation which was based on the use of one embedded computer to log the data and a workstation or PC to analyse and then display it.

The current design of the data capture interface computer system is based on an examination of the data flow through the IME system as a whole conducted during the initial phase of this research. The data is inherently parallel from its origins through to the final processing stage. The four read heads produce asynchronous square wave trains in parallel which latch values from the counter on parallel latches which are read into the computer system via parallel interfaces. The counter data is also processed by the computer as parallel streams. From this analysis it was decided to replace the sequential algorithm used in the original design for data collection with a parallel solution based upon a generalised computer architecture. For expediency this design was implemented using the VME computer components available at the time although the generalised nature of the design would allow it to be implemented using different equipment with the same functionality. The design was called the conceptual computer architecture to emphasis its technology independence. Various implementations can be envisaged as future developments based on this framework.

As well as an analysis of the fundamental requirements of the IME data capture interface the conceptual computer architecture drew heavily from parts of the literature survey conducted in support of the project. Published descriptions of various commercial and experimental computer based data acquisition systems were reviewed. None of these were exactly analogous to the system which needed to be created to support the IME as almost all had some kind of analogue to digital conversion step. The use of direct memory access (DMA) controllers to allow the output of high speed analogue to digital converters (ADC) to be directly loaded into a computers memory was noted as a common feature of many data analysis systems [76,77]. It was also described as a way of implementing multi-processor communication [78,79].

The computer architecture required to support the IME is functionally divided between data collection, performed by the data capture interface, and data processing, analysis and display. As shown schematically in figure 31 the output channels of the IME interface circuit are each independently interfaced to a data collection unit (DCU) which is linked across a shared memory interface to the computer system which performs the measurement calculations.

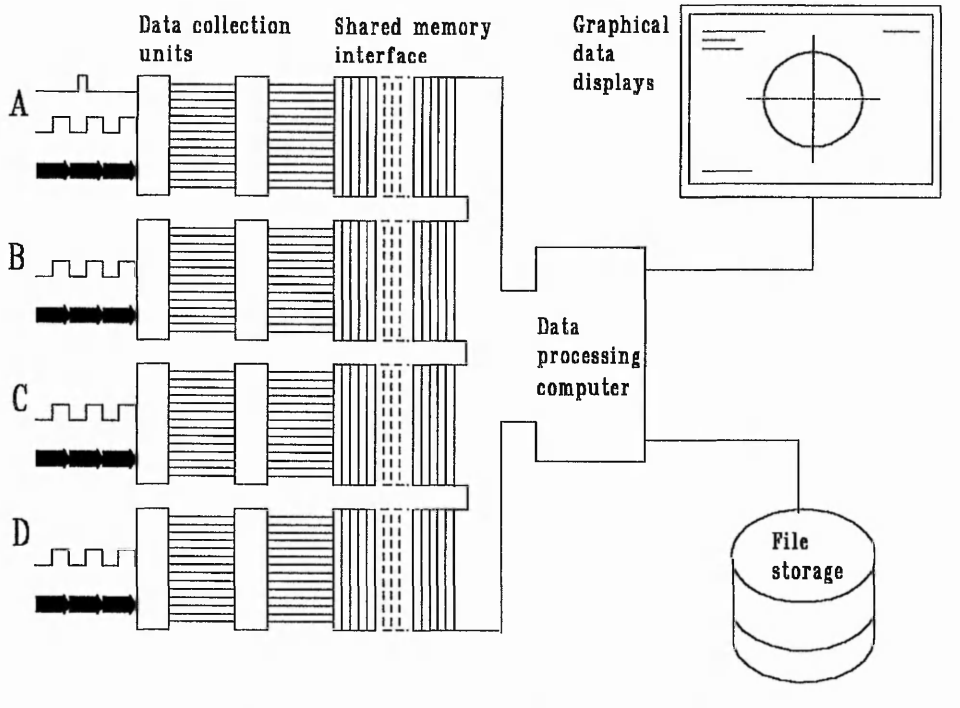


Figure 31. The conceptual computer architecture.

Figure 32 shows the functional components which make up a DCU. A parallel interface (PI) is used to link the latches of the IME interface circuit to the data bus of the DCU. The count values are read from the PI into the dual port first in first out (FIFO) memory of the DCU under the control of the memory management unit (MMU) responding to the falling edges of the IME output. The processing computer system sees the counter values across the shared memory interface as a four column table. The start of data collection is synchronised by the once per revolution signal from one read head. Use of this signal allows the data collected to be precisely related to the angular position of the external shaft. After it has been set by the computer system the DCU which receives this input signals the others via the shared memory interface. The number of counter values required is written to all the MMUs and they fill the data table accordingly. Data collection then stops and the values in the table are then processed to give values for displacement and angular velocity. The user has control of the collection and display of this data via a graphical user interface (GUI) running on the data processing computer. The system can gather data either in blocks, for a certain number of rotations, or continuously, averaging the resulting values for display or writing them to disk for historical analysis.

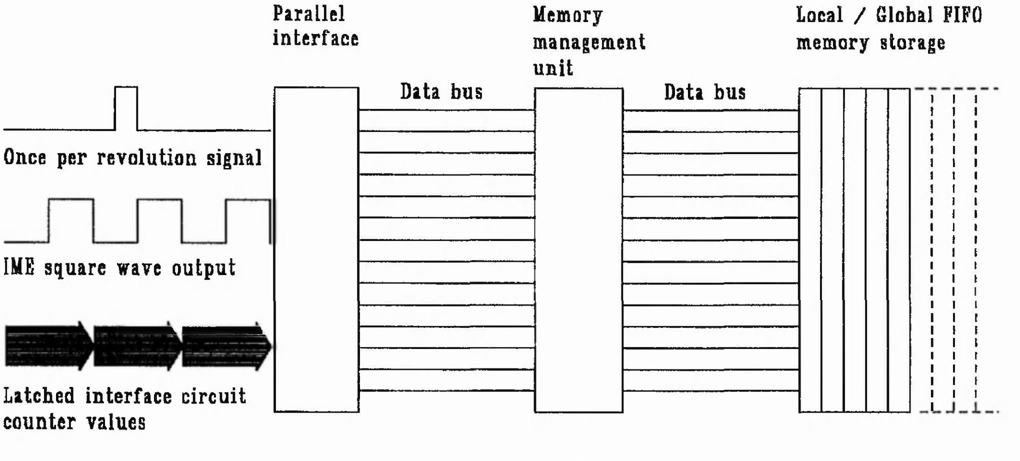


Figure 32. Data collection unit.

In the experimental system the DCUs were implemented with four monoboard 16 MHz MC68000 computers connected via a VME bus to global memory. Each processor acts as the memory management unit between its board's parallel interface and local memory. This parallel approach allows the system to collect data at shaft speeds up to 8000 rpm from a sensor with a 1024 grating line disk. A data rate of 136.5 kHz per channel. The IME interface circuit is housed in the same enclosure as the VME bus. A fifth MC68000 on the VME bus acts as bus controller, reading the data collected out of the global memory and transmitting it to a PC, via a 16-bit wide parallel connection, where the data processing program runs under a Microsoft windows GUI.

The DCUs have the same function as DMA controllers and these are the components with which they would ideally be implemented. Mated to the latches of the interface circuit each DMA controller could load data into the memory of a PC based system which might have a digital signal processing (DSP) co-processor to actually do the calculation of displacement and subsequent analysis. Despite not being able to make an ideal implementation of the conceptual architecture due to financial restriction, that which was made worked very well and proved the general operation of this design.

The highly accurate angular velocity monitoring aspect of Orton's work [1] was subsequently developed by Steiger [80] as part of a method of improving the smoothness of stepping motor based, multi-axis, continuous path motion for machine tools. Stieger's work has itself become the subject of further development by MacManus and Stout. To support the development of Steiger's machine control algorithm a new hardware architecture has been designed for the data capture interface based on a PC interface card [81].

3.3 The data analysis and display computer software.

The experimental IME system used two separate computer platforms to host software for the analysis and display of the sensor data. A 386 PC was used for direct control over the upload of data from the embedded data capture interface and a UNIX workstation was used for the computationally intensive numerical analysis of the data. This section explains the basic function of this software and the differences between each platform.

3.3.1 Software structure

Figure 33 is a data flow diagram which illustrates the structure of analysis and display software. The source of the counter data, upon which all subsequent processing takes place, is the main practical difference between the two platforms. The PC based system takes this data directly from the data capture interface. It therefore has to check the data for errors and correct these where found. This data is saved as a file to allow the UNIX program to access it across the local area network (LAN) to which both machines are connected. To minimise file size the data is saved in its primary form as 16-bit timer counter values rather than as displacement or any derived data type. The programs on both platforms perform the same basic calculation to derive shaft centre displacement. This calculation is explained in detail in the next section. The PC program displays the displacement values chiefly to allow the data loaded by the system to be visually checked and the alignment of the rig's elements adjusted.

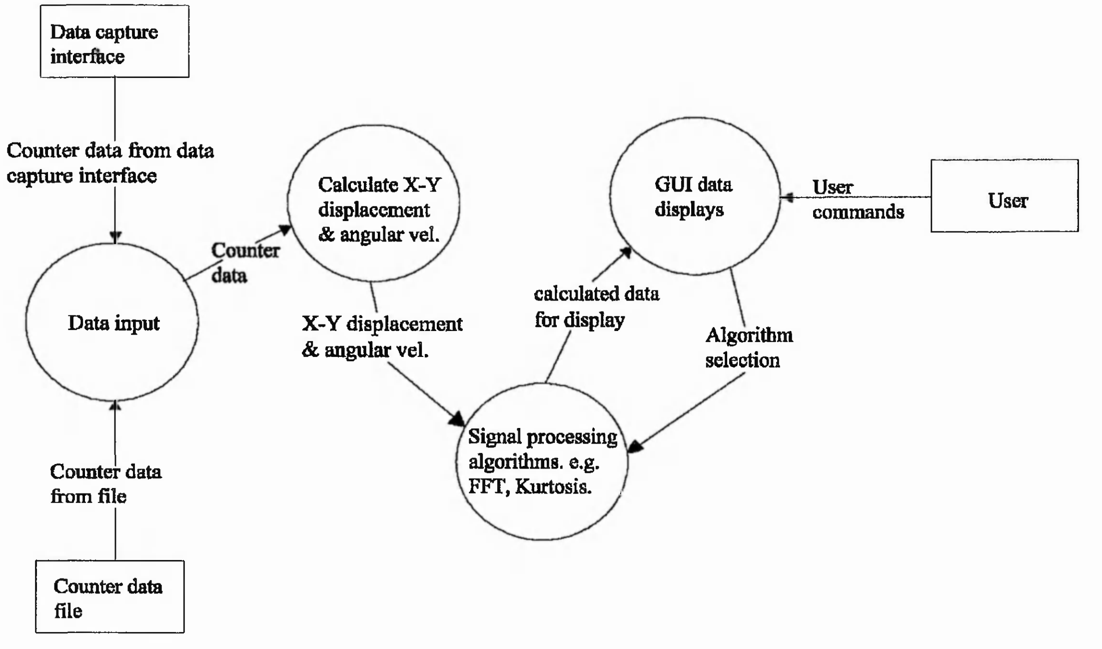


Figure 33. Data flow diagram of the data analysis and display software.

The main, UNIX based, data analysis program allows the display of displacement, average-removed or smoothed displacement, angular velocity, shaft centre velocity, shaft centre acceleration and time interval between data points. Each of these data types is calculated for the two directions of measurement of the IME. The display allows for the scaling and scrolling of the data recorded. The display may also be printed out. Analysis of the data is carried out with a variety of DSP algorithms. This DSP functionality in the experimental software is provided by a specialist library of DSP routines. These routines provide the program with the ability to do FFTs, various digital filters, correlation and apply statistical techniques. The analysis program integrates these routines into their own set of GUI sub-menus, allowing them to be selected and applied in a flexible way to the displacement and other derived data types.

3.3.2 The basic displacement calculation

The analysis computer system sees the output of the IME as a series of counter values from each read head, one for each line on the encoder grating disk. The angular difference between read heads is calculated from these counter values. The start of data collection from all four reads heads is synchronised by a once per revolution signal marker on the encoder disk. This is taken from the output of one read head. From this point the successive timer counter values are stored in a four column table in the computer's memory the format of which is similar to table 2 below. Each row consists of two pairs of values from the diametrically opposed read head pairs A-B and C-D.

Table 2.

Angular position	Counter values			
	A	B	C	D
0	A507	A60B	A5B7	A5AD
1	A68B	A78D	A73B	A730
2	A80F	A910	A8BE	A8B3
3	A994	AA92	AA40	AA36
4	AB17	AC15	ABC5	ABB9

The row number indicates the angular position, at intervals of the encoder grating line spacing, from which the readings were taken. If the disk is centred on the diameter joining the read heads then a grating line will pass each read head simultaneously. In this case the numerical difference between pairs of values in each row will be zero. If the disk is off centre, and there is an angular difference between the read head pairs, then counter values at the same angular position, indicated by the row number in the table will be different. The grating line passes each read head at a different point in time and therefore latches different counter values. This difference in time is used to find the precise angular difference between the read heads by means of time base interpolation. The counter is free running.

Figure 34 shows a graph of angular position against time for read heads A and B with an inset showing the format of the memory table. In this example A is shown to be leading B as would be the case from the read heads shown in figure 23.

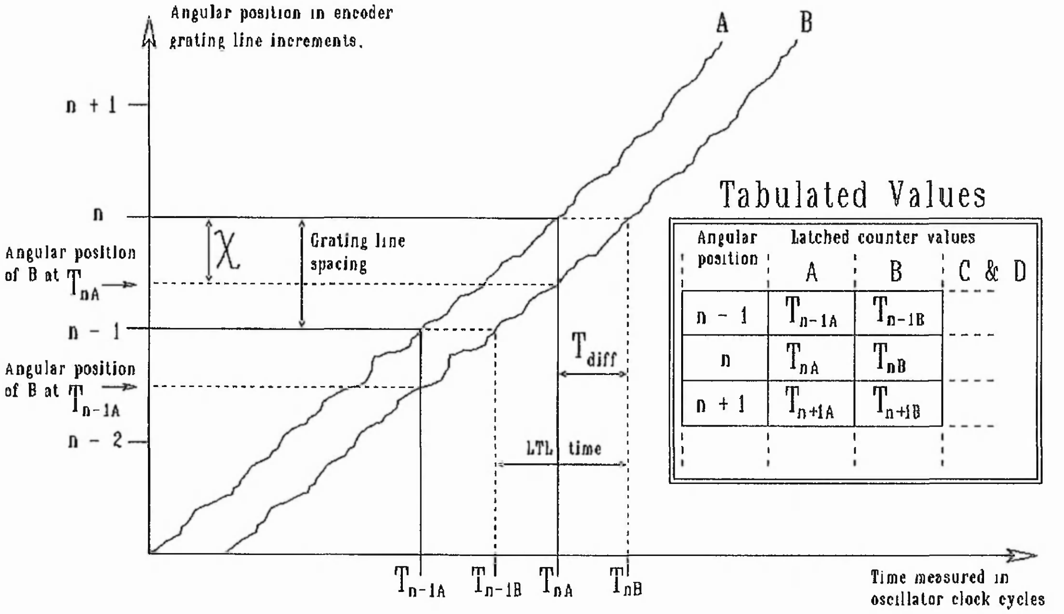


Figure 34. Calculation of relative angular position by time base interpolation.

At angular position **n** the time value recorded from read head A, $\mathbf{T_{nA}}$, is less than that recorded at read head B, $\mathbf{T_{nB}}$. To find χ, the difference in angular position at this point, the angular position measured by B at $\mathbf{T_{nA}}$ must be found by interpolating between the values **n** and **n - 1** measured by A. Assuming a constant angular velocity over the distance **n - 1** to **n**, the difference in time, $\mathbf{T_{diff}}$, measured between $\mathbf{T_{nA}}$ and $\mathbf{T_{nB}}$, will bear the same percentage relationship to the line to line time $\mathbf{T_{nB}}$ - $\mathbf{T_{n\text{-}1B}}$ as χ bears to the encoder grating line spacing, **n** - (**n** - **1**). This allows χ to be found. θ is then given by the relationship shown in equation 10.

$$\theta = \pi - \chi \qquad (10)$$

The above algorithm is a simple interpolation between two points. This is the method which has been used as the basis for the experimental processing software. To increase the margin of accuracy interpolation could be performed between a greater number of points using more complex methods. In the experiments which have been conducted as part of this research the accuracy of the simple algorithm has proven sufficient for the system to be evaluated.

Chapter 4

The Initial Experimental Investigation

4. The initial experimental investigation

This chapter details the first stage of the experimental investigation into the application of the IME to the condition monitoring of rolling element bearings. The aim of this initial investigation was to determine the extent to which conventional approaches to bearing monitoring could be applied to the signal produced by the IME. The chapter's first section describes the form of the signal produced by the experimental IME system. Next the result of a simple experiment conducted to test the response of this signal to actual bearing condition is presented. The remainder of the chapter explains how the analysis of the IME signal from a bearing with an induced defect was conducted. The chapter ends with the conclusion, drawn from this analysis, that a model needed to be developed to explain the response of the IME signal so that an interpretation method for could be devised for it.

4.1 The form of the IME signal

Chapter 3 described how the theory of the IME predicts that shaft centre position can be derived from angular measurements recorded as time base interpolated counter values. This section examines the initial results from the experimental IME system which showed for the first time what the signal from this new device actually looked like.

Figure 35 is an example of a set of displacement traces calculated from a data file collected at the start of the investigation. This figure is a screen capture from the display of the data analysis program. The top trace is from the vertical measurement direction, the horizontal measurement direction is at the bottom. The y axis of each trace measures shaft centre displacement in microns. The x axis in both cases is angular displacement from 0 to 2π in 1024 increments. The data shown came from a single rotation of a tightly adjusted RHP

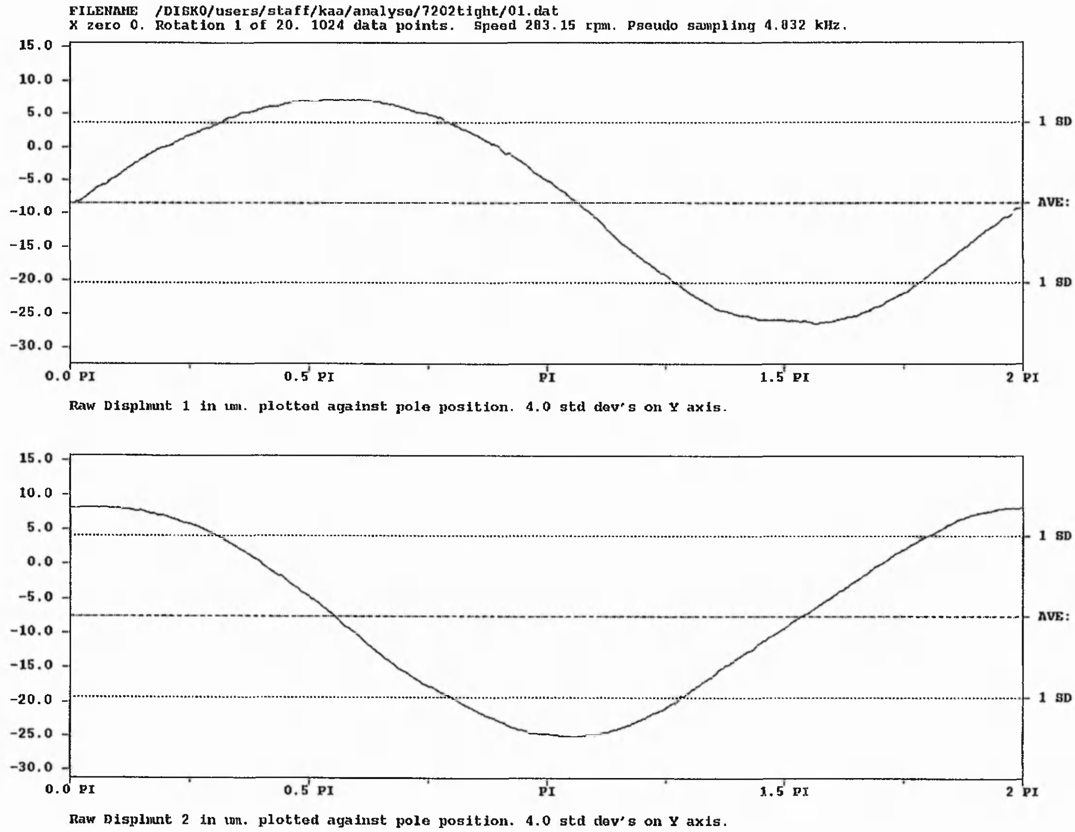


Figure 35. Displacement in two dimensions measured by the IME.

7202 BETN 40° angular contact bearing rotating at 280 rpm. These traces illustrate the general form of the signal that the sensor produces. The wave form from each axis of the sensor is sinusoidal, with a 90° degree phase difference between the two signals.

Each trace is composed of 1024 individual data points. Each data point on the two channels is the result of interpolation between a pair of timer counter values from each opposing read head pair. This results in one displacement value per grating line for each channel. As there are 1024 lines on the disk of the experimental IME, 1024 displacement values are calculated on each channel for each rotation. The angular increment between each value is equal while the time interval is variable. This is a principal difference between this type of data, its subsequent analysis, and the conventional method of sampling data at equal time intervals. Since the timer counter values provide a continuous time record it is possible to resample the data to derive the displacement at equal time intervals. However this would result in the loss of the inherent angular position information. Consequently it has been regarded as advantageous to look at the sensor data in terms of equal intervals of angular displacement as in figure 35.

The sinusoidal form of the signals can be explained from the construction of the device with reference to figure 36. The centre of the encoder disk, X, is not expected to coincide with the actual centre of the shaft rotation, R. The point tracked by the sensor is the nominal centre of the encoder disk, determined by the grating lines, orbiting around the actual centre of rotation. Each axis of measurement records this motion as a sinusoid. When plotted against one another, read heads C-D on the y axis, A-B on the x axis, the circular trace shown in figure 37 is produced. This orbit represents the motion of point X.

The motion of the centre of the disk will be modified by the motion of the centre of rotation. Changes in the displacement of the rotational centre Δx, Δy are recorded by directly related motion of the encoder trace. The sensor measures this displacement relative to the zero position indicated by the intersection of the diameters joining the read heads, X'. The distance between the centre of rotation and the centre of the disk is the eccentricity e. In this example of simple motion the eccentricity is the radius of the disk centre orbit. Here the terms shaft centre and centre of rotation are used interchangeably.

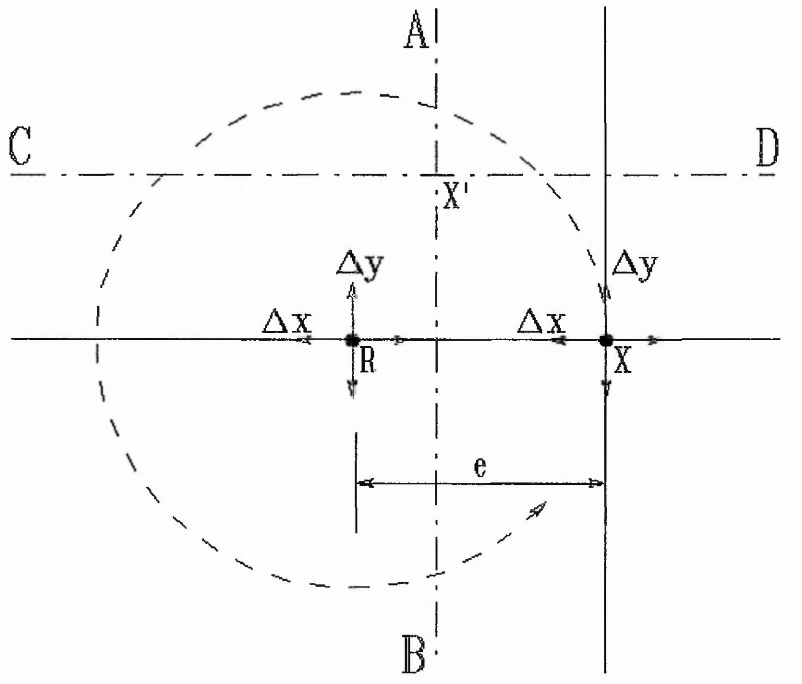


Figure 36. The relationship between the encoder disk centre, the centre of shaft rotation and the signal produced.

At this stage, when the model for the signal had been established, sources of possible error could be predicted. These were errors due to the device's method of operation, rather than those specifically due to the hardware used for this implementation. The first phase of analysis, before any experimental work, was to develop an initial method of signal processing to account for these errors.

The position indicated as the centre of the encoder disk might be subject to change during a rotation due to changes in grating line spacing. Fluctuations in speed will also cause a loss of accuracy in the position indicated as the number of available interpolation steps will decrease and effectively increase the quantisation of the data points. Any effect the disk has on the data will be cyclical. The same error will occur at the same point on successive rotations. The IME is unique in that each data point has an angular displacement position

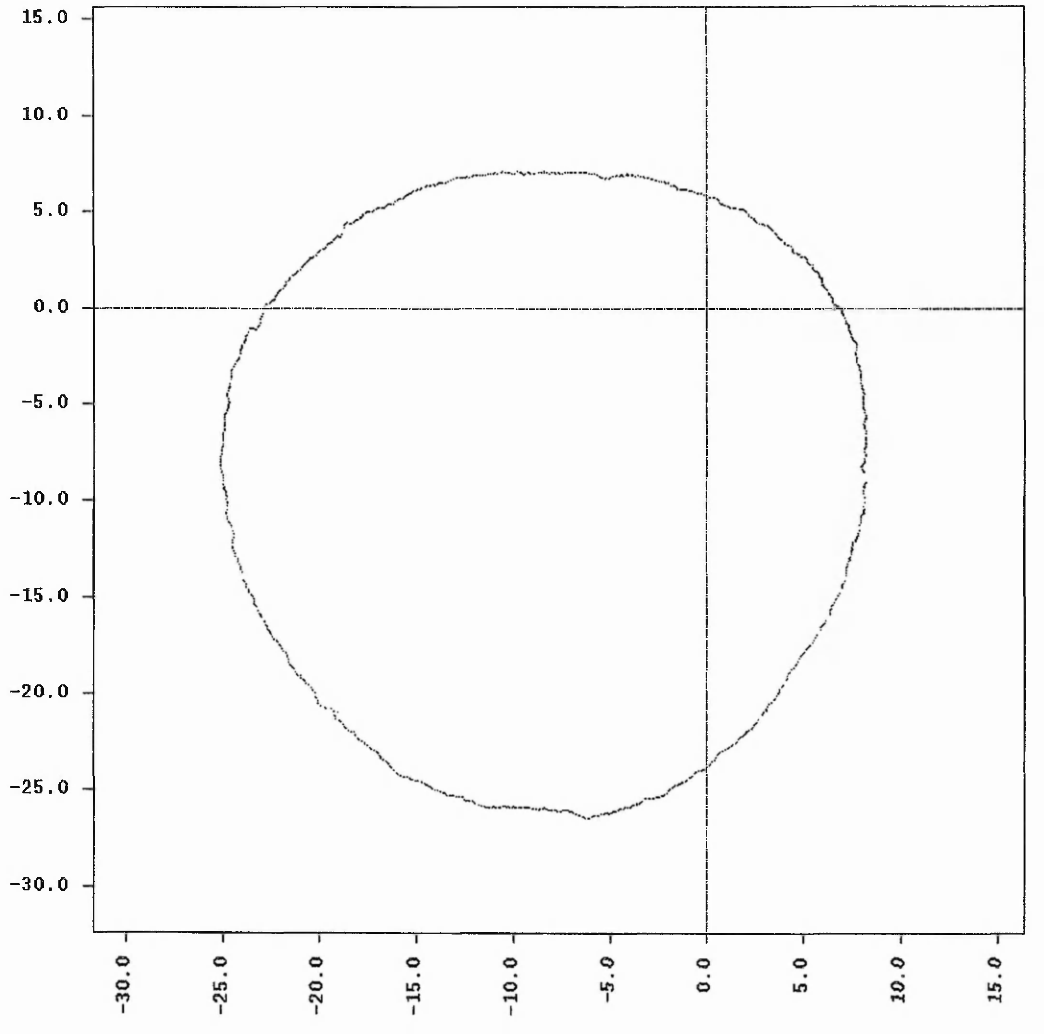


Figure 37. Displacement values in microns from the two axes of measurement plotted against each other for a single rotation.

associated with it. This means that the displacement value for exactly the same angular position on successive revolutions can be compared.

If the centre of rotation were fixed in the x-y plane then the orbit recorded by the sensor would be a perfect circle except for deviations caused where the displacement values were modified by disk errors or speed fluctuations. By averaging the measurement at each angular position a reference value can be established for each position that accounts for any repeated error. The rotation shown in figure 37 is the first of a data set consisting of twenty similar rotations. Figure 38 shows the traces for all twenty rotations overlaid. The traces show a high degree of coincidence. Points from the same angular position plot within ~1 μm for the same angular position. For angular position n, which can be any one of the 1024 positions in the rotation the average displacement d, over twenty rotations for that position, AVEdn, is given by:

$$AVEdn = d1n + d2n + d3n \ldots d20n / 20$$

Where dXn is the value at position n on rotation X.

The 1024 average measurements will each define a zero position for the associated angular position. Assuming a fixed centre of rotation subtracting this average value from the values recorded at this point on each rotation will give zero. This is equivalent to subtracting the value of the eccentricity *e*, shown in figure 36. Removing this average from the displacement leaves the displacement of the shaft centre relative to its average position. For real data this is an approximation since it has been observed that actual bearings give a variable motion of the shaft centre due to different ball or roller diameters as successive elements pass through the load zone. For this reason the traces in figure 38 are not exactly coincident while the shape of successive traces is similar.

Subtracting the average displacement value for each angular position from the displacement data points in the two traces shown in figure 35 gives a trace of the displacement of the shaft centre. This is shown in figure 39. This signal is referred to as smoothed displacement since the irregular sinusoidal waveform has been smoothed out. Plotting

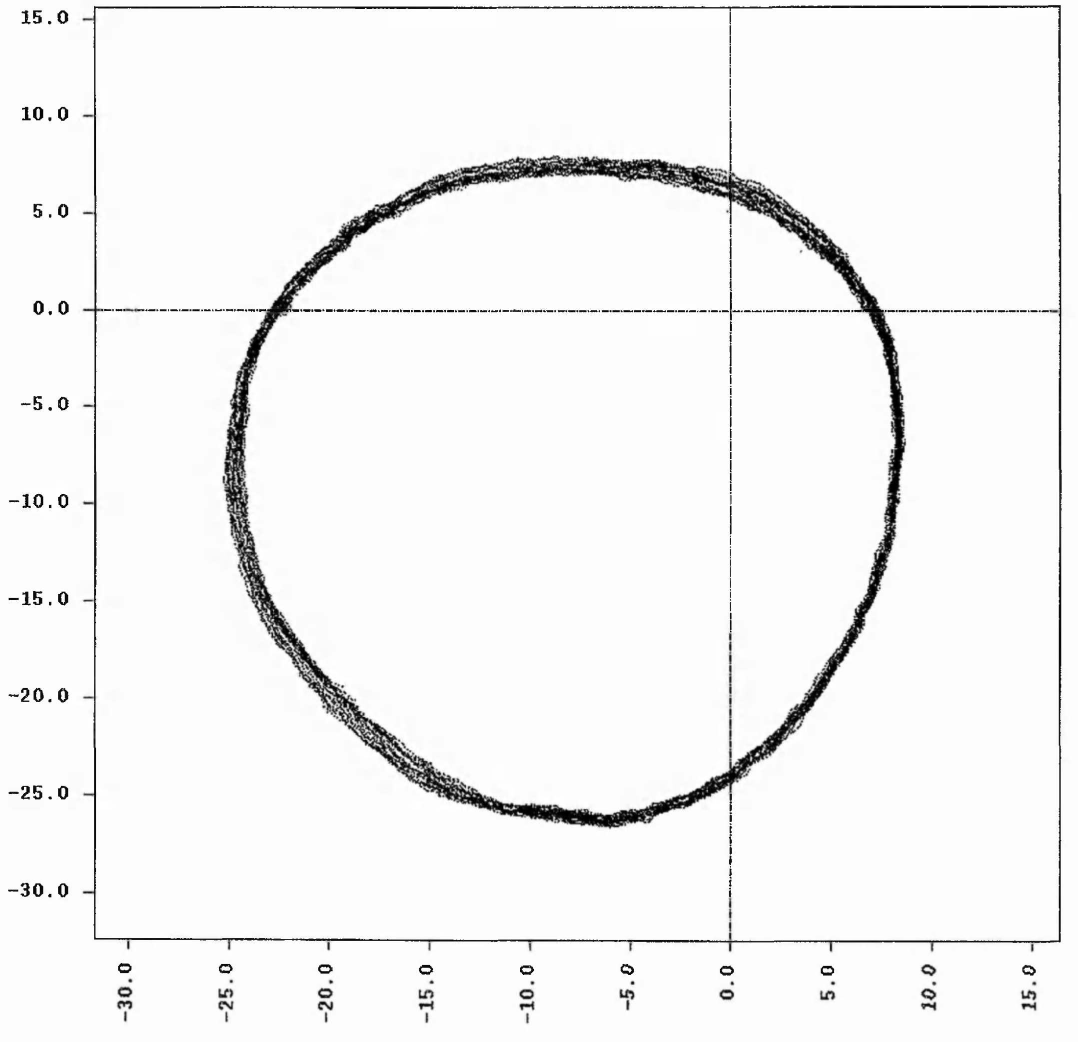


Figure 38. Displacement traces for 20 rotations overlaid to show degree of coincidence.

these traces against one another produces a plot showing the motion of the centre of rotation in two dimensions. An example of this motion is shown in figure 40.

Other methods have been considered to remove the average sinusoid from the displacement data. In particular generating sine waves and subtracting these from the displacement data has been investigated. The results of this method have not been found to give better results than the simple averaging method described above. The average sinusoid is not a perfect sine wave as it contains systematic errors due to imperfections of the encoder disk. Sine waves generated to fit the data do not take these errors into account.

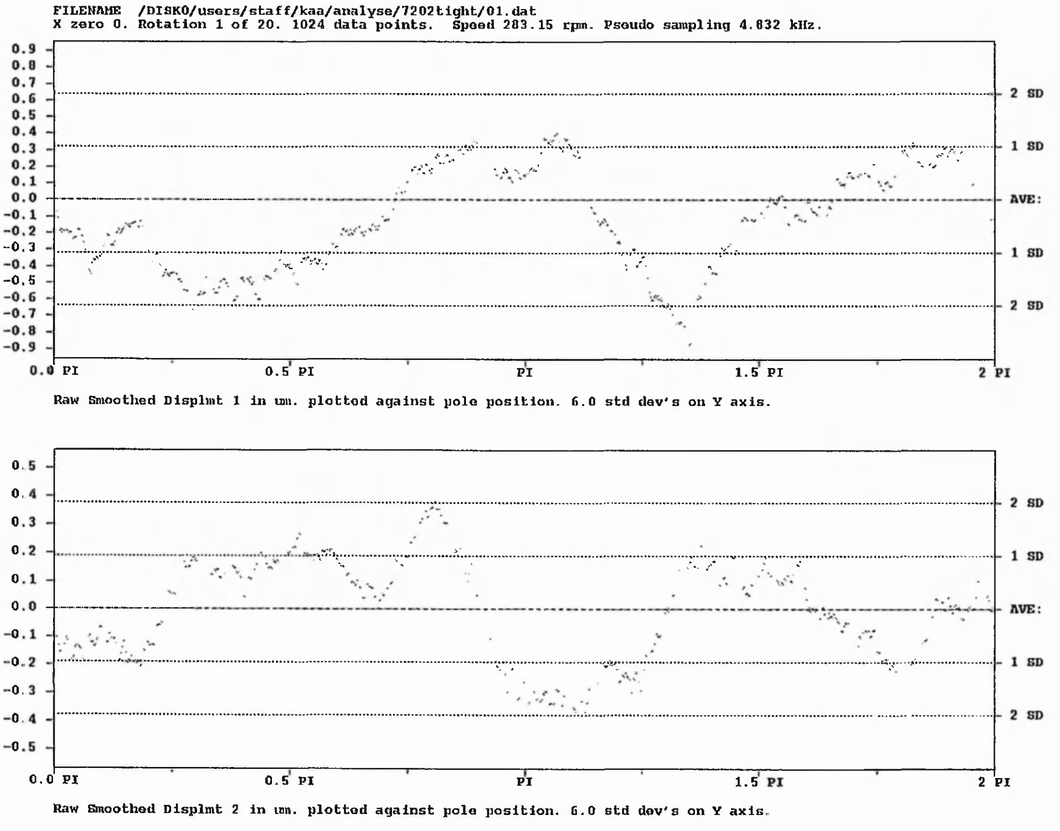


Figure 39. Removal of the average sinusoid from the traces in figure 35 produces these traces showing the motion of the centre of rotation.

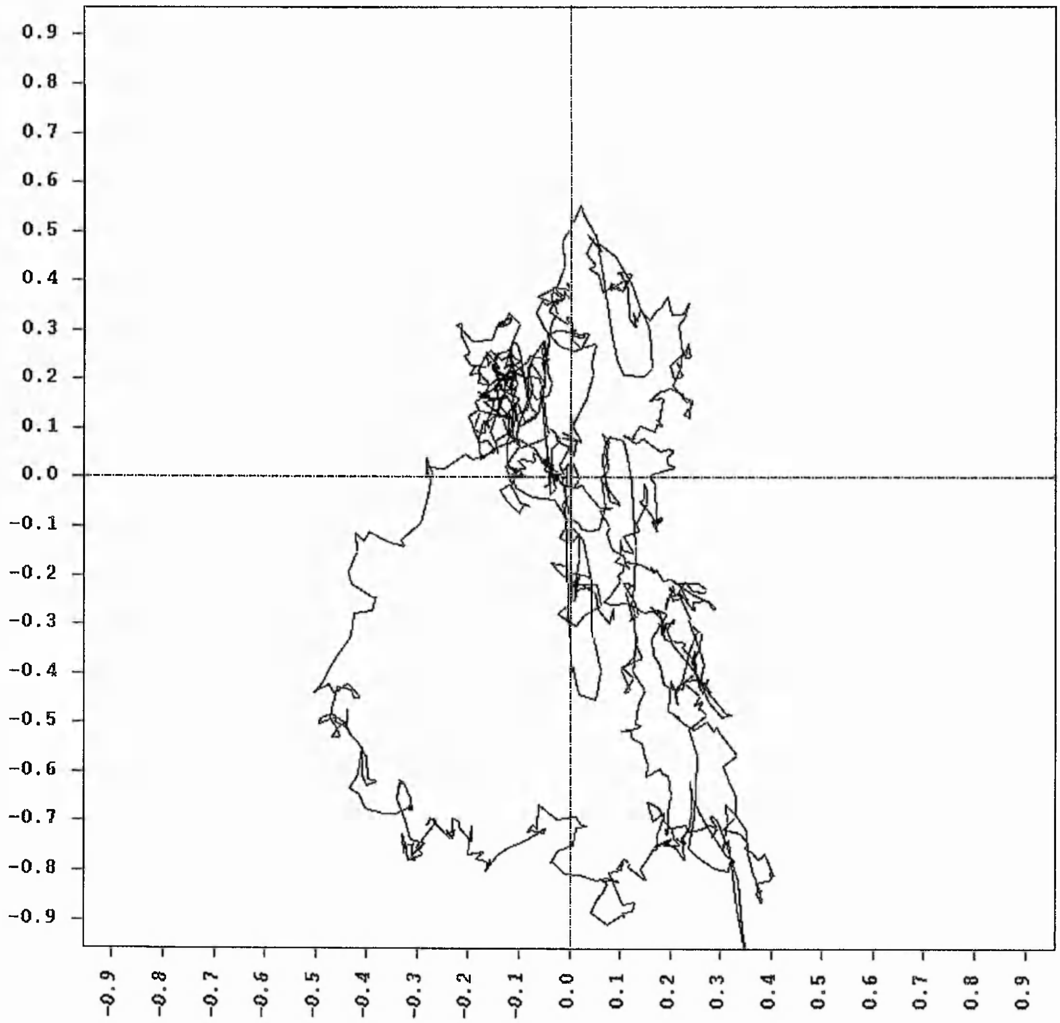


Figure 40. Combined plot of the two traces in figure 39 showing the motion of the centre of rotation in two dimensions. Scales in microns.

4.2 Shaft movement experiment

The data presented so far in this chapter has come from an angular contact bearing with the inner race tightly adjusted. By changing the relative position of the inner and outer races of an angular contact bearing it is possible to increase the amount of freedom between the balls and the outer race. This effect was used as a way to perform an initial test of the IME's ability to detect changes in bearing condition. This test simulated the possible condition of a badly adjusted or grossly worn bearing. When the IME was used to monitor this condition, its output produced a very interesting set of results which revealed phenomena, which from the literature surveyed, have not been seen previously with this level of detail from rolling element bearings.

Figure 41 shows the braided net-like pattern produced when 14 successive traces from the loosened bearing rotating at 300 rpm are overlaid. The x and y axes are displacement in microns as before. With close examination, it is clear that this pattern is formed from 7 distinct paths traced by the shaft centre on successive rotations, each repeated twice.

Figure 42 shows how the first seven rotations build up to give the pattern shown in figure 41. From the top left of the figure, showing the first two rotations, to the bottom left, showing the pattern given by the first six, the area circled shows the starting and ending points of the pattern. The pattern is clearly incomplete. In the bottom right of the figure the complete pattern built up by seven rotations is shown. In this case the pattern is closed, indicating that after the seventh rotation the internal elements of the bearing have returned to their starting positions. The pattern then repeats in the same way over the next seven rotations to give the complete pattern in figure 41.

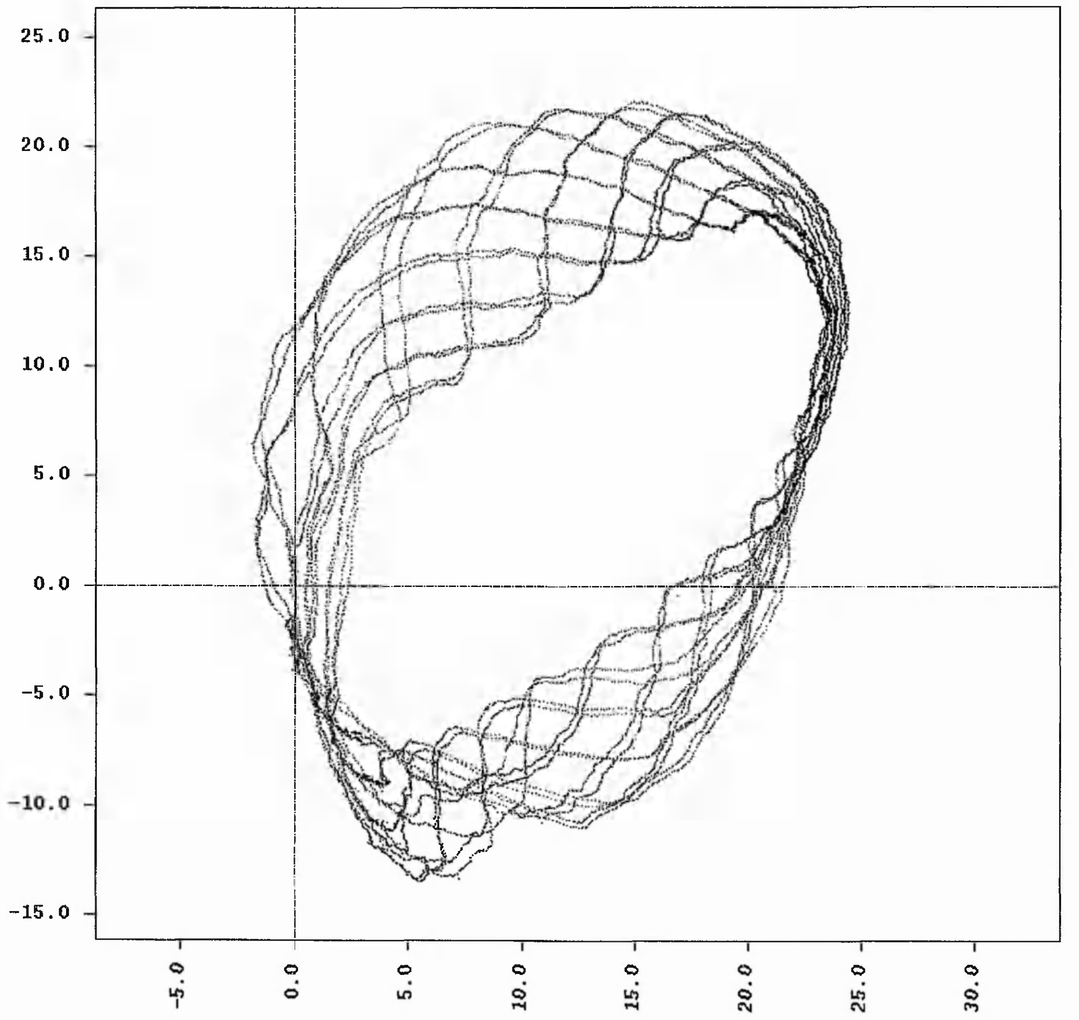


Figure 41. Displacement traces from 14 rotations of a loosely adjusted angular contact bearing.

The interpretation of this, and similar results, is that each path taken by the displacement traces represents a particular set of positions taken up by the bearing's internal elements during each rotation. Using equation 2, explained in section 2.3.1, the number of rolling elements which pass any point on the inner race during a rotation can be calculated from the internal dimensions of the bearing elements. In the case of the 7202 BETN bearing used for the above experiment this number is 6.57. The number of rotations taken for all the rolling elements to return to any given initial position with respect to the inner race will

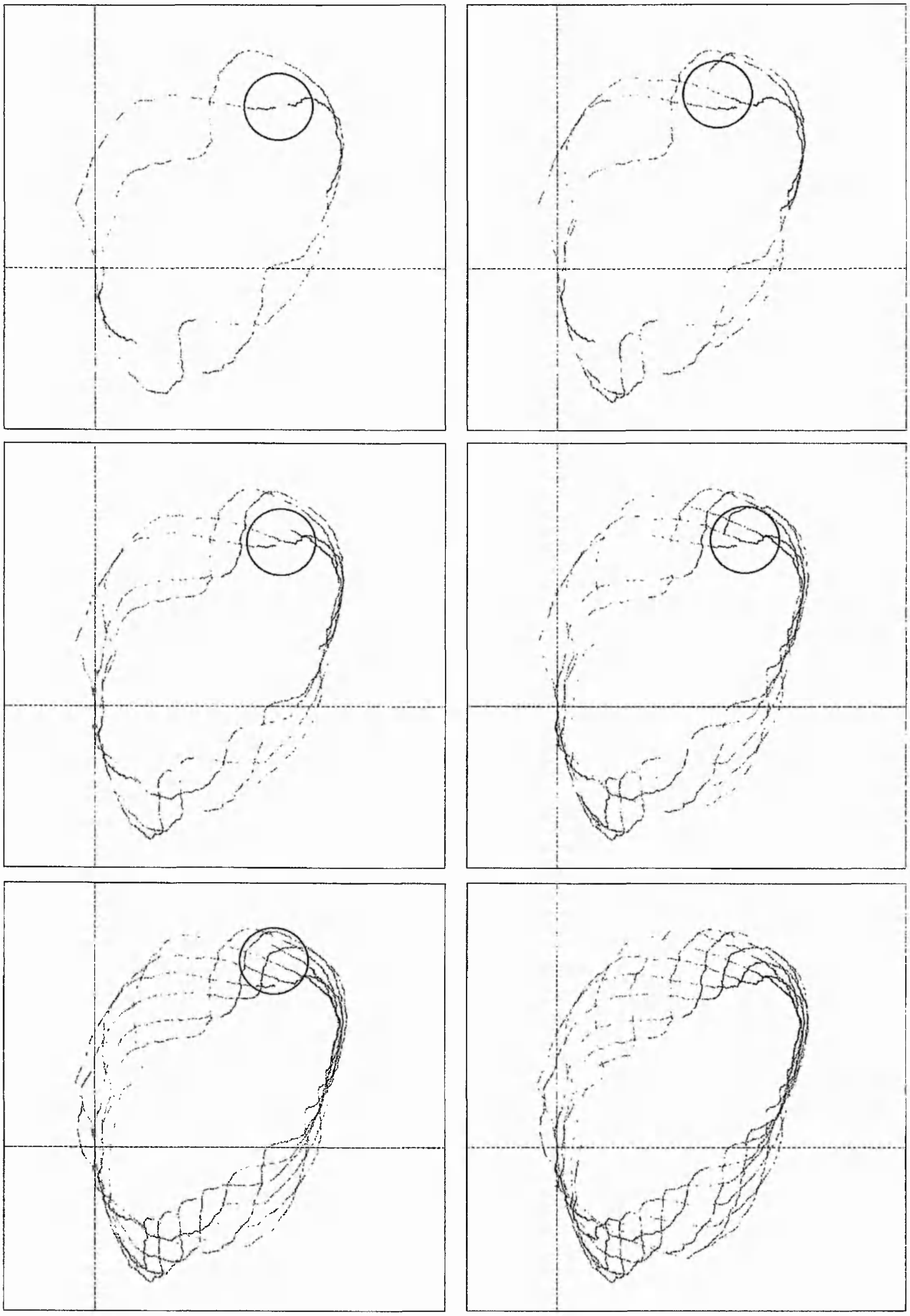


Figure 42. The pattern in figure 41 is built up by separate displacement traces over 7 successive rotations. The circled area highlights the beginning and end of each trace showing how the pattern returns to the same point at the end of the seventh rotation.

be the first whole multiplier of 6.57. Calculation shows this to be seven. This directly confirmed the observation made experimentally from the IME displacement traces. The coincident parts of all the traces are likely to be the points at which the shaft's motion is naturally constrained by its weight and the influence of the rest of the rig.

At higher speeds complex motion of the shaft centre can be observed within a single rotation of the loosely adjusted bearing. An example of this data is shown in figure 43, a single rotation monitored at 630 rpm. The IME has clearly been able to follow the skips and bounces of the shaft centre after the inner race and rolling elements strike the outer race. The displacement trace shows the acceleration of the shaft centre as increased distance between data points. As might be expected, as the centrifugal forces increase with rotational speed, the degree of motion increases and the patterns become increasingly less regular until the paths followed appear essentially random.

The significance of these results to the project was the fact that they clearly showed that the IME system was capable of resolving changes in the state of a bearing from the motion of a shaft. This was the first proof that the theory of the IME correctly predicted the performance of the experimental sensor system. The resolution of the results was particularly important. The data points of the displacement traces recorded during this series of experiments were clearly directly related to each other and free from any error or fuzziness which might have been due to calculation or data collection errors. As can be seen the data points lie so closely together that they look like a single line. In figure 43 the motion has clearly been followed as the shaft centre has accelerated.

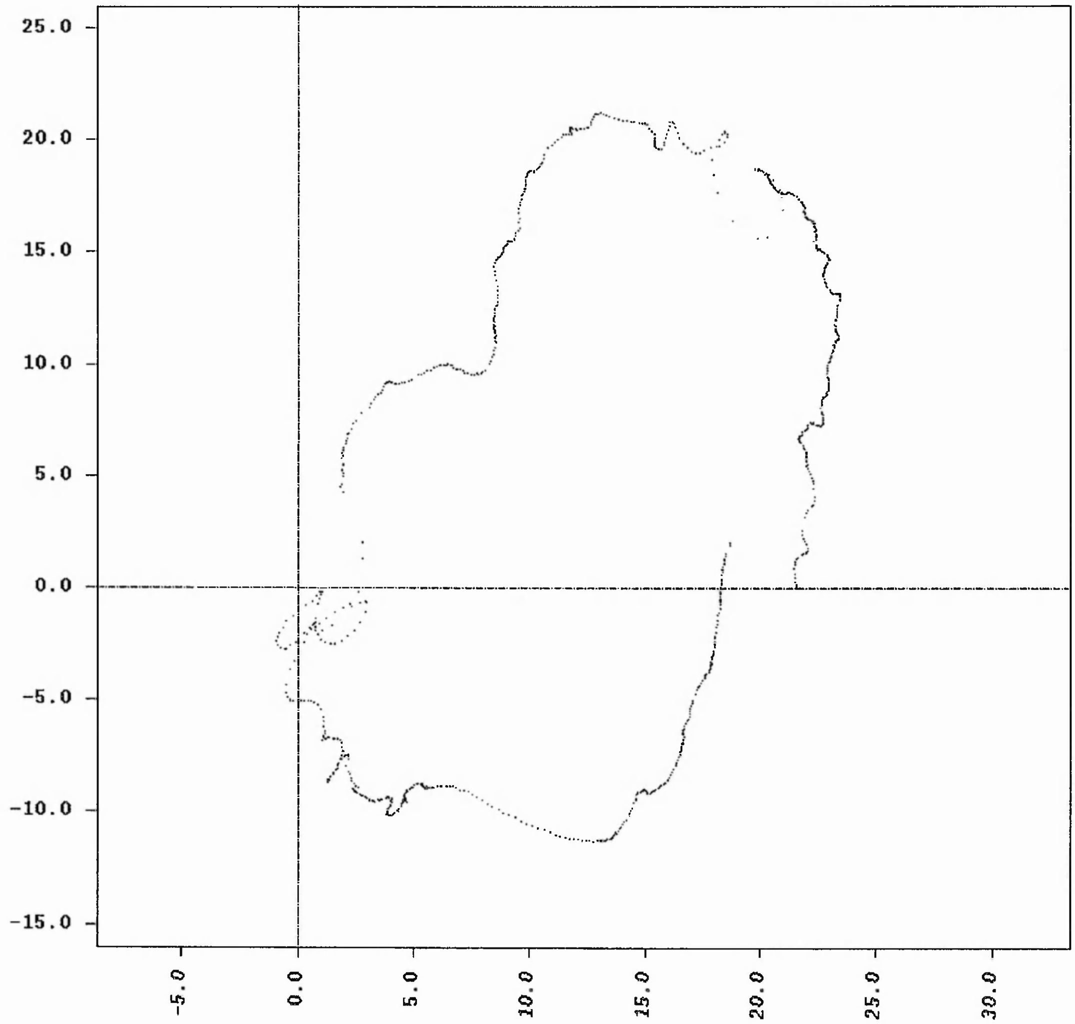


Figure 43. Complex shaft motion from a loosely adjusted angular contact bearing.

Detailed treatment of these results from the angular contact bearing was not carried out as the main aim of the project was to establish the ability of the system to resolve bearing defects. A decision was taken to concentrate on the obvious relative accuracy of the experimental sensor rather than attempting to quantify its absolute accuracy. The degree to which 1μm measured by the sensor actually corresponded to 1μm of shaft motion could not be investigated within the scope of this project. Suitable equipment was not available to measure these tolerances independently. Such an investigation could offer an interesting direction for future work.

4.3 Initial bearing defect experiments

The methodology adopted for the conduct of the initial set of bearing defect experiments was based upon examples of other work taken from the literature read in support of the project. Most of the literature surveyed had presented the spectral analysis of bearing vibration as the main conventional method of bearing condition monitoring. Using the equations given in section 2.3.1 the frequencies in a vibration signal's frequency spectrum can be related to defective bearing components. The IME signal is totally different from that obtained from a case mounted velocity pick-up or accelerometer using the high frequency resonance technique or a similar process. However, since the IME monitors the same basic process, bearing rotation, it was expected that the same information could be extracted from its signal. Defining the method of this extraction was the aim of the experimentation and was expected to constitute a fundamental part of the project. The initial approach adopted was to use conventional frequency analysis techniques on the signal to see if the calculated defect frequency was present.

For the initial bearing defect experiment SKF NJ202 single row roller bearings were used. As well as being the size required to fit the experimental rig, the inner race could be freely removed from this type of bearing. This allowed a defect to be made relatively easily on the inner race of the bearing. This defect was produced by grinding a ~1mm flat diagonally across ~5mm of the bearing's inner race. A defect of this size was obviously far larger than any normally occurring bearing defect would be. A large defect was used as the aim of this initial experiment was to investigate the ability of the IME to detect the presence of a defect and it was not known what the lower limit of the IME's sensitivity would be.

An unmarked bearing of the same type was used as a reference for comparison with the marked bearing. In retrospect this was not the best way to arrange the experiment as it would have been better to have looked at the bearing to be tested before it was marked and then compared the results obtained.

4.3.1 Comparative accelerometer experiments

To provide a comparison for the results expected from the IME, a pair of accelerometers were used to obtain vibration signals from the experimental rig in the conventional manner. Unfortunately it was not possible to collect this data at the same time as data from the IME since both sets of experimental equipment could not be brought together in the same location. Figure 44 shows the positioning of these accelerometers on the experimental rig and their relationship to the IME's channels. Accelerometer signals were collected and examined from both the marked and unmarked bearings using a spectrum analyser. The bearings were run at very close to a constant speed. This allowed the position the characteristic bearing frequencies should occupy in the frequency spectrum to be calculated. These frequencies in Hertz and in cycles per rotation are shown in table 3. As the IME records data at constant angular intervals the frequency in terms of the rotation is used. This allows the identification of cyclic events which occur relative to the same point on each rotation which give integer frequencies. These differ from events related to the bearing elements which give non-integer frequencies.

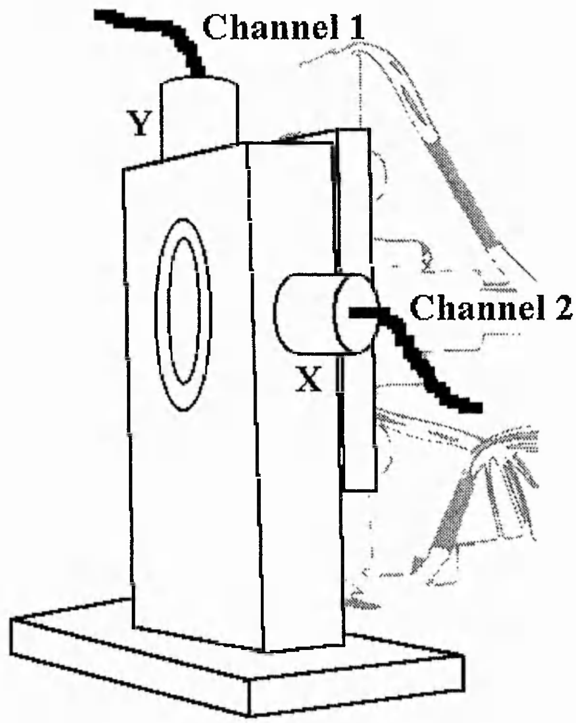


Figure 44. Accelerometer placement on the bearing test pillar of experimental rig.

Table 3

Calculated defect frequencies for an NJ202 bearing rotating at 2650 rpm, 44.1 Hz.		
Defect type	Frequency	Frequency in cycles per rotation
Inner race defect	296 Hz	6.72
Outer race defect	189 Hz	4.28
Rolling element defect	191 Hz	4.29
Cage defect	17.22 Hz	0.39

Figure 45 is an example of the signal produced by the accelerometer monitoring the y axis of the marked bearing seen in the time domain. The figure is taken from the screen of the spectrum analyser. Although it is possible to discern definite impulsive components in this waveform, not seen from the unmarked bearing, relating these to the calculated bearing defect frequencies proved impossible. In the time domain no obvious regularity could be

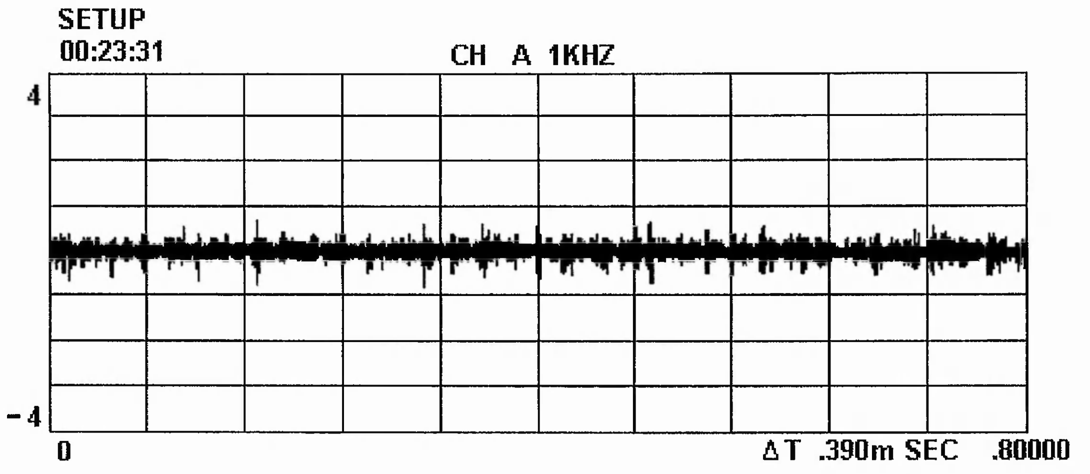


Figure 45. Time domain accelerometer signal recorded from the y axis of the marked bearing.

observed in the signal. Attempts were made to fit time intervals corresponding to the calculated frequencies between the impulses seen in the signal. These failed and, in the time domain at least, the signal appeared

An example of the frequency spectra of the accelerometer signals recorded from the marked bearing is shown in figure 46. The upper spectrum is from the x axis with the y axis below. In the case of the frequency domain, although distinct peaks were seen in the spectra, no precise relationship could be made between the data and the calculated frequency for an inner race defect, around 296 Hz. There were peaks which seemed to correspond to the outer race and rolling element defect frequencies between 180 and 200 Hz, but these seemed no more distinct than others whose origin could not be linked to any part of the bearing. The main cursor asterisk in both traces is at 333.75 Hz. The prominent peak at this frequency lies outside the range of the frequencies expected from the bearing. In figure 46 the only peak that could be conclusively explained appeared at around 44 Hz which was the frequency of the shaft's rotation.

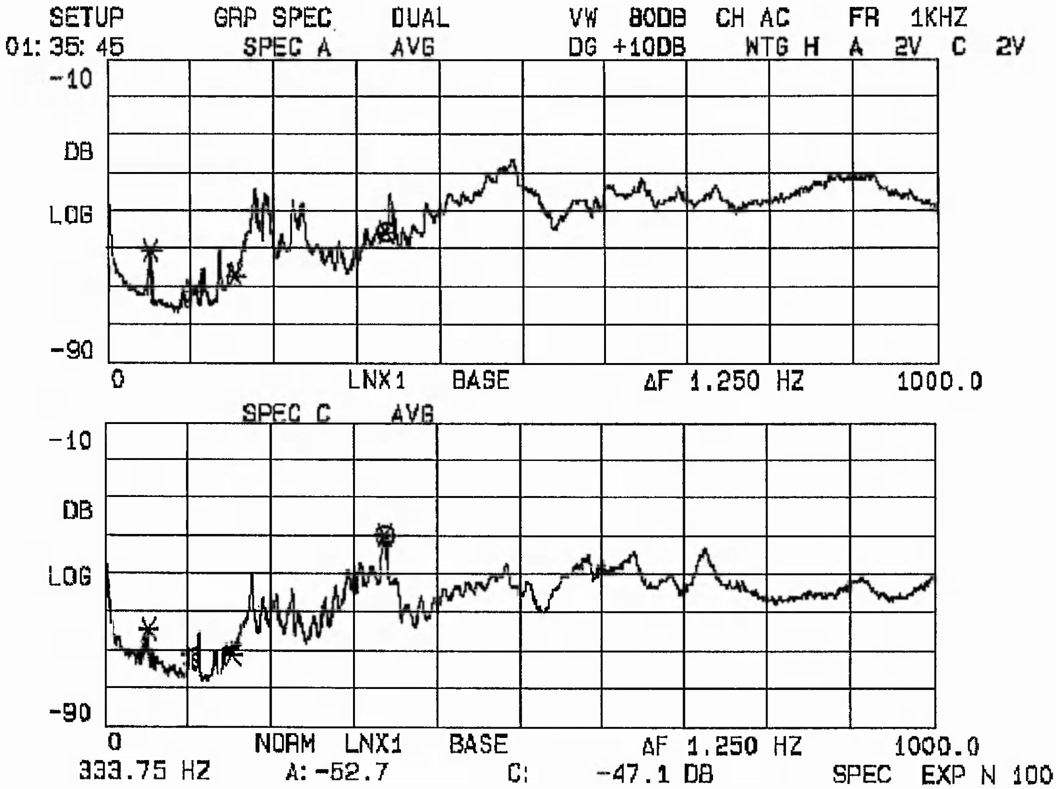


Figure 46. Accelerometer frequency spectra recorded from the marked bearing. x axis top, y axis bottom.

The frequency spectra of the signal from the good bearing, figure 47, were dominated by harmonics of the frequency of rotation, in this case 42.5 Hz. The asterisks in figure 47 indicate the harmonics of this frequency. All the peaks in the frequency range from 16 - 285 Hz, where the bearing element frequencies would be expected, appeared to be harmonics. The shape of these traces, with their relatively widely spaced harmonics, differs from those seen from the marked bearing where there appears to be a much more closely spaced harmonic.

While the form of the signals from the two bearings was observed to be different, a lack of experience in interpreting data of this type meant that it was not possible to accurately

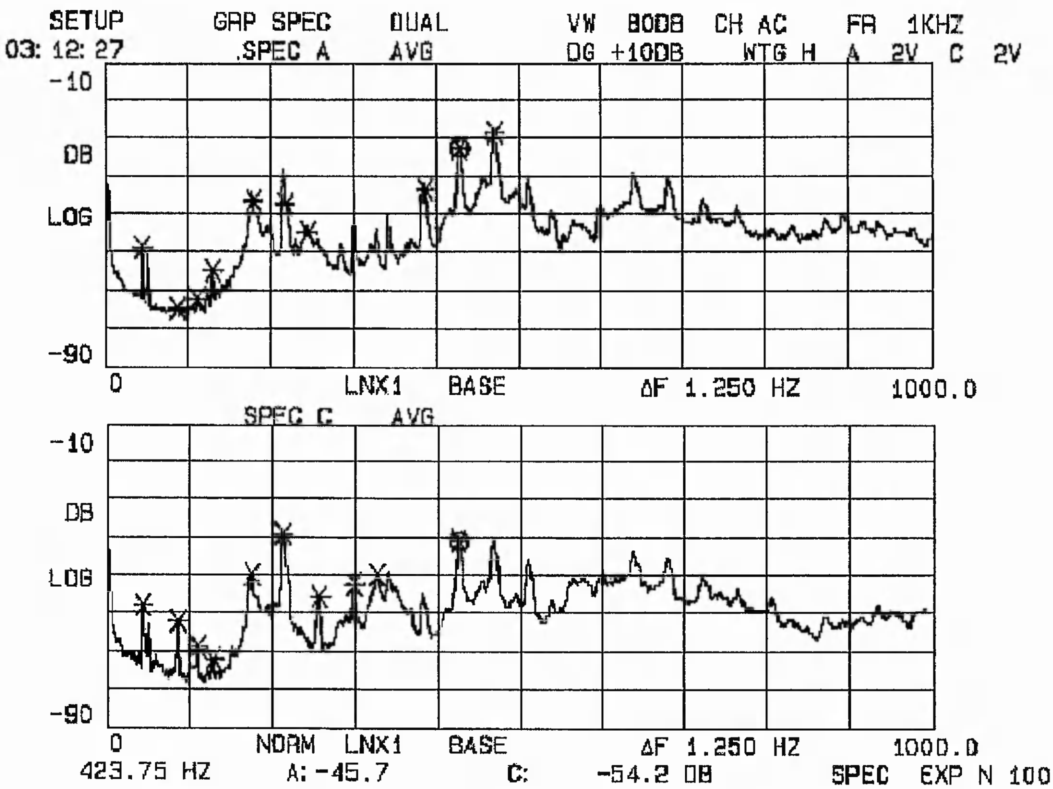


Figure 47. Accelerometer frequency spectra recorded from the unmarked bearing. x axis top, y axis bottom.

determine the source of these differences. In particular in the tests that were conducted it was not possible for the marked bearing to be readily identified as such from it's frequency spectrum. It was not possible to draw a clear conclusion from this set of experiments. It had been hoped that a clear frequency or set of frequencies would be identified that could be used as a target for the subsequent investigation with the IME.

4.3.2 IME signal analysis: unmarked bearing

The initial approach taken to the analysis of the IME signal was based on the published results of work by Philips [54] with an experimental fibre optic based proximity sensor, described in section 2.2.1.3. Since this device monitored displacement within a bearing directly it was considered that its output signal would be resembled by that of the IME.

Figure 49, reproduced from Philips' paper, shows an example of this displacement signal, taken from a bearing in good condition. Each cycle in the waveform results from the flexure of the bearing's outer race when loaded by one of the bearing's rolling elements. Philips describes how the character of this waveform is determined by the surface finish of the raceways and balls as well as their diameter variations. The presence of a defect on one of these surfaces was shown to produce a distinct discontinuity or spike in the displacement signal, illustrated in figure 48.

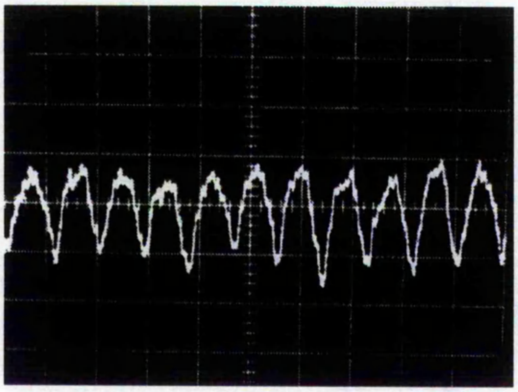


Figure 49. Ball passing waveform measured at the outer race of an undamaged bearing by a fibre optic proximity sensor [54].

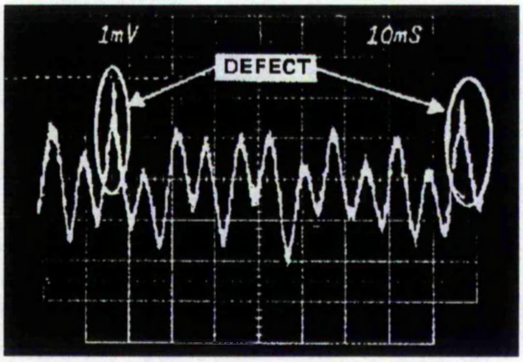


Figure 48. Ball passing waveform from a defective bearing showing the spike associated with a damaged ball [54].

The shaft centre motion monitored by the IME was expected to show a similar relationship to the passage of individual balls or rollers between the races during each rotation. This motion was expected to be at the frequency characteristic of the bearing and to be modified by the presence of a defect on any part of the bearing. The experimental analysis of the IME signal began on this basis.

The two variables in the initial experimental conditions were loading and rotational speed. In a small number of the initial set of tests, a flywheel was used to add some radial load to the experimental system. Experiments were conducted at rotational speeds between 300 and 3000 rpm. Although differences were noted in the resulting signal from the IME at different speeds, these were primarily changes in the shape of the trace. These changes were ascribed to centrifugal effects on the rolling elements of the bearing. The main conclusion drawn from this set of experiments was found to be independent of speed.

Although load was found to be a factor in the results the small load represented by the flywheel did not make a noticeable difference. The NJ202 bearings used were rated up to 9800Kg. The load from the flywheel did not exceed 2Kg. The primary reason for its use was to increase the experimental rig's moment of inertia to limit the possibility angular vibration affecting the results.

Figure 50 is an example of displacement in the vertical axis measured by the IME during a single rotation of the unmarked bearing. Although some information can be gained from

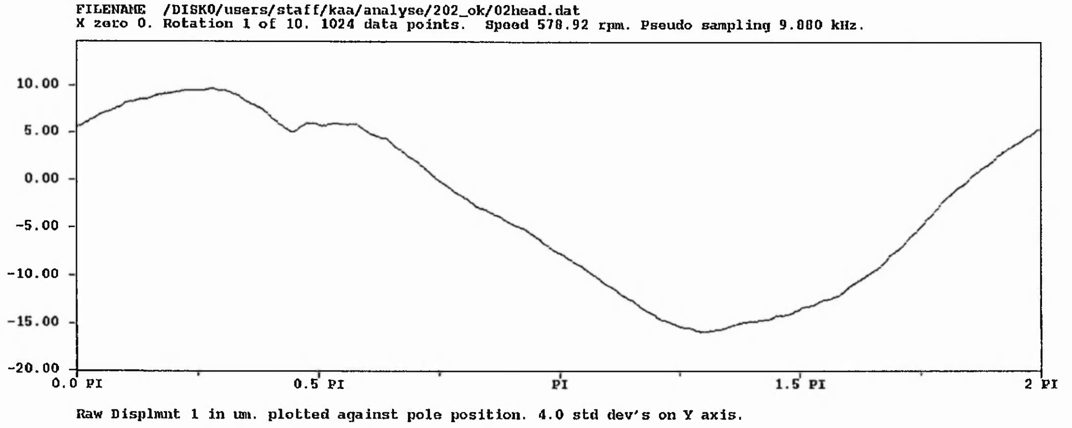


Figure 50. A typical displacement trace for a single rotation measured in the vertical axis from the unmarked NJ202 bearing at 578 rpm.

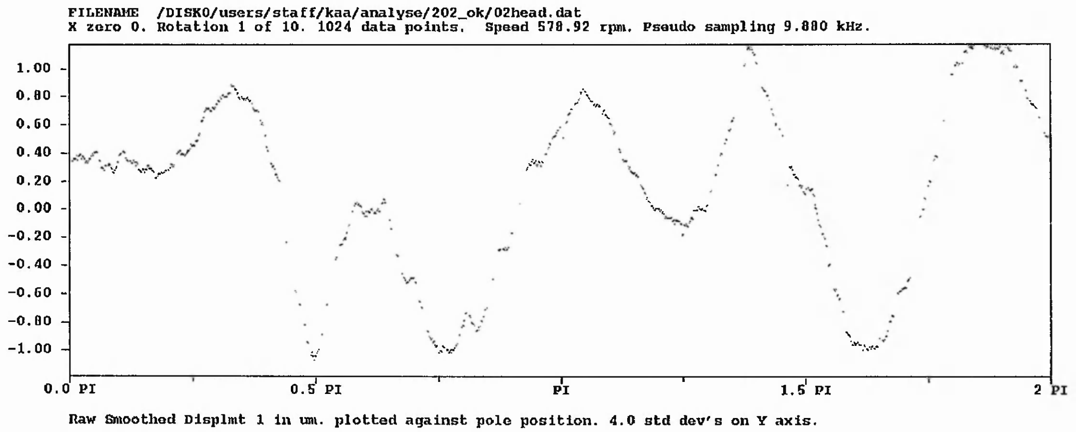


Figure 51. The smoothed displacement calculated from the displacement trace in figure 50.

this raw sinusoidal signal, identifying shaft centre motion requires the removal the calculated average value for each angular position. This process results in the smoothed displacement trace in figure 51.

The picture of shaft centre motion that emerged from the smoothed displacement traces revealed a pattern similar to that which was expected with one important difference. The series of peaks and troughs in figure 51 with their irregular sinusoidal shape and approximately even frequency are strongly suggestive of the ball passing frequency pattern in figure 49. The consistency with which patterns similar to this were seen in the IME signals indicated that they were related to the motion of the internal elements of the bearing. From the bearing frequency equations 6.72 cycles per rotation are to be expected from the passing of the rollers of an NJ202 bearing relative to its inner race. In this example there are clearly five cycles in the trace. In all the experiments conducted with the unmarked bearing the IME showed between approximately 4.5 and 5.5 cycles per rotation in the smoothed displacement signal. The cycles varied in the degree to which they were pronounced, both between rotations recorded at different times and within the same set of rotations.

The analysis and display program contained a library of digital signal processing algorithms. The FFT was used extensively to generate frequency spectra from the IME signal. Applying this technique to the data shown in figure 51 produced the spectra shown in figure 52. This is typical of the results produced in the initial stage of the data analysis. Spectra such as these did not produce a definite result proving that the IME signal was related to any of the frequencies in table 3 as they were understood at the time.

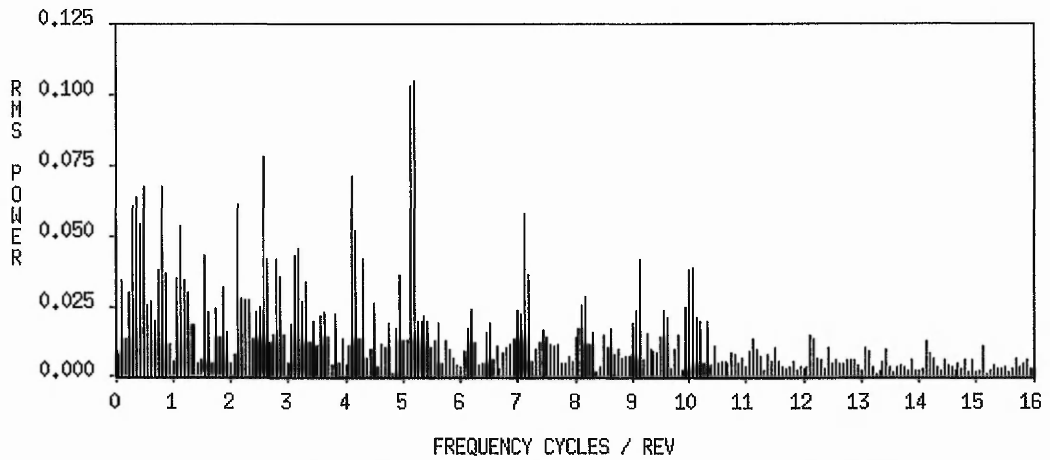


Figure 52. An example of the frequency spectrum derived from the smoothed displacement measured from the unmarked bearing.

4.3.3 IME signal analysis: marked bearing

In contrast to the lack of a clearly definable signal component from the unmarked bearing, the IME signal from the marked bearing showed that the sensor was able to clearly able to detect the induced defect. Figure 54 shows an example of the signal as it was initially recorded from the marked bearing. The overall shape of the signal is still sinusoidal, however a sharp impulse is clearly visible in the trace. Another feature of the signal from the marked bearing was the higher frequency component visible to the right of the impulse.

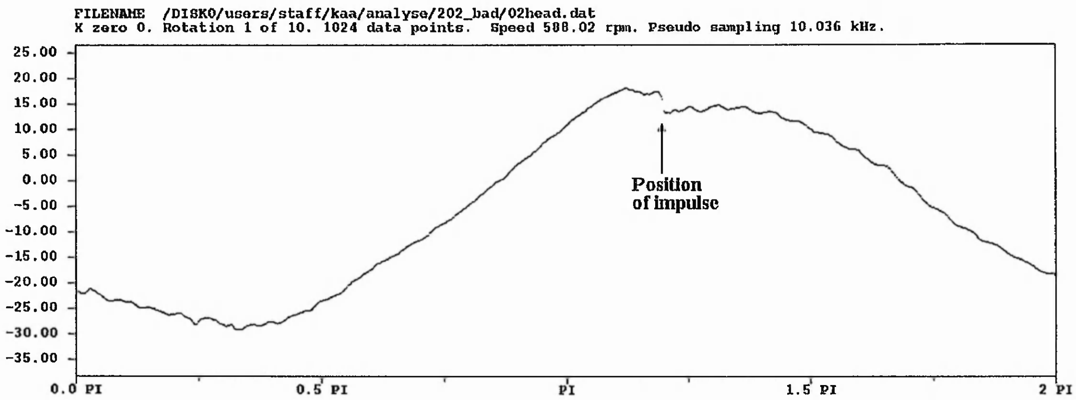


Figure 54. The y axis displacement measured by the IME from a single rotation of marked NJ202 bearing showing the position of the impulse due to the damage.

The smoothed displacement in figure 53 shows this clearly. It can also be seen that this frequency is superimposed on the waveform with the frequency of approximately 4.5 times per revolution also seen from the unmarked bearing.

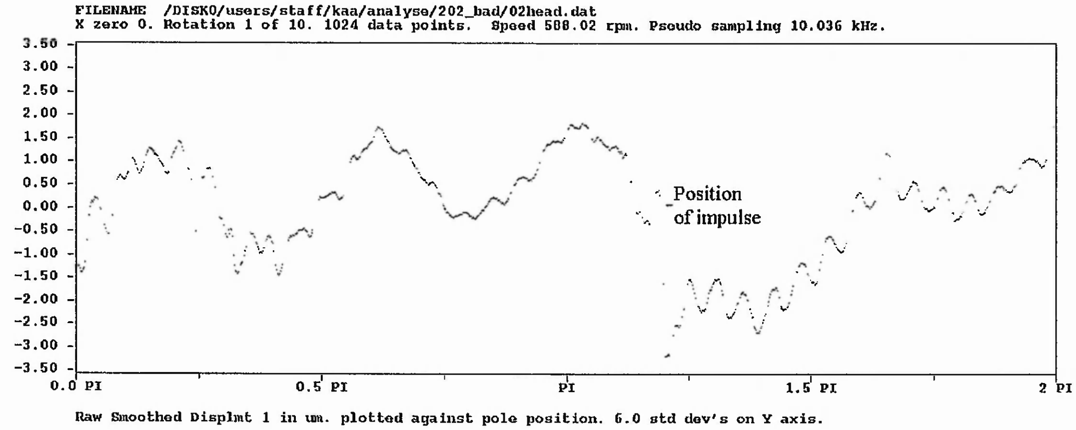
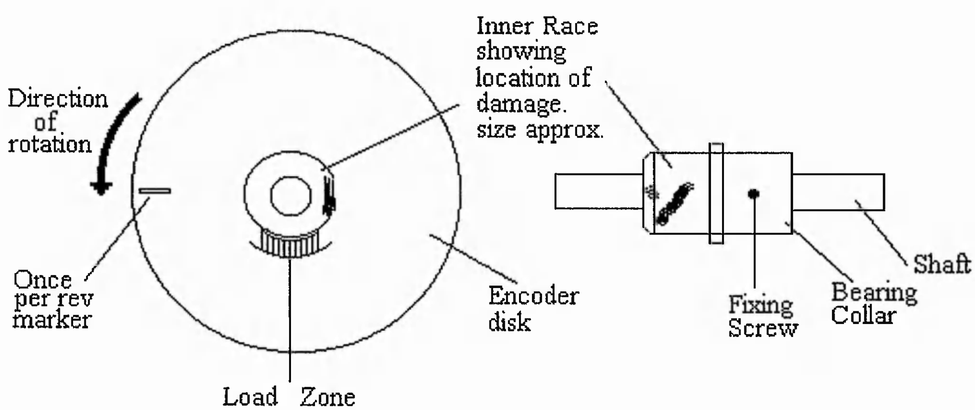


Figure 53. Smoothed displacement measured from the bearing with a defect.

The interpretation placed on these results was that the impulse, caused by the defect, resulted in a response from the shaft-bearing system which manifested itself as the higher frequency.

The first part of this interpretation was tested by changing the position of the defect relative to the disk. As the start of data collection was synchronised with one point on the disk, if the impulse were related to the defect it would occur at a different angular position in each rotation. In a set of four experiments the defect on the inner race of the bearing was positioned at 0°, 90° 180° and 270° offsets to the once per revolution marker on the encoder disk. In each case the impulse was observed to shift angular position in the displacement trace. Figure 55 illustrates the experiment. The defect is shown offset 180° from the once per revolution marker with the shaft rotating anti-clockwise. The IME system starts recording data when the once per revolution marker is in the 9 o'clock position. It will rotate 270° degrees before the defect enters the bearing's load zone. The smoothed displacement trace in figure 55 shows that the impulse occurred at the position



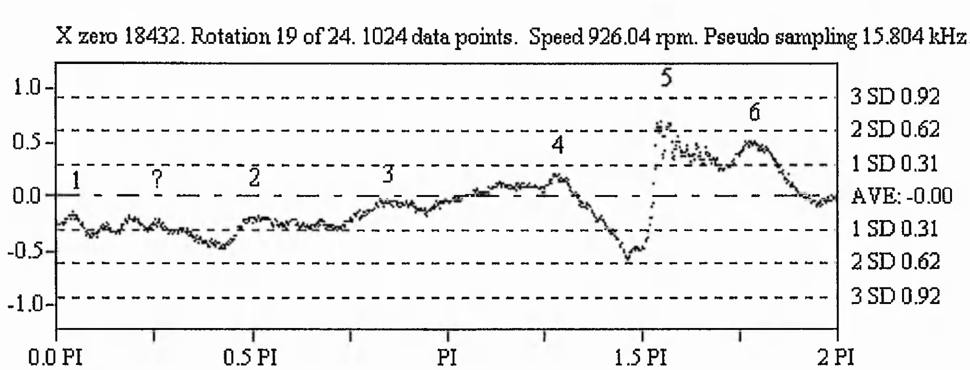


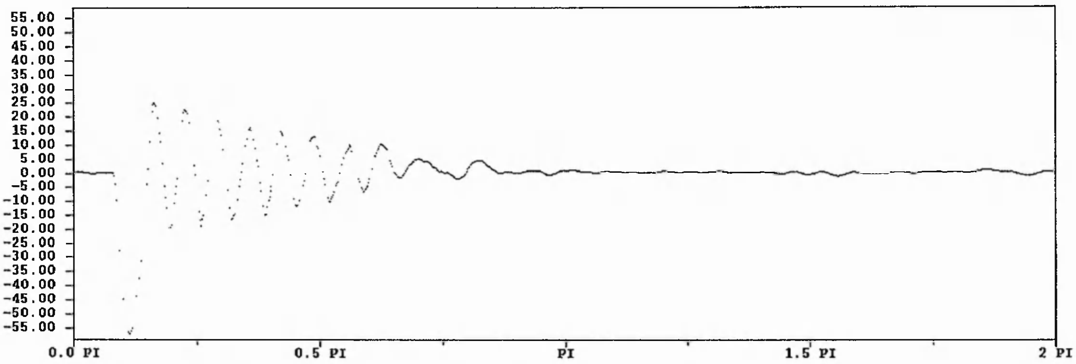
Figure 55. Position of defect relative to angular position of impulse in the smoothed displacement trace.

predicted.

In the course of the project this was a more important experimental result than was immediately realised. For the first time it had been shown that the IME system was definitely able to resolve a bearing defect from shaft motion. The key fact however was that the bearing defect created an impulse in the displacement signal. This impulse was related to the passage of the defect through the load zone of the bearing. It did not occur at the characteristic inner race defect frequency. For this to have happened an impulse would have to have been generated each time the defect passed a roller. Although there are some peaks in the smoothed displacement trace, for example those shown numbered in figure 55, these did not occur consistently in the way the impulse did.

4.3.4 Shaft impulse response experiment

Despite the lack of a clear defect frequency the focus of experimentation remained on detecting specific frequencies within the displacement signal. The impulse due to the bearing defect clearly excited a resonant response in the shaft. It was reasoned that exciting the shaft externally should produce a similar response. A small plastic hammer was used to produce a sharp, impulsive, blow to the shaft of the experimental rig as close a possible to the IME disk. The shaft was supported by the unmarked bearing. Figure 56 shows an example of the response of the shaft displacement to this event. The top trace, channel 1, is from the vertical axis, the lower trace, channel 2, is from the horizontal. Both traces clearly show how the shaft is displaced by the initial impact of the hammer blow followed by a period of damped oscillation. The frequency of this oscillation can be compared with that shown in figures 54 and 53, examples from the defect impulse. Note also how the



Raw Smoothed Displmt 1 in um. plotted against pole position. 60.0 std dev's on Y axis.

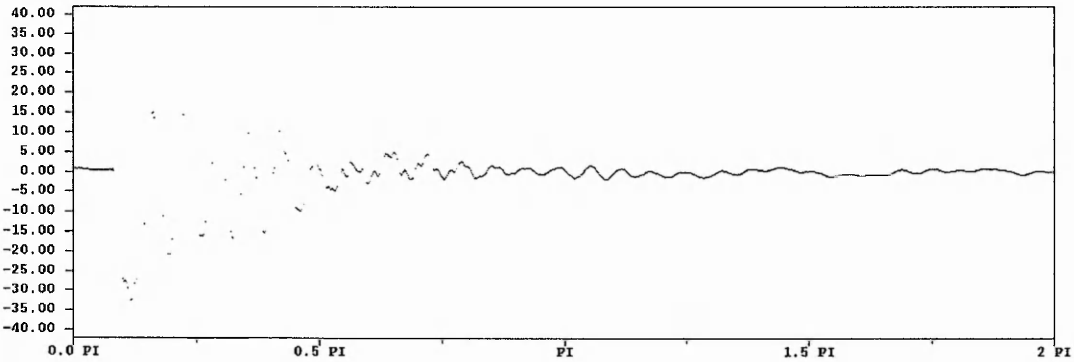


Figure 56. Displacement traces showing the response of the shaft to a short, sharp, impulsive hammer blow.

oscillation seen in figure 54, although very much less pronounced, dies away in a similar manner suggesting a damped response.

Plotting the displacement traces from this set of experiments against each other gave further evidence that the IME signal provided a correct picture of shaft motion. In each of the hammer blow experiments the direction of the blow was different. The shaft was struck from directly above, from the horizontal and at 45° to the horizontal on both sides. The aim of this was to see if the direction of the blow could be measured accurately by the IME. In all cases the orbit plot produced by combining the displacement data from the IME's two axes of measurement allowed the direction of the force applied by the hammer blow to be accurate determined.

Figure 57 is an example of the orbit plot from the one of these experiments. It is formed from the displacement data in figure 56. In this case the hammer blow was directed against the shaft from the right of the rig at 45° to the horizontal. The orbit plot clearly shows the motion of the shaft's response to this blow as it moves sharply away to the lower left. Note how the position of each displacement point can be seen as an extending smooth curve which turns over on itself before retracing to the point where the blow was struck. As the displacement points move into the lower left portion of the graph they become closer together. As they move back, the points become more widely spaced again until they move

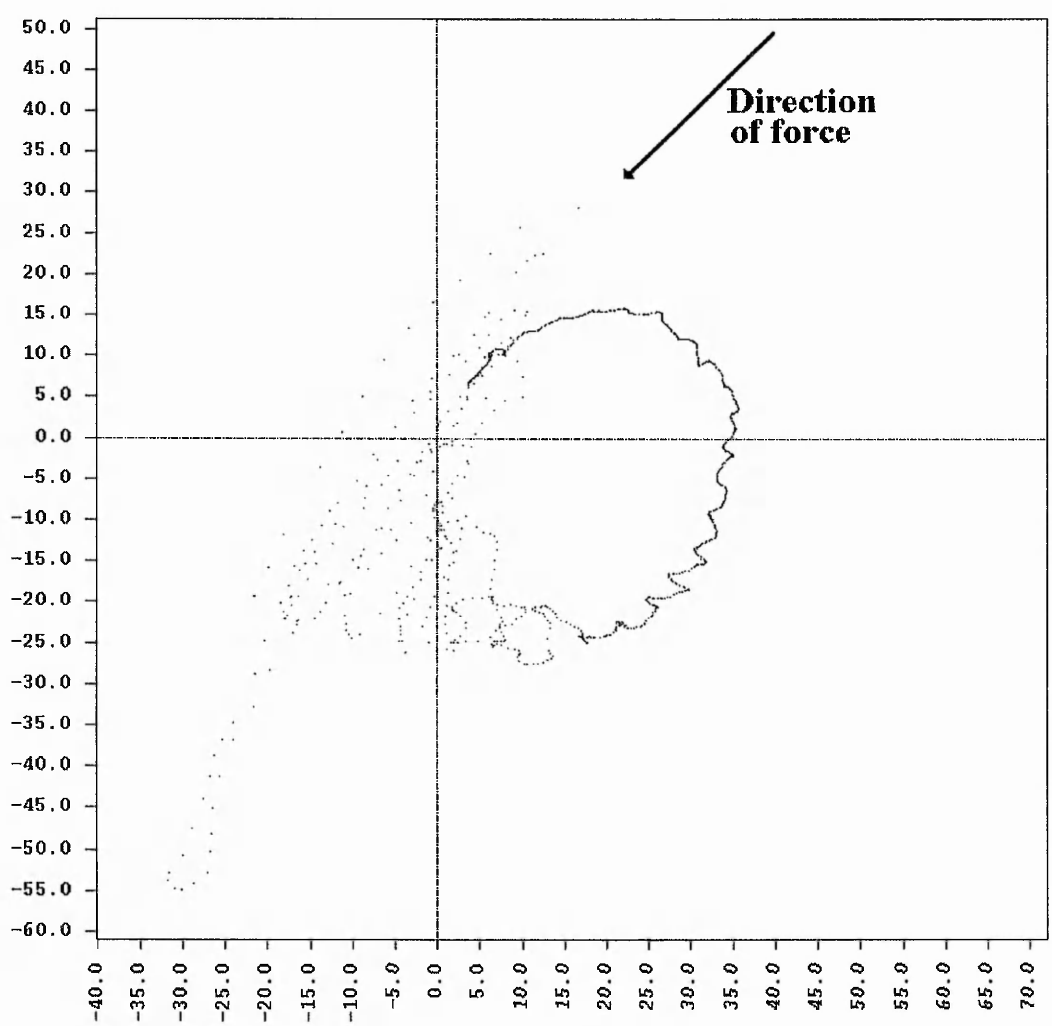


Figure 57. The impulse response of the shaft in two dimensions.

beyond the point at which the blow was struck. The reverse then happens with the points becoming closer together as the trace extends into the upper right portion of the graph. The distance between each point is a measure of the displacement of the shaft centre in each angular measurement interval. As the shaft centre accelerates it is displaced more in each angular interval so the measured points are further apart. The pattern observed is therefore consistent with the pattern of acceleration which would be expected from the simple harmonic motion of the shaft vibration.

If the bearing defect were exciting the shaft in the same manner as the hammer blow the frequency at which the shaft responded to the two events should be similar. After the displacement traces from the hammer blows had been studied, frequency spectra were generated. These were compared with spectra from the marked bearing.

Figure 58 is an example of the frequency spectra obtained from the marked bearing. The frequency range in the figure, from 0 to 8 cycles per rotation, is the range in which the bearing defect frequency should lie. As has been explained no features were identified which could be used to differentiate spectra like this from spectra produced by the unmarked bearing, figure 52.

When the spectra from the impulse response experiments were studied they showed that the frequency of the shaft's response lay outside the range expected for bearing defects. Figure 59 shows the FFT produced from the IME data file of which figure 56 was a part. When this data was studied, two features immediately stood out. The impulse response of the shaft has produced a large, broad, peak between 26 and 34 cycles per revolution. The

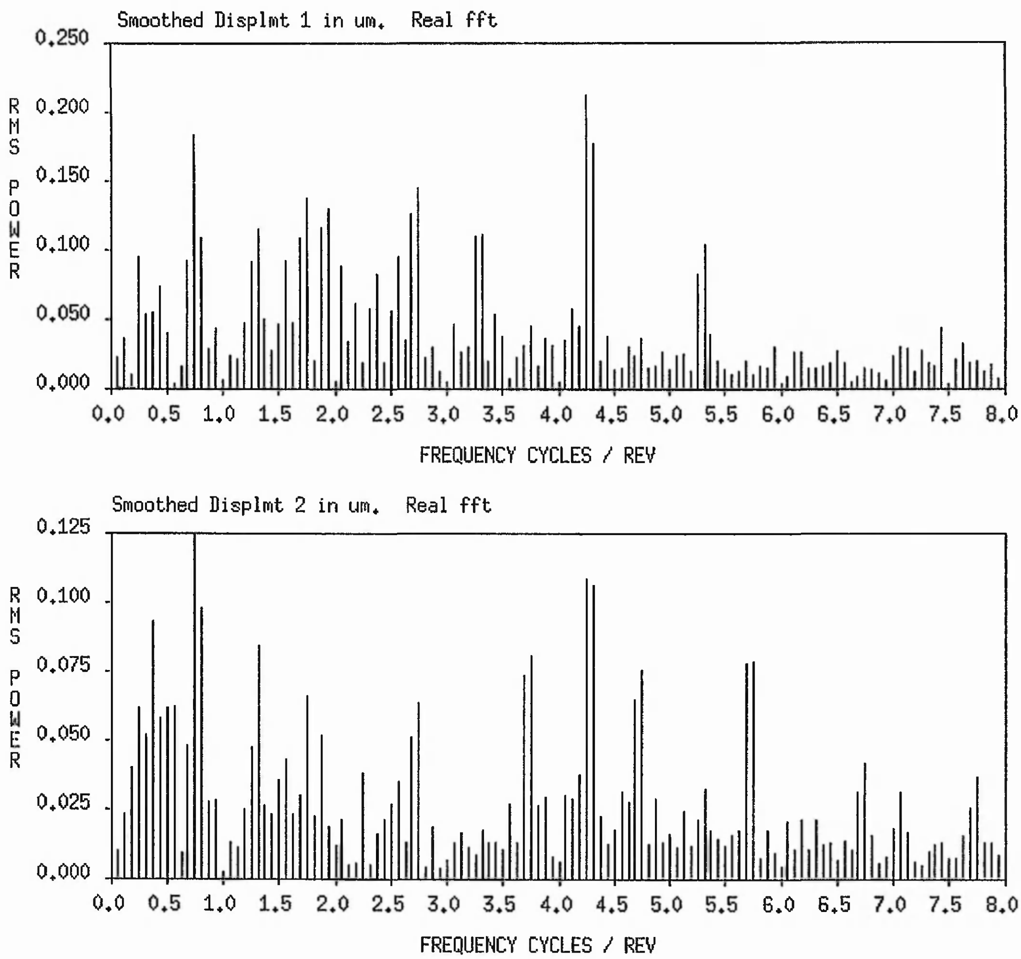
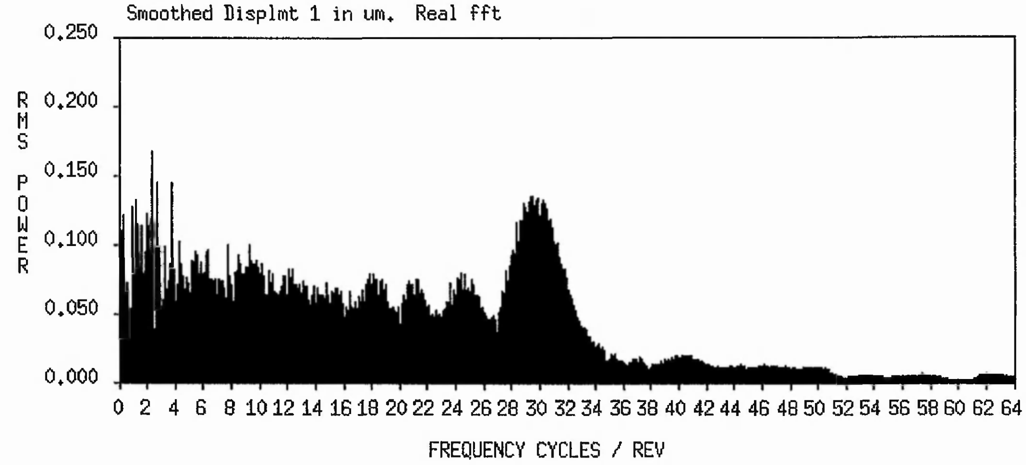


Figure 58. Typical smoothed displacement frequency spectra from both measurement axes of the IME monitoring the marked bearing.

overall spectra is also more energetic. Compare the rms power level across the frequency range in figure 59 with the level across the same frequency range from the defective bearing, shown in figure 60. The response to the impulse is not a sine wave, but simple harmonic motion in which the frequency and amplitude change from a maxima to zero. The increased rms level represents this decreasing impulse as the energy spreads out across the frequency range.

The small peak in the spectra from the defective bearing in the same region as the large



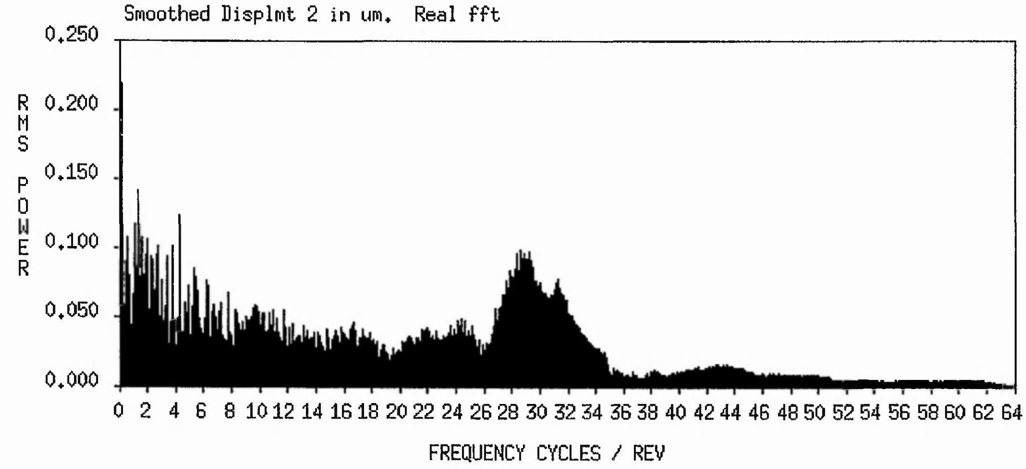


Figure 59. Frequency spectra resulting for the shaft impulse response experiment

peak from the impulse test suggests that the defect is exciting the same impulse response, as was initially deduced.

It was concluded at the end of the shaft response experiments that the IME was able to detect the motion of the shaft due to the effect of the test defect. This motion took the form of a sharp impulsive acceleration of the shaft centre which was caused the shaft to oscillate. The oscillation decayed away over approximately one rotation. It was thought that the frequency of the impulse response would be the indicator of the defect and that this

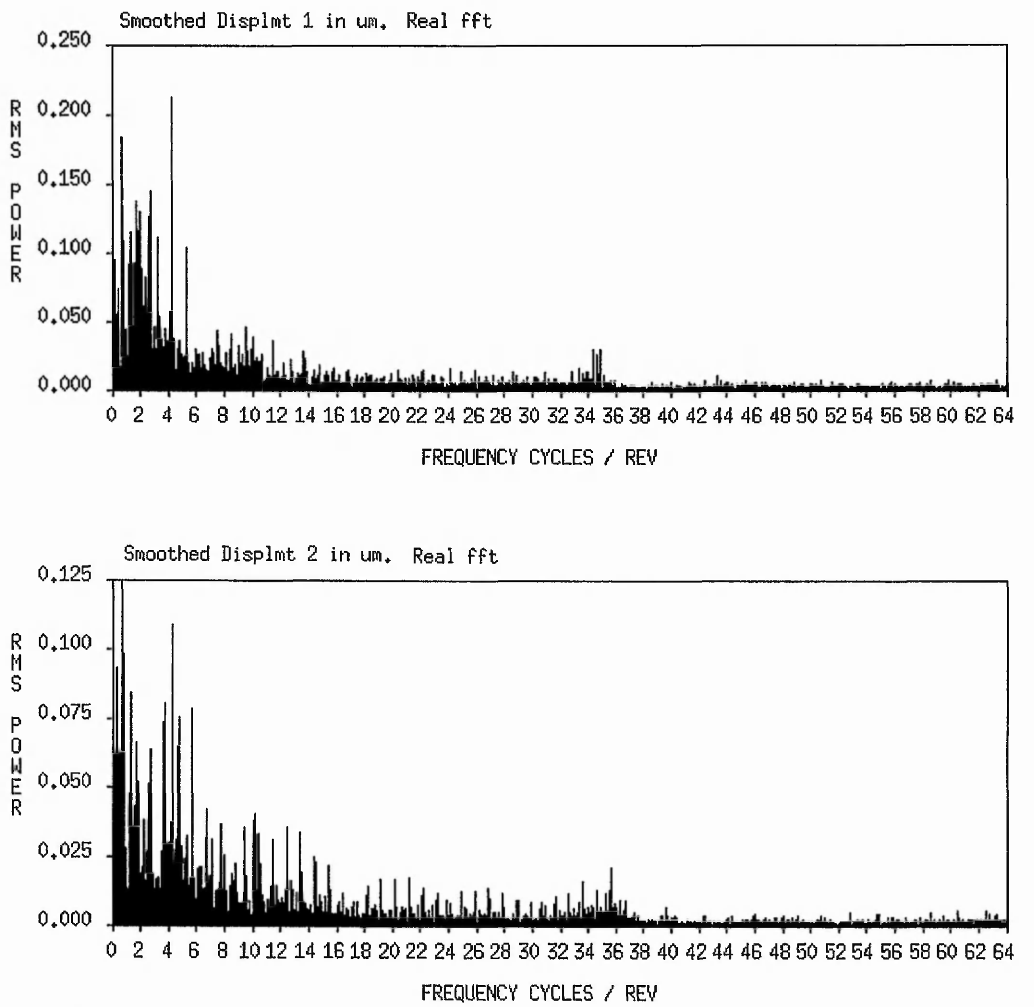


Figure 60. Increasing the scale of the frequency spectra in figure 58 shows a small peak between 34 and 36 cycles per revolution above the frequency range expected for bearing defects.

frequency could be identified from a frequency spectrum. Subsequent work was based around these conclusions. Essentially this meant that frequency analysis became the main tool used in trying to interpret the signals with the emphasis placed on the region between 26 and 34 cycles per rotation.

The problem with this approach was that it was based on detecting the effect of the impulse rather than the impulse itself. In practice almost all attempts to classify the peaks of the frequency spectrum into two classes failed. Some differences seemed to come from the way the rig was adjusted. The retaining screw pressure on the outer race and misalignment of the shaft relative to the bearing test pillar both affected the shape of the displacement signal and hence its frequency spectrum. The main conclusion drawn from the observations made was that there was not a clearly visible difference between the frequency spectra from the unmarked and the marked bearing.

On one occasion it seemed that in a series of tests that the ball passing waveform was clearly present. A paper was written based on this set of results [82]. An example of the waveform is shown in figure 61. This smoothed displacement signal from the sensor was similar to that described as due to ball passing in other experimental work involving novel displacement sensors [54,57].

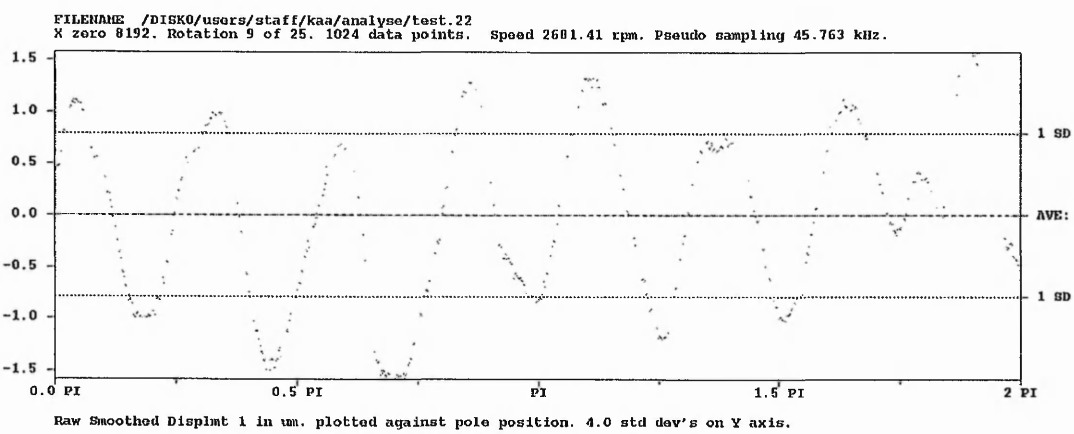


Figure 61. Smoothed displacement trace from defective bearing showing a frequency at 7.5 times per rotation thought to be the ball passing frequency.

The frequency spectra of this set of results all showed that this frequency agreed exactly with that calculated from equation 2. An example of this is shown in figure 62. The explanation of this is complicated by the fact that the dimensions used to calculate this

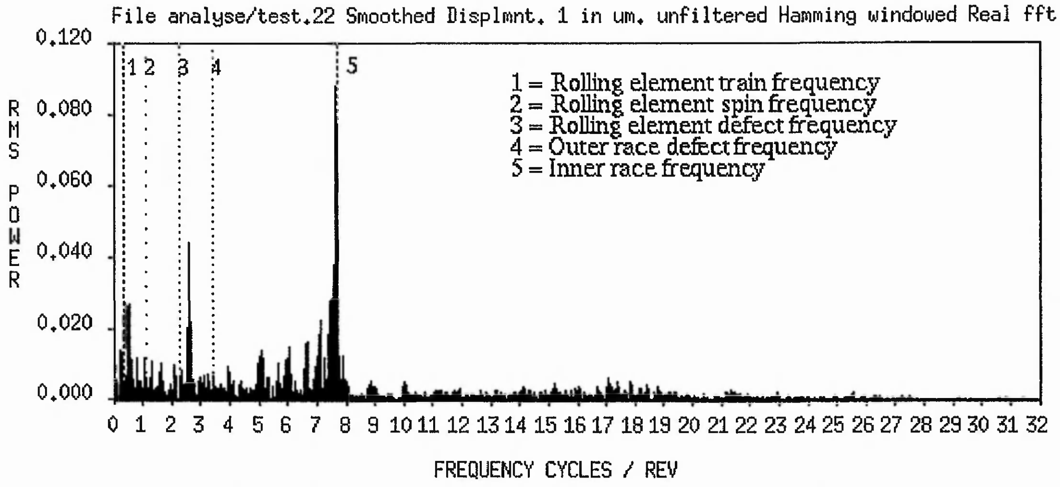


Figure 62. Frequency spectrum of smoothed displacement signal showing a peak at the position calculated for the inner race defect frequency.

frequency which were supplied by a bearing manufacturer were incorrect. The value given by these dimensions was equivalent to 7.5 cycles per revolution. The correct figure for the frequency given by an inner race defect was 6.7 cycles per revolution for a bearing of this type. Although allowing for slippage, 7.5 is quite close to the calculated figure, this still does not explain the way that this set of 8 data files recorded this frequency when others before and after did not. The reason thought likely to account for this is the adjustment of the rig which may have placed a load in a different position for this set of data.

4.4 Initial experimental conclusions

At the end of the initial set of experiments described in this chapter it was concluded that while the IME was able to detect changes in displacement related to bearing condition, there was no clear interpretation framework for the sensor signal. The hypothesis that the ball passing frequency would be easily detectable had not been proven. Instead it had been shown that this frequency was almost never present in the IME signal.

The failure to be able to consistently repeat the results shown in figures 61 and 62 and the overall failure of frequency analysis to provide a clear differentiation between the marked and unmarked bearings led to further research into the conventional methods of bearing condition monitoring. The aim was to look at the models used to explain the form of signals from conventional sensors and then develop a similar model for the IME which would predict where and in what form the presence of a defect could be detected.

Chapter 5

Creation of a model to explain the effect of a bearing defect on the IME signal

5. Creation of a model to explain the effect of a bearing defect on the IME signal

5.1 Introduction

Chapter 4 explained the nature of the IME signal and described how initial experiments had shown that it could be directly related to shaft motion. These experiments had not been able to show the expected relationship between the sensor signal and an induced bearing defect. This chapter describes how a model to explain the effect of a bearing defect on the IME signal was developed based on the initial experimentation and published work of other researchers in this field. The experiments used to test this model are then presented.

5.2 The McFadden and Smith model for single defect bearing vibration

The starting point for the development of a model to explain the IME signal was an examination of work done by McFadden and Smith. They set out to explain the frequency spectra from bearings monitored by accelerometers using the HFRT. For a description of the HFRT see section 2.3.2.7. In a review of the HFRT McFadden and Smith concluded that the method of obtaining a vibration spectrum from an accelerometer was well established, but that the understanding of the spectrum remained strictly limited [60]. In order to provide the increased understanding required, they developed a model to describe the vibration produced by a single point defect on the inner race of a bearing operating under constant radial load [64]. This model incorporates the effects of the load distribution in the bearing, the motion of the rolling elements, the variations in the transfer function and the decay of the resonances excited by the defect. This model was subsequently extended to describe the vibration caused by multiple point defects [65].

The feature of their model relevant to the IME sensor is the use of the load zone to redefine the frequencies expected from a defective bearing. The geometrically based defect frequency equations given in section 2.3.1 and used in the interpretation of vibration spectra from accelerometers are only directly applicable when the bearing is subject to a unit load distributed evenly about its whole circumference. This condition applies to axially loaded thrust bearings, but not to radially loaded deep groove ball bearings or roller bearings. The model treats the vibration produced by the impact of a single point defect with another bearing surface as an impulse, represented by the impulse function $\delta(t)$. The severity of the defect is represented by multiplying the impulse function by the constant d_o. The frequency of the impulses is determined by the location of the defect. For an inner race defect this frequency, f_d, is given by equation 2. Figure 63(a) shows how the vibration due to the defect is modelled as a series of impulses of amplitude d_o extending to infinity in both directions from any point in time with a separation, T_d, which is the reciprocal of the frequency f_d.

The fourier transform of an infinite series of impulses is also an infinite series of impulses. Figure 63(b) shows this series of impulses of amplitude $d_o f_d$ extending to infinity in both directions with a separation between the impulses equal to f_d.

For a radially loaded bearing the defect frequencies are modulated by the extent of the

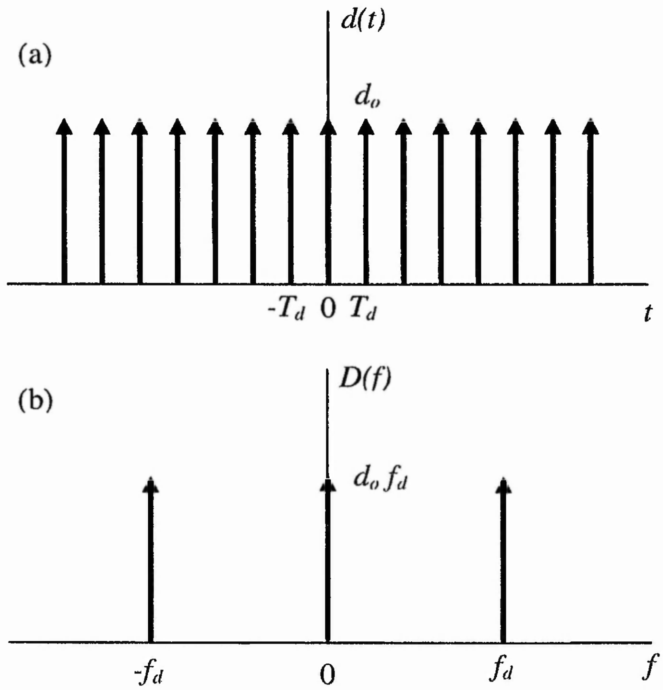


Figure 63. The impulses produced by an inner race defect under unit load.
(a) Time history; (b) spectrum [64].

load zone. Impulses are only generated when the bearing elements strike each other under load. The greater the load, the greater the magnitude of the impulse. The distribution of the load around the circumference of a rolling element bearing is defined approximately by the Stribeck equation [64]:

$$q(\theta) = q_o[1 - (1/2\varepsilon)(1 - \cos\theta)]^n \tag{11}$$

q_o is the maximum load intensity, $q(\theta)$ is the load at angle θ, ε is the load distribution factor and n = 3/2 for ball bearings and 10/9 for roller bearings. θ_{max} is the angle at which q becomes equal to zero defining the angular extent of the load zone. The terms q_o, ε and θ_{max} are all functions of the diametral clearance of the bearing and the applied

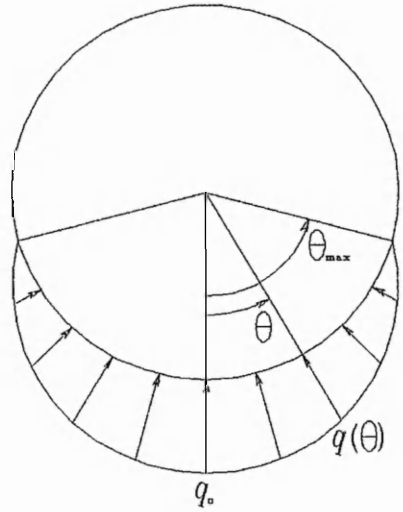


Figure 64. The load distribution in a bearing under radial load [64].

load. For a bearing with positive clearance, $\varepsilon < 0.5$ and $\theta_{max} < \pi/2$. In this case the load distribution is of the form shown in figure 64.

If the inner race of a bearing is rotating at a constant speed f_s revolutions per second, then the instantaneous load at a point on the inner race as a function of time, *q(t)*, is obtained by substituting $2\pi f_s t$ for θ in equation 11 which results in equation 12.

$$q(t) = \begin{cases} q_o\left[1 - (1/2\varepsilon)(1 - \cos\theta)\right]^n & for\ |\theta| < \theta_{max} \\ 0 & elsewhere \end{cases}$$

The function *q(t)*, is periodic as a given point on the inner race of the bearing will pass through the load zone with each revolution of the shaft. Figure 65 shows this function as a series of load curves extending to infinity in both directions from any point in time. The load curves in this illustration are idealised as semi-circles. In actuality their form is determined by the Stribeck equation. The separation of the centre of each load

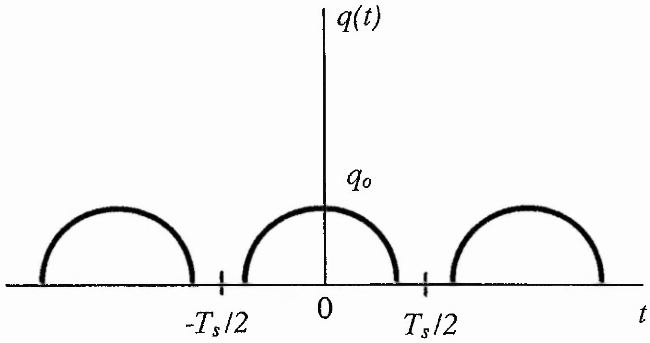


Figure 65. The time history of the load on the inner race of a bearing under a radial load [64].

distribution curve is equal to the period T_s, the reciprocal of the shaft rotation frequency f_s.

If it is assumed that the amplitude of the impulse caused by a defect is proportional to the load on the rolling element when it strikes the defect then the amplitude of the impulses is given by multiplying the amplitude under unit load *d(t)* by the actual load distribution *q(t)*. Figure 66 illustrates the product *d(t)q(t)* graphically. The frequency of *d(t)* is not necessarily an integer multiple of the frequency *q(t)* as the relationship between their frequencies is determined by the bearing geometry. Hence the pattern of impulses produced by the product *d(t)q(t)* is different from one shaft revolution to the next.

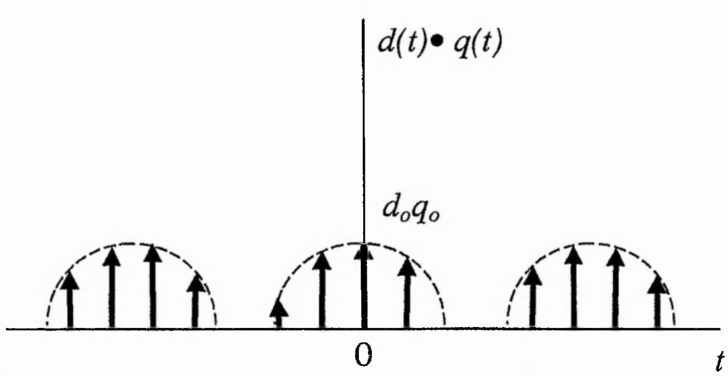


Figure 66. The time history of impulses produced by an inner race defect under a radial load. [64]

In their model McFadden and Smith went on to consider the effect of the vibration due to the defect on the machine structure and its transmission to the accelerometer. As the principal advantage of the IME is that the motion due to a defect is not required to pass through the machine structure this part of their model was not directly applicable. A model for the motion of the shaft centre due to a defect had to be developed to replace it.

5.3 The application of the McFadden and Smith model to the IME

The McFadden and Smith model can be applied to the IME by substituting the measurement of time with the measurement of angle. For a single inner race defect in a bearing with a uniformly distributed unit load a number of impulses equal to the inner race defect frequency will be produced at equal angular intervals during each shaft rotation. The angular interval in radians between impulses, θ_d, is given by equation 13. Substituting the number of grating lines on the IME disk for 2π in equation 13 allows the calculation of the distance between each impulse in the actual IME signal.

$$\theta_d = \frac{2\pi}{\textit{inner race defect frequency}} \tag{13}$$

The function *d(t)* defined by McFadden and Smith is replaced for the IME by $d(\theta)$, shown graphically in figure 68.

The actual distribution of load in a bearing is defined in angular terms in the McFadden and Smith model, but converted with reference to time to calculate its effect on the

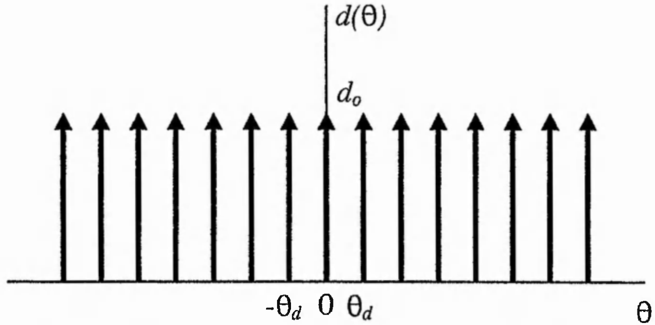


Figure 68. The angular distribution of impulses produced by an inner race defect under unit load.

impulse frequency. For the IME the angular extent of the load zone can be used directly since the Stribeck equation defines load as a function of angle. The load distribution curves defined by this function are shown in figure 67.

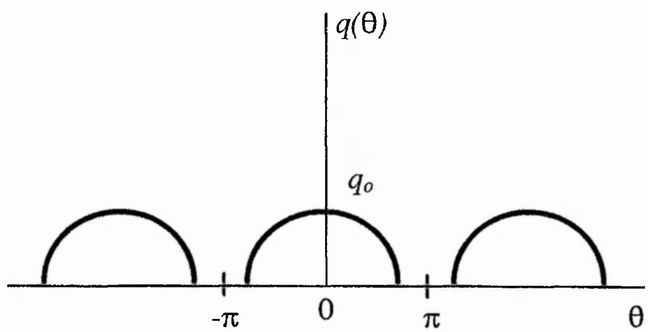


Figure 67. Load as a function of angle for a point on the inner race of a bearing under a radial load.

The magnitude of the impulse produced by a defect is proportional to load. For the IME the series of impulses produced under unit load $d(\theta)$ is multiplied by the actual load distribution $q(\theta)$ to find the actual magnitude at any angular position. The result of $d(\theta)q(\theta)$ is shown in figure 69.

By applying the McFadden and Smith model to the IME in this way it can be predicted that the number of impulses the sensor would detect in each rotation will depend on the angular extent of the load zone. The magnitude of each impulse will depend on the

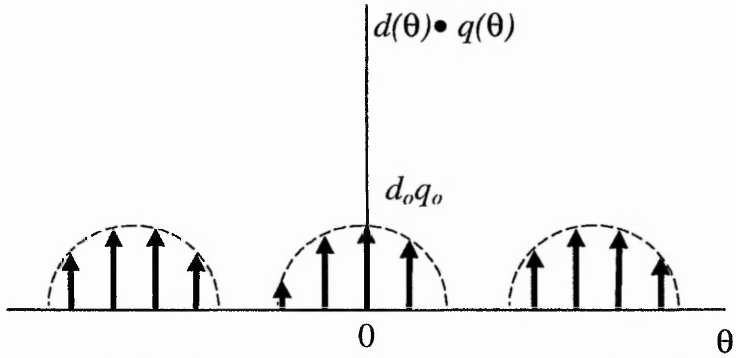


Figure 69. The effect of angular load distribution on the magnitude of impulses produced by an inner defect under a radial load.

position of the bearing elements between which the impulse was generated in the load zone. The inner race defect frequency only accurately predicts the number of impulses which would be generated from each shaft rotation in the case where $\varepsilon = 1.0$ and the load zone extends throughout the complete circumference of the bearing. In other cases the IME would record a number of impulses given by equation14.

$$number\ of\ impulses = inner\ race\ defect\ frequency \times \frac{2\theta_{\max}}{2\pi} \tag{14}$$

Having used the McFadden and Smith model to predict how many impulses the IME should record in each rotation it was then necessary to develop a model for the motion of the shaft centre in response to each impulse.

5.4 Motion of shaft centre due to an inner race defect

The causes of shaft centre motion can be divided between two sources. Those which originate as a result of external systems to which the shaft is fixed, such as dynamic imbalance of a flywheel or gear meshing, and those which are due to the relative motion of the bearing elements. In the development of this model only motion due to the bearing elements is considered as this is the primary indicator of bearing condition. However it will be shown that the model provides the ability for the two types of motion to be discriminated on the basis of the angular interval of the signal.

For an undamaged bearing the motion of the shaft centre is due to the varying distribution of the rolling elements in the load zone supporting the inner race as it rotates. If the surfaces of the races and elements were perfect, shaft centre motion would result purely from the changes in loading due to the combination of positions that elements were able to take up in the load zone. The resulting differences in load distribution would produce small, regular and repeating changes in the shaft centre position as each rolling element passed through the maximum load position. In practice the surface irregularities due to manufacturing errors of the bearing components combine with the load distribution variation to produce more complex motion which repeats over a longer period. These irregularities are of two types and have different effects [59,83]. Surface finish errors, which take the form of a roughening of the contact surfaces of the bearing, can be expected to cause random motion of the shaft centre. Geometrical errors, which take the form of varying diameter rolling elements and races with a waviness of a few cycles around the circumference, will result in periodic motion of the shaft centre. This motion will repeat over the number of rotations taken for the elements to return to any given initial position relative to the load zone.

Investigation of the effect of these surface irregularities on the results of conventional vibrational analysis has found that the vibration spectrum of a normal bearing can have a similar pattern to that of a bearing with minor damage [59]. This occurs because an irregularly shaped rolling element can excite the machine structure at the same frequency as one with a defect. As the IME senses the motion of the bearing via the shaft directly, these geometric irregularities present a lesser problem. The motion of the shaft due to rounded surface irregularities is different to that expected from bearing defects, which are generally angular.

Fatigue in the load zone between the elements and the races of a loaded bearing can lead to loss of material from the bearing surfaces. Spall defects of this type produce sharp sided pits on the bearing's contact surfaces and angular particles contaminating the bearing lubricant. In the initial stage of this type of bearing damage, spalls are localised and most of the bearing's contact surfaces remain undamaged. Following the initial spalling, wear accumulates rapidly due to the damage caused by rolling surfaces impacting the spalls and the freed particles. The angular nature of the spalls and debris particles allows their effect on the motion of the shaft centre to be modelled as the transition from a smooth surface, where the race-roller contact is undamaged, to an irregular surface where damage has occurred.

Figure 70 is a diagram illustrating the model developed to predict the motion of the shaft centre in response to an inner race spall. Defects in other locations can be treated in a similar way. The size of the defect is exaggerated for clarity and the effect of other elements on the depth of penetration is ignored. This model has been loosely based on an explanation of the shock pulse bearing monitoring method by Collacott [4]. In figure 70 D_i is the outside diameter of the inner race, D_{re} is the diameter of the rolling element, n_i is the rotational speed of the inner race, y is the depth of penetration of rolling element into the spall, x is the width of the spall A - B, V_{re} is the velocity of the rolling element at point of impact, A is the point on the edge of spall about which the rolling element pivots, B is the point of impact between the rolling element and inner race, R is the centre of rotation and $q(\theta)$ is the load acting through point B.

The model assumes that from supporting the inner race's undamaged surface a rolling element will rotate around point A into the spall as the inner race pivots down to impact on the element at point B. The force of this impact is determined by the velocity of the rolling element at the point of impact, V_{re}, and the load acting through that point $q(\theta)$. The magnitude of $q(\theta)$ is dependent on the angular position of B in the load zone. It is this impact which generates the stress wave that propagates as an impulsive vibration through the machine structure. The depth of penetration y depends on the width of the spall, x, and the diameter of the rolling element, D_{re}, combined with the force of the impact. The motion of the bearing continues and the inner race pivots about point B to return to being properly supported on its undamaged surface by the rolling element. Slipping relative to the outer race may occur at this point.

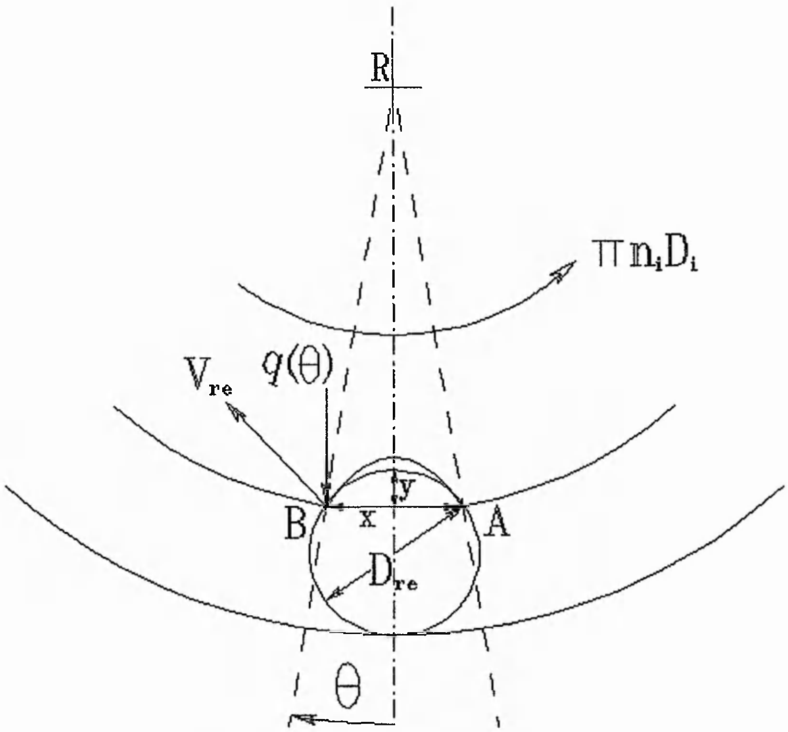


Figure 70. Model for the effect of an inner race spall on shaft centre motion.

The motion of the shaft centre which results from the defect is related primarily to the pivoting of the inner race at points *A* and *B* due to penetration of the element into the race rather than to the impulsive vibration. As the inner race pivots around point *A* the shaft centre *R* will move as the end of the radius *R* - *A*. This motion, comprising horizontal and vertical components, is directly related to the depth of penetration since this determines the amount of rotation that occurs before the impact at point *B*. Similar motion of the centre will occur when the inner race moves around point B. The motion around point *A* will be faster than that around point *B* since at point *B* the load acts against the motion. As the shaft is part of the machine structure, the impulsive vibration due to the impact at point *B* will affect the motion of the shaft centre. A certain amount of resonant vibration of the shaft will occur which will be visible as fluctuations in displacement immediately following the effect directly due to the defect described above.

The resultant effect of this on the IME signal can be predicted as a rapid change away from the displacement values due to normal shaft motion, followed by a slower change back to these values with an oscillatory component. This model explains the impulsive motion recorded from the large induced defect described in the previous chapter.

5.5 Experiment to test the assumptions made by the model

An experimental approach was adopted in order to test the assumptions made in the development of the model described in the preceding sections. The initial experiment used was similar to the one used by McFadden and Smith to confirm their model [64].

Initially the IME output signal from each of two new SKF NJ202 roller bearings was recorded. Different speeds were used between 1000 and 3000 rpm and load was changed by the addition of a flywheel for some tests. No significant change in the signal was noted for these conditions. The IME recorded smooth, sinusoidal displacement traces from both bearings, illustrated in figure 71. Displacement, on the y axis, is plotted against angular position for a single rotation from 0 to 2π for both channels. After a profile had been built up of the signal from each bearing, a defect was induced on the inner race of one by electric discharge machining. This defect was 0.2 mm deep, 0.2 mm across and extended across the width of the inner race. This compares to a 0.5 mm wide defect used by McFadden and Smith. Comparison was

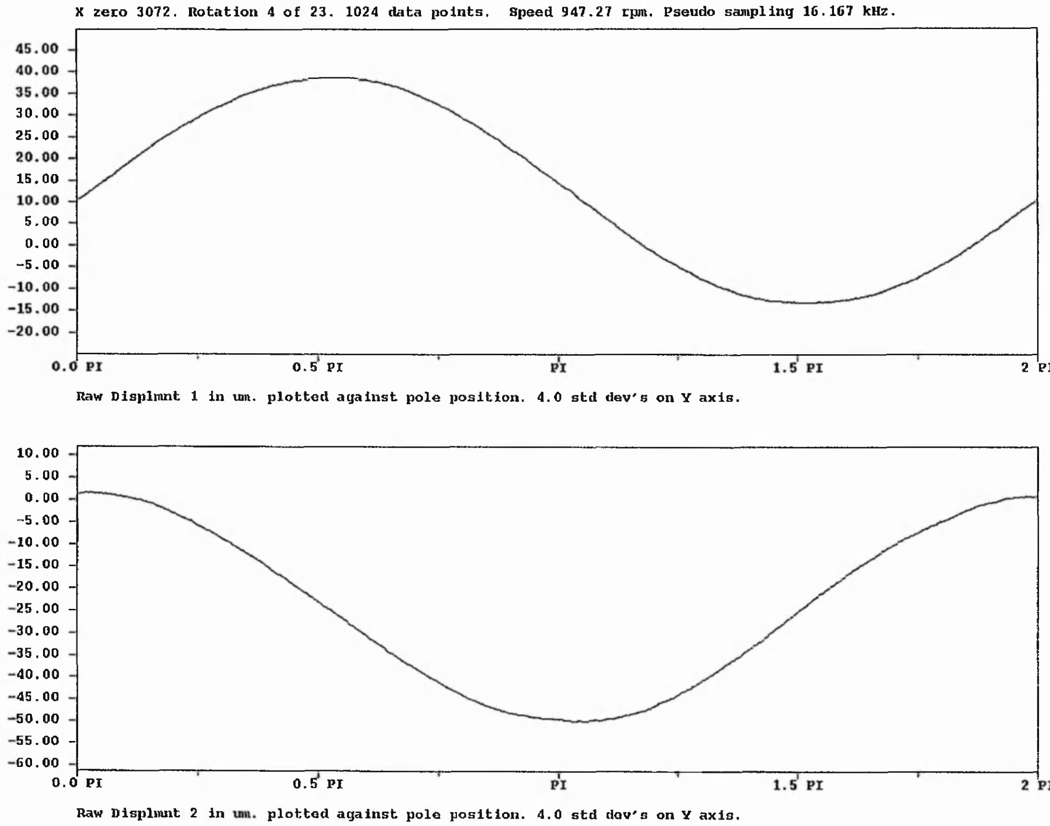


Figure 71. IME displacement signal from a new, unmarked NJ202 roller bearing.

then made between the data collected previously for the marked bearing, the bearing without the defect and the behaviour predicted by the model.

After the defect was induced the change in the signal was quite clear. The smooth sinusoids previously seen were now disrupted by small peaks of displacement. The displacement traces in figure 72 clearly show two small peaks occurring in the second quarter of the rotation on both channels with a third lesser peak on channel two. Also clear is the waveform which appears superimposed on the sinusoid after the two peaks on both channels. These examples are typical of the signal recorded after the defect had been made on the bearing's inner race surface.

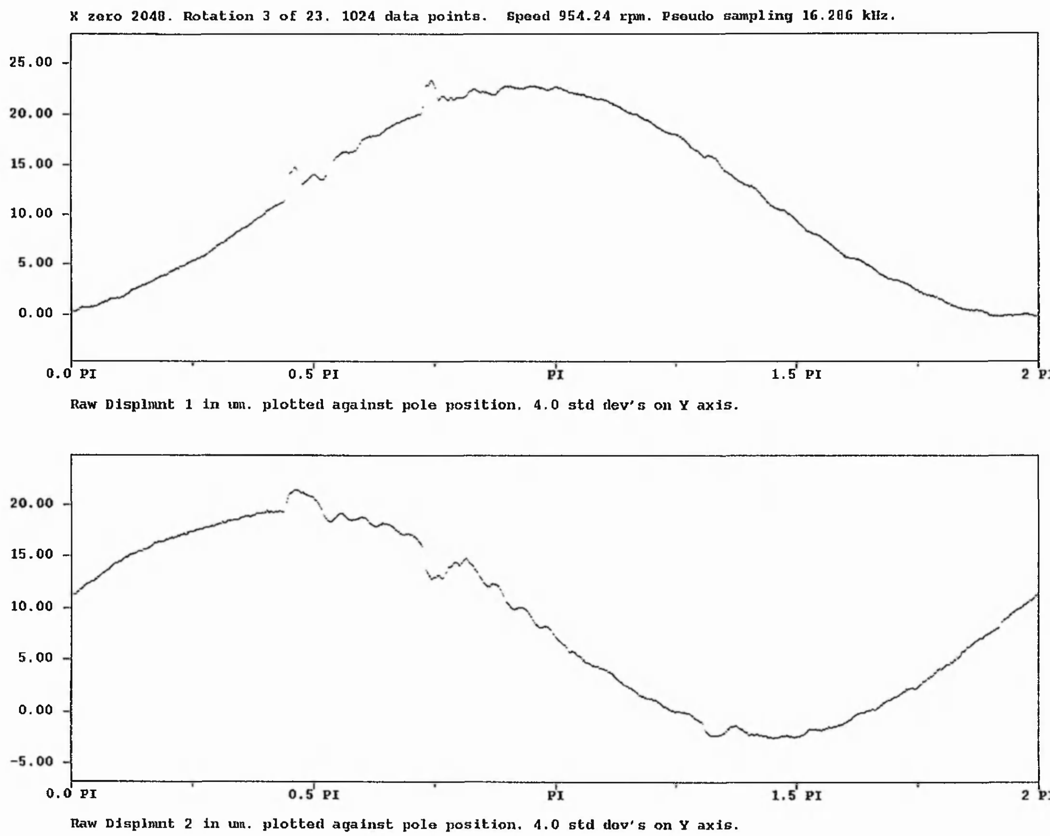


Figure 72. IME displacement signal from bearing with induced defect.

The immediate supposition that these two peaks are due to the defect is confirmed by their angular separation, which is as predicted by the model. The inner race defect frequency for an NJ202 is 6.71 cycles per rotation. Applying equation 13 the expected angular separation of the impulses, θ_d , can be calculated as 0.937 radians equivalent to 149 grating lines. This is the angular distance between the two peaks in figure 72. This spacing was consistently observed. The fact that two peaks are seen indicates that the load zone extends over a space greater than the distance between two elements, approximately 1.872 radians.

The shape of the peaks confirms that the motion of the shaft centre is as predicted by the model. The displacement data points from normal shaft motion are clustered so closely together they form a continuous line. At the point where the first rolling element strikes the defect, point *A* in the model, there is a sudden, sharp jump in the displacement values to a maximum peak. The spacing of the displacement points is greatest at the start of this motion and decreases toward the maximum displacement, showing that this motion is impulsive acceleration. From the peak there is then a slower, but still sharp, return to the previous sinusoidal trend. The motion suggested by the displacement values immediately after each peak is that of an oscillatory resonance due to the impulse generated when the defect and the rolling element meet.

Although the peaks on the two channels have the same spacing they are different in shape and direction of displacement. This difference can be understood with reference to the model. Channel 1 measures the motion of the shaft in the vertical direction. Channel 2 measures motion in the horizontal. The radial load on the shaft, in this example, acts primarily in the vertical direction. The displacement in this direction is therefore against the load and the motion of the shaft centre due to the impulse is of a short sharp nature as the rolling element rolls into and out of the defect. In the horizontal direction the motion is due to the pivoting of the shaft centre as an extension of the line between it and the defect. This motion is less influenced by the load and is of a longer duration. The shaft is more free to move in this axis. In the first peak on channel 2 displacement is initially upwards, indicating motion to the left. In response to the second impulse the initial motion is clearly in the opposite direction, to the right. This is explained by considering the relative position of the two impulses and the load zone. Since the load zone extends around the lower portion of the bearing, for radial load, and there are only two impulses in this example, the assumption can be made that there is one impulse on either side of the position of maximum load. In this case it would be expected that the pivoting motion of the shaft centre in the horizontal axis predicted by the model would be reversed either side of the maximum load position. If the rolling element strikes the defect to the left of the maximum load position then the rotation of the line R - A in figure 70 will be toward the centre of rotation. If the defect is struck to the right of the maximum load position then the motion will be away from the centre of rotation. Since the IME measures motion relative to the centre of rotation this explains the shape of the channel 2 trace in figure 72.

In terms of the displacement measured by the IME the defect was easily detectable in the traces recorded from the marked bearing. The transition of the rolling element from the smooth surface of the inner race to the angular defect and back again gave the results predicted by the model. The fact that this transition is seen to give a rapid change in displacement means that the presence of such a defect will most easily be detected from examination of the shaft centre acceleration. This will be the case regardless of the magnitude of the peak displacement. This measure of shaft motion will therefore be more sensitive to the presence of a defect than to its size.

Differentiation to find the rate of change of the displacement between each angular position on the disk is used to find the shaft centre's velocity. Shaft centre acceleration is calculated by repeating the process with these velocity values. Shaft centre acceleration derived in this way from the IME signal has units of microns per radian squared. For examination of the shaft centre acceleration channel 1 is used as the displacement in the vertical axis is the most rapid as this was the loaded axis during this experimentation.

Figure 73 is an example of the smoothed displacement recorded from channel 1 of the IME for a single rotation. The sinusoidal average of the signal shown in figure 72 has been removed to show in greater detail the impulsive motion due to the defect and the subsequent resonant frequency. The large difference in the relative spacing of the displacement points in the impulse compared to those in the rest of the trace indicate how the relative velocity and acceleration will mark a transition in the signal.

Figure 74 shows the velocity and acceleration traces calculated from the smoothed displacement signal in figure 73. With the data in this format it becomes easier to see the sharp transition that the shaft centre motion makes between normal motion and motion due to the defect. The scale below the velocity trace is marked in three different expressions for the impulse spacing given by equation 13. All the impulses in

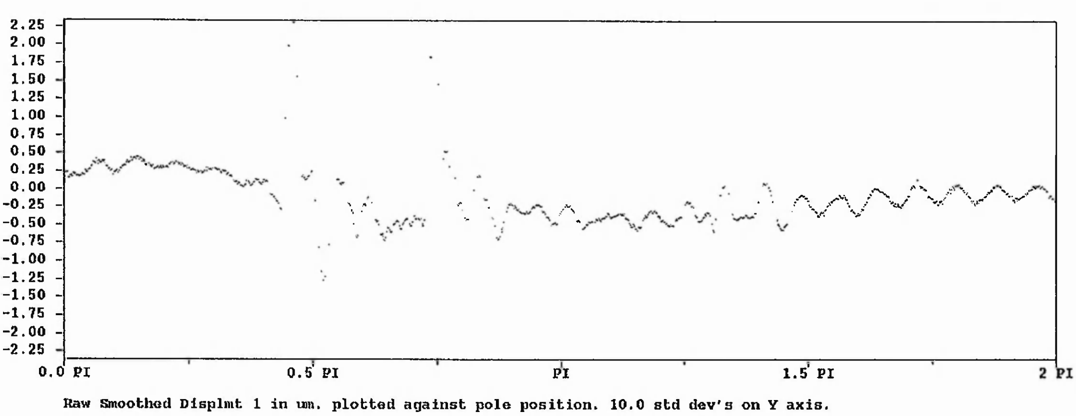


Figure 73. Smoothed displacement signal from bearing with induced defect

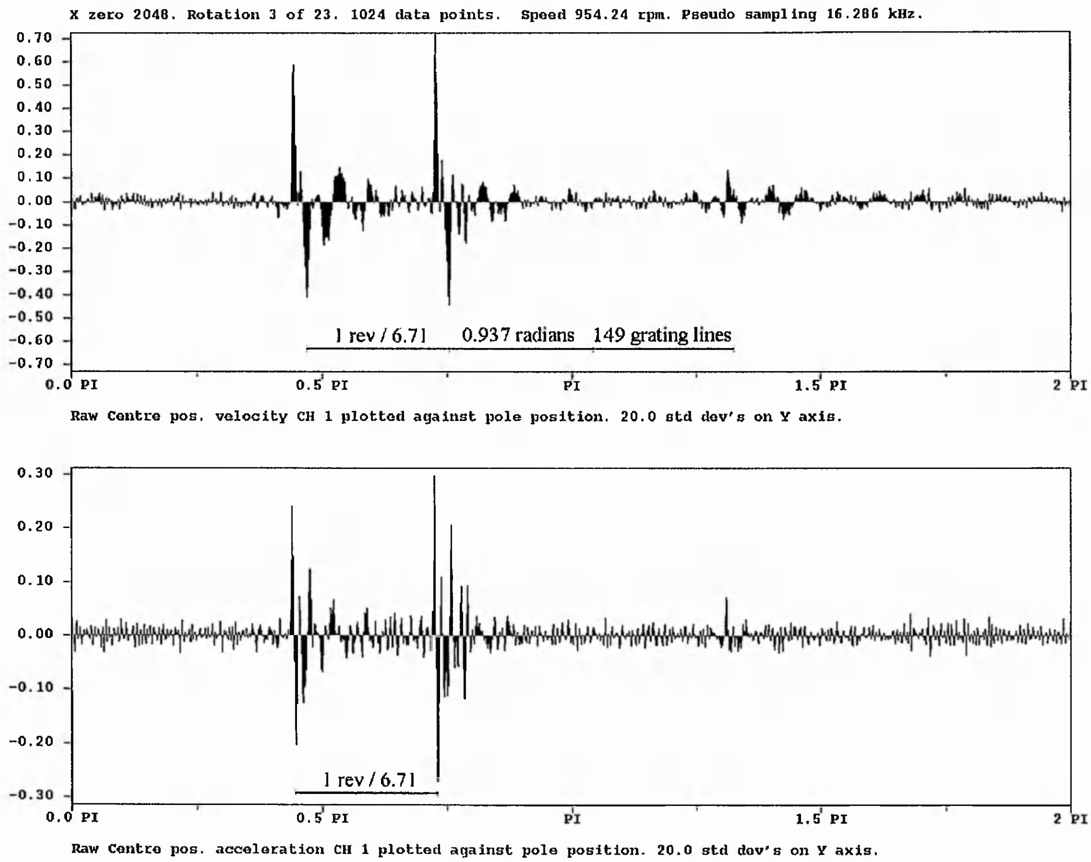


Figure 74. Velocity (top) and acceleration (bottom) signals from bearing with induced defect

this trace and the many others collected as part of this experiment were observed at the same approximate spacing. The marginal differences observed were attributed to the slippage likely to occur between the bearing elements as the bearing was very lightly loaded.

As well as the two main impulse peaks in both traces that define the position of the load zone, a third smaller impulse occurs in the rotation just after 1.25π. This is not related to the load zone and therefore cannot be explained by the model. Observation of the traces showed that impulses of a much smaller amplitude than those in the load zone were seen to occur in a position diametrically opposite to the load zone. From this experimental observation it is was deduced that IME was picking up the response of the shaft centre motion to the rolling elements passing over the defect when it was in the top of the bearing.

Over successive rotations the position and amplitude of the impulse due to the defect were observed to change in the manner predicted by the model. Figure 75 shows the shaft centre acceleration over four successive rotations. A dotted line at the 0.5π position in each trace provides a reference point against which the relative motion of the impulses can be judged. A scale bar of constant length below each trace indicates the calculated defect spacing.

The adaptation of the McFadden and Smith model to the IME predicts that the amplitude of an impulse due to a defect will change due to the variation in load with angular position, $d(\theta)q(\theta)$, as shown in figure 69. The pattern of amplitude variation seen in experimental results and illustrated in figures 75 and 76 confirms this. The maximum amplitude impulse is seen in the first trace in figure 75 where it occurs approximately 0.625π radians into the rotation. A second much smaller impulse occurs to the left of the main one separated by the calculated defect spacing, $d(\theta)$. In the second trace the position of the main impulse has moved to the left by approximately 0.03125π and its amplitude has decreased. This apparent motion is due to the fact that the load zone is fixed in relation to the IME measurement, but on each rotation the rolling elements will be in different positions relative to the defect on the inner race. The smaller impulse has moved by the same angular amount and has also decreased in amplitude.

To the right of the main impulse in the second trace a third impulse has appeared. The two smaller impulses in the trace are separated to the left and right of the main one by $d(\theta)$. Both of these smaller impulses mark the angular extremities of the Stribeck load zone. Between these positions the amplitude of an impulse will increase to the maximum value seen in the first trace and then decrease again as the angular position changes. Beyond these positions to the left and right no impulses due to contact in the load zone will occur.

In the third and fourth traces in figure 75 as the impulses move to the left in each successive rotation their amplitude changes with their relative position. The main impulse decreases in amplitude as it moves further from the centre of the load zone. The smaller impulse increases as it moves in from the rightmost extreme of the load zone left toward the centre position defined by the maximum amplitude seen in the first trace. In the rotation trace shown in figure 76 the two impulses are approaching equal amplitude.

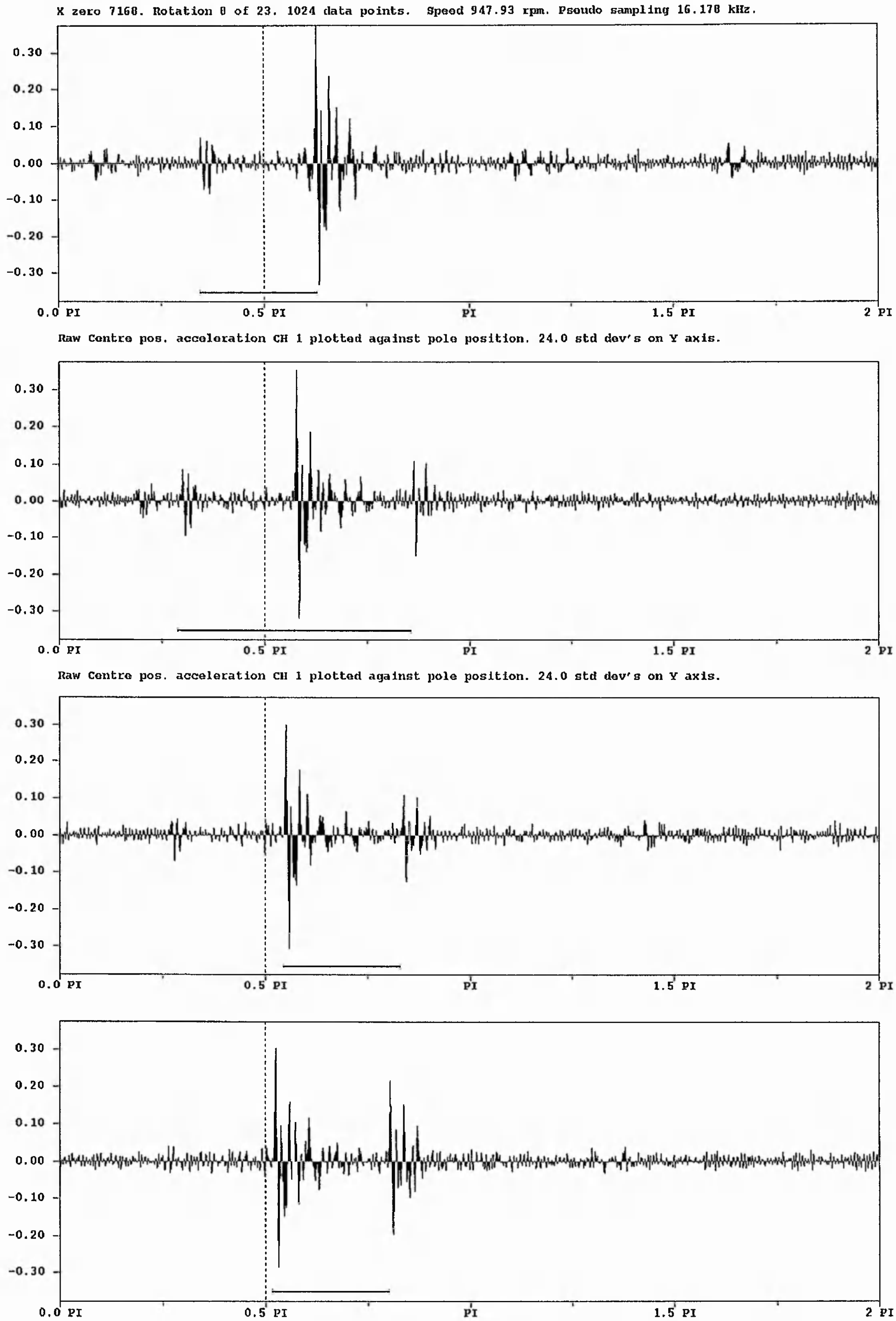


Figure 75. The change in amplitude with relative position of the shaft centre acceleration impulse resulting from the bearing defect over four successive rotations. Scale bar indicates the interval calculated from the bearing defect frequency. Dotted line at 0.5π for reference.

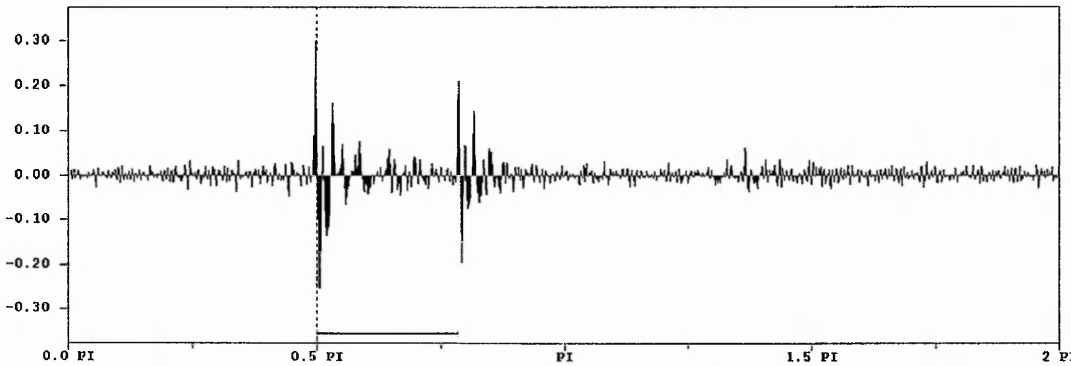


Figure 76. After five rotations the most prominent impulse has moved to the 0.5π position and decreased in amplitude. While the spacing between the impulses remains the same, the amplitude of the second impulse increases as it moves toward the right.

5.6 Extending the IME model to allow for bearing defect detection

In order to be effective as a condition monitoring technique the IME must allow the detection of incipient faults, as well as the large, fairly obvious one presented in the preceding example. The change from a random acceleration signal to an impulsive one is the key to this.

5.6.1 The kurtosis coefficient

In a properly lubricated, defect free, rolling element bearing the motion recorded by an IME sensor will be related to the normal contacts between the bearing's races and its elements. The shaft centre acceleration in this case will result from the random roughness of the normal bearing surfaces. This random roughness will produce a signal in which there is a random amplitude distribution [83]. Figure 77 shows the random amplitude shaft centre acceleration signal recorded by the IME from an undamaged NJ202 bearing.

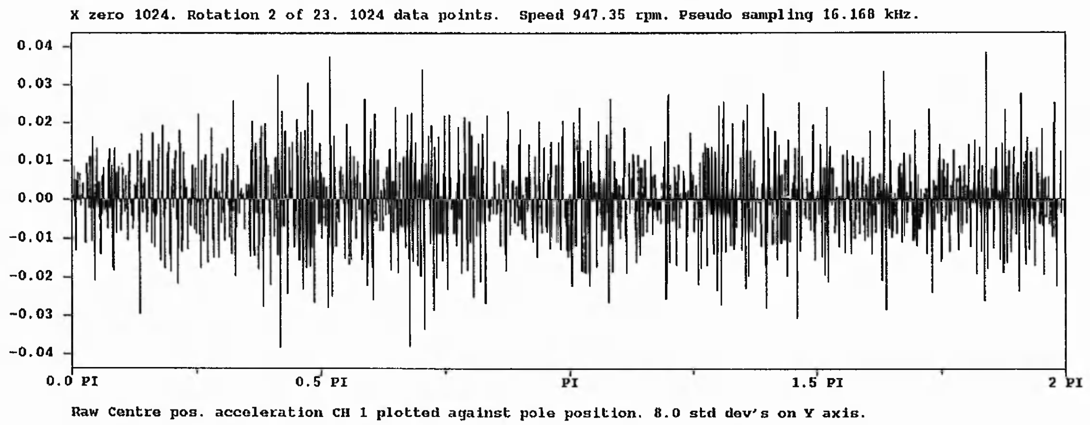


Figure 77. Shaft centre position acceleration recorded by the IME from the undamaged NJ202 bearing showing random amplitude distribution. Kurtosis coefficient 3.37.

The random amplitude distribution of this acceleration signal produces the approximately gaussian curve shown in figure 78. This distribution curve will change with the development of a spall-type defect on one of the bearing surfaces. The amplitude of any regularly spaced impulses in the acceleration signal will cause the tails of the amplitude distribution curve to increase. To detect this change the acceleration signal must be analysed statistically using a parameter which is sensitive to this change in the amplitude distribution.

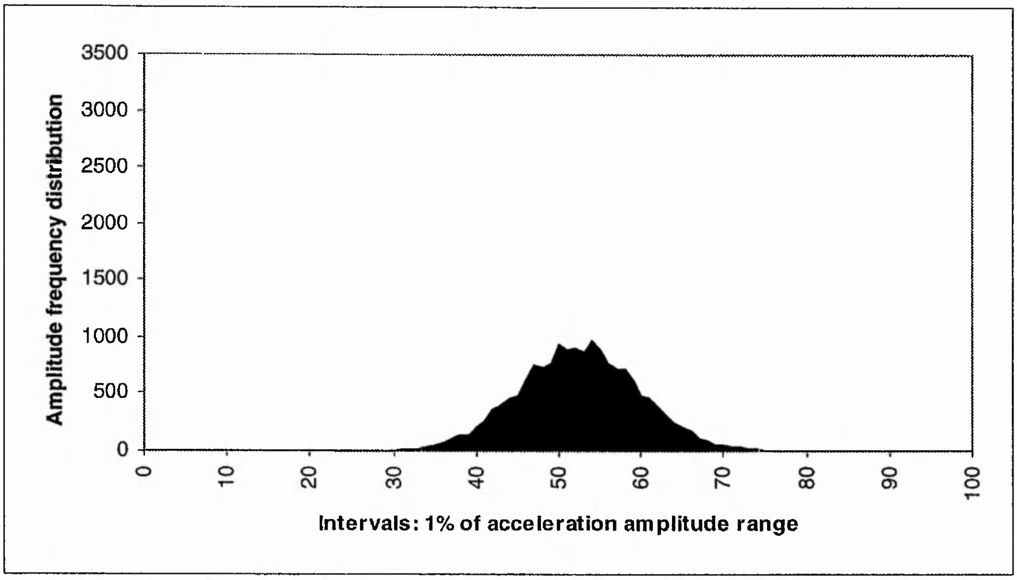


Figure 78. The amplitude distribution curve from the shaft centre acceleration signal in figure 77. The random acceleration produces an approximately gaussian curve.

Dyer and Stewart, who used accelerometers for bearing condition monitoring, proposed the statistical parameter kurtosis for the detection of the change from a random to an impulsive signal [62]. Kurtosis is a dimensionless, statistical parameter which remains constant for an undamaged bearing irrespective of load and speed, yet changes with damage. When applied to a signal it gives a measure of its amplitude distribution. The kurtosis coefficient is calculated as the fourth statistical moment, M_4, normalised to the root mean square value [4], equation 15.

$$kurtosis\ coefficient = \frac{M_4}{(RMS)^4} \tag{15}$$

The kurtosis coefficient of a purely random noise signal has a value of 3.0. As a signal becomes more impulsive its kurtosis coefficient increases above 3.0. The kurtosis coefficient is sensitive to the size and distribution of the impulse peaks in a signal. The average value of the kurtosis coefficient calculated from the shaft centre acceleration signals of the two undamaged bearings was 3.011.

In figure 79 the shaft centre acceleration from one rotation of the bearing with the induced defect is shown. The amplitude range of the bulk of the signal is still the same as that from a normal bearing. This is because the majority of the motion in the bearing is still between normal bearing surfaces. The signal is however dominated by the impulse peaks generated by the defect. The small number of these large amplitudes in the signal produces a distribution curve like that shown in figure 80. The impulse peaks have extended the amplitude range of the signal so that the distribution tails out

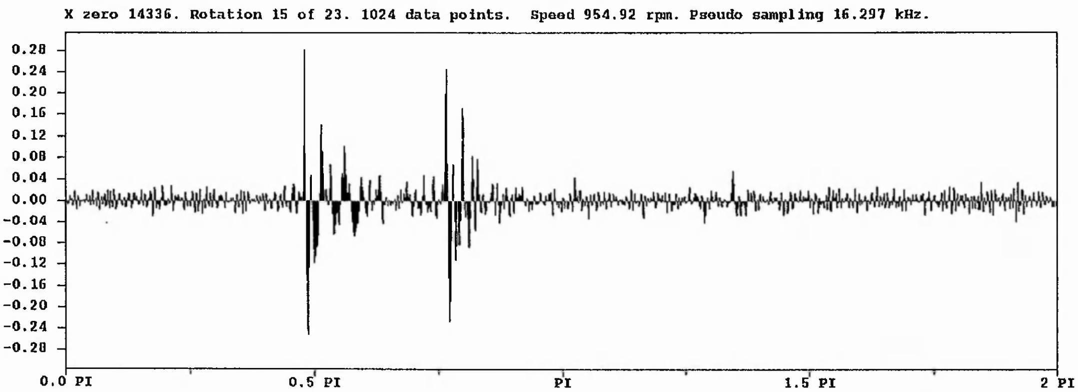


Figure 79. Impulsive shaft centre acceleration from NJ202 bearing with induced defect. Kurtosis coefficient 39.39.

from the central peak. The average value of the kurtosis coefficient recorded from the bearing with the induced defect was 32.451.

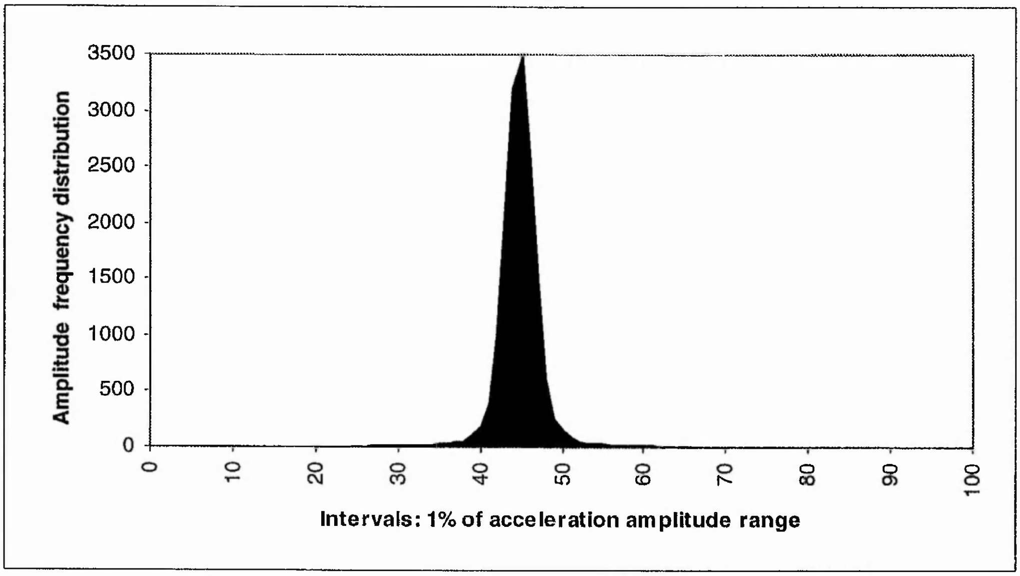


Figure 80. The amplitude distribution of the defective bearing's shaft centre acceleration signal in figure 79. The impulsive components of the signal have given rise to the tails of the distribution curve. As these develop the value of the kurtosis coefficient increases.

From these results it is clear that a calculation of the kurtosis coefficient from the IME's shaft centre acceleration signal can determine the difference between the signal from the marked and the unmarked bearing. Further work is required to prove the ability of kurtosis to accurately determine the onset of bearing damage from the IME signal. This would require an endurance rig capable of conducting an accelerated bearing life test.

5.6.2 Measuring signal similarity

As well as the use of kurtosis the frequency distribution of the IME's shaft centre acceleration signal could also be analysed by the use of a similarity measure. Deciding whether two statistical distributions differ or are consistent is a recurring problem throughout many branches of the sciences. In the case of the IME the amplitude frequency distribution of a newly installed bearing could be calculated and then compared to subsequent values to detect the development of the impulsive shaft centre acceleration due to a defect. Currently the chi-squared statistic is the most commonly used technique for comparing statistical distributions [84]. To compare two distributions designated R and S the chi-squared statistic is calculated using equation 16 where R_i is contents of the *i*th bin of distribution R and S_i is the contents of the corresponding bin in the second data set.

$$\chi^2 = \sum_i \frac{(R_i - S_i)^2}{R_i + S_i} \tag{16}$$

Thacker et al. have suggested that the Bhattacharyya metric may offer advantages over the chi-squared statistic in the field of system identification from frequency distribution data where large differences exist between the distributions [85].

The angular nature of the IME's data means that the frequency that any bearing defect will produce can be determined by calculation. It will not be dependent on, or variable with, the speed of shaft rotation as is the case with accelerometer signals. The effect of the load zone will be a constant factor determining the impulse amplitude. On this basis the likely amplitude distribution in an IME signal from a particular bearing with a defect can be simulated. By calculating similarity measures between this simulated signal, the actual signal from a newly installed bearing and the signal from same bearing subsequently, the direction of change in the signal's amplitude distribution could be monitored. This would allow the trending of the acceleration signal and the ability to determine where the trend was heading based on the initial divergence from the normal random amplitude distribution. Figure 81 shows how the IME model is extended by the use of a similarity measure.

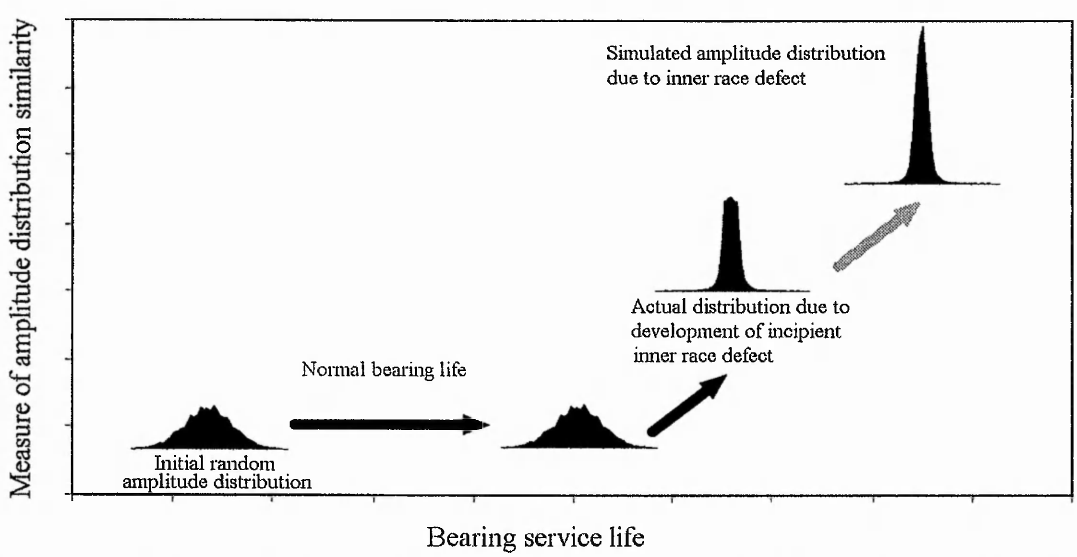


Figure 81. The use of a similarity measure over the service life of a bearing to determine the presence and development trend of the an incipient defect from the amplitude distribution of the IME shaft centre acceleration signal.

5.6.3 Discrimination of the source of an impulsive signal

Monitoring of the kurtosis level of the shaft centre acceleration allows the presence of impulses in the signal to be detected, but the source from which they originate cannot be determined. Discriminating between defects on the rollers and the inner or outer races is not generally useful since in the event of damage on one part the whole bearing will be replaced. However it is important to be able to differentiate between impulsive signals that originate in the bearing supporting the shaft and those which might arise from other sources in a machine, such as gear meshing. The nature of the IME signal, collected at equal angular intervals, allows the direct measurement of the angular distance between impulses in the signal. This is an advantage over conventional vibration sensors which need a separate angular motion sensor to relate their signals to the rotation of the shaft. Knowledge of the geometry of a bearing's races and rolling elements allows the frequency of the impulsive signal due to any spall-type defect to be calculated. From this frequency the angular interval to be expected in the IME signal between each impulse from such a defect, expressed as a number of grating lines, can be calculated.

Once impulses in the shaft centre acceleration signal are detected from a rise in the kurtosis level, the angular interval between the impulses can be compared to the values calculated for bearing defects. On an amplitude distribution curve the impulsive components of a signal, which produce a high kurtosis coefficient, will lie above the level of twice the standard deviation. This is the threshold level used by the experimental IME system above which the amplitude of the signal counts as an impulse [86]. For each impulse in twenty rotations of data the system measures the angular distance to impulses in the three succeeding and preceding rotations of data. By dividing the distances measured between impulses in the signal by the intervals calculated for bearing defects the system can determine whether or not the source of the impulses is part of the bearing. At the same time, by dividing the length of a rotation by the interval measured between impulses, the system is able to determine if the interval is an integer fraction of a rotation. The impulses due to bearing defects are always spaced at intervals that are non-integer fractions of a rotation as a result of bearing geometry. Shaft vibration sources such as gear meshing or fan blade passing always result in an integer number of impulses per rotation corresponding to the number of gear teeth or fan blades present. This difference provides the method by which the two types of vibration source can differentiated.

5.7 Conclusions from the experimental testing of the IME model

At the end of the second phase of experimentation all of the questions about the IME and the nature of its signal raised by the first set of experiments had been successfully answered. A model had been developed and confirmed experimentally to explain the response of the IME signal to a bearing defect. This model had been extended to allow the detection of the development of bearing defects in their initial stages by statistical analysis of the IME signal. A method of discriminating between the impulses due to bearing defects and those from other sources was also added to the basic model and tested experimentally. This model and its application are two of the principal claims to novelty that are made in this thesis.

A high level of confidence had been achieved in the IME system at the end of the second set of IME experiments. It had clearly been proven in experimentation that the IME could monitor a bearing accurately. In three further experiments which were conducted it was shown that the experimental IME system was able to successfully detect lubricant contamination, shaft loading and bearing corrosion. Details of these experiments are contained in appendices G, H and I.

Chapter6

Conclusions and future work

6. Conclusions and future work

6.1 Conclusions

The overall objective set at the start of this project was the investigation of a new type of shaft displacement transducer to be based on the patented, but unproven, IME principle. This aim was successfully achieved by the construction of an experimental bearing condition monitoring system based on a prototype IME sensor and its application in a series of experiments to monitor the condition of rolling element bearings with induced defects or simulated fault conditions. The project was conducted in four phases.

In the initial phase of the project the objective was the implementation of the IME system hardware. The sensor itself was successfully constructed based on the patent application. However the electronic circuitry and the computer hardware which were described proved to be inadequate. It was found that the data rate from the fully implemented four read head sensor could not be captured by the system proposed in the patent. This finding resulted in a completely new analysis of the interface requirement for an IME sensor. A hardware independent conceptual architecture was designed and implemented to handle the required data collection rate. Together with the software that was written to run on it, this implementation formed the basis of the experimental system with which the rest of the project was carried out.

With an experimental IME system successfully constructed the second phase of the project began with an attempt to interpret the signal the sensor produced. The initial approach adopted was to look for the geometrically derived defect frequencies in the signal from a bearing with an induced defect. This was compared with the signal from an unmarked bearing. The displacement signals and their frequency spectra were examined. Various filters and averaging techniques were used, but none were able to reveal a distinctive feature in the frequency spectrum which could be related to the calculated defect frequencies. Those results which did show a partial agreement were not consistently repeatable. During this period a conventional accelerometer and spectrum analyser were also used to monitor the bearings in an attempt to confirm the frequencies actually being generated. This experiment was inconclusive. No frequency that could be directly attributed to the bearing defect was identified. At the end of this series of experiments it was concluded that the IME signal could not be interpreted from its frequency spectra using the conventional approach applied to accelerometer signals.

To provide an interpretation framework for the IME signal a model was required which would explain the effect of a bearing defect on the signal. The third phase of the project began with the development of this model. A combination of published work by various authors and the knowledge of the IME signal that had been gained in the initial experimentation was used to construct the new model. A mechanism for the response of the IME signal to the mechanical motion of the supported shaft resulting from the contact between an inner race defect and a rolling element was developed. From this the model was able to predict that the defect would result in an impulsive response in the acceleration of the shaft centre. The modulation of these impulses would be dependent on the load distribution in the bearing. The fact that the motion is impulsive led to the suggestion that the onset of bearing damage could be detected from the kurtosis coefficient of the signal combined with other statistical analysis techniques. The assumptions made in this model were confirmed by experimentation which showed that the IME could clearly detect a large, but still realistically sized, defect on a bearing inner race. A method of discriminating between the impulses due to the defect and those due to other machine components was proposed based on the angular separation of the impulses and their relative position between successive rotations. This potentially gives the IME the important attribute of being able to monitor individual bearings in a complex machine with many sources of impulsive vibration.

In the fourth and final phase of the project a set of experiments were successfully conducted showing the ability of the IME system to detect three types of pre-failure bearing condition. Lubricant contamination, race corrosion and shaft loading, all of which can lead to bearing failure, were clearly shown to produce changes in the IME signal predictable from the model which had been developed [86,87,88].

Successful application of the IME principle to the condition monitoring of rolling element bearings will represent the first completely new technique in this area for a decade. This project has proven that the novel, and previously untested, IME concept is itself valid. An experimental sensor of this type has been constructed and shown to give some of the many types of information about a rotary system predicted by the IME's theory. The principal claim to novelty made by this project is in the manner of interpreting the IME signal to yield an accurate indication of the condition of a rolling element bearing.

It is a conclusion of this project that the recent development of bearings with integrated encoder elements provides the possibility of extending the IME sensor from the experimental laboratory desktop to industrial applications. By using the integrated encoder elements an IME sensor could be incorporated into a standard bearing. This would allow the creation of a new class of sensorised bearings, developed with current technology and using the IME principle. These would be able to sense a number of machine parameters as well as their own condition. Within an intelligent machine tool, a smart bearing of this type would allow a higher degree of control and monitoring information integration than is currently possible for rotating machines.

6.2 Future work

There are two main areas into which future work on this project could fall. There is still scope for further development of the mathematical theory behind the IME. Work in this area would allow the accuracy of the sensor to be confirmed.

Although the IME has been shown to work experimentally, the main area of future work will be in the translation of its potential into an industrial reality. The growing importance of computer integrated manufacturing (CIM) is increasing the need for better machine monitoring and fault detection. High reliability, achieved through effective maintenance, is an essential prerequisite for the successful operation of the sophisticated machine tools in a CIM system. Rolling element bearing monitoring is important in this context as machine tool bearings are critical components with the potential for unexpected failure.

Consideration must be given to the factors involved in scaling the IME sensor to suit a real application. The most obvious solution would be the use of an optical disk similar to the one used in this project's experimentation. Figure 82 shows how an optical encoder could

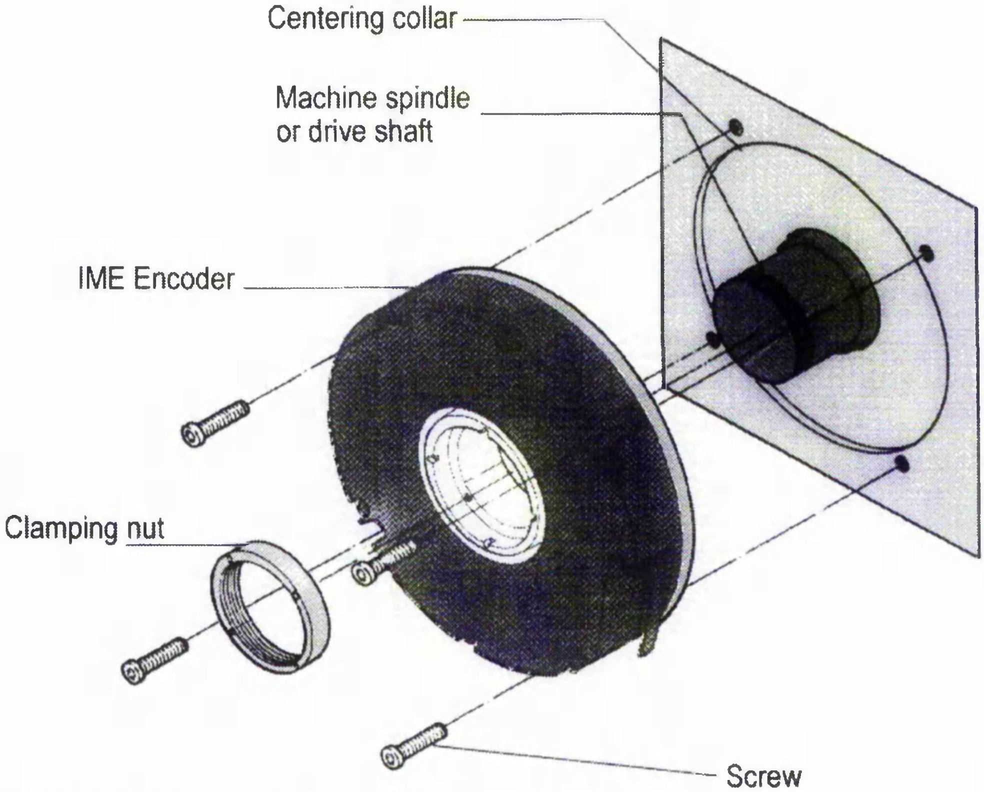


Figure 82. An illustration of the possible mounting of an IME device based on an optical encoder of the Heidenhain RON type. The sensor in this case is added to the existing machine structure.

be mounted onto an existing machine's drive or spindle shaft. Although suitable for some applications, adding the sensor to the machine in this manner is not concluded to be the ideal solution. The rotary motion sensors already built into machine tools may, with the addition of an extra sensing element, be able to implement the functionality of an IME sensor. This would allow the sensor to be integrated into the existing machine control structure. In this way the machine's motion control computer could monitor the condition of its bearings, as well as angular motion with the much improved accuracy given by time base interpolation. Using sensors already designed into machine tools or other industrial equipment would also satisfy the desire to use robust technologies more suited to potentially challenging industrial environments than the current experimental sensor.

Figure 83 illustrates a typical rotary motion sensor of the type that might be found in a variety of industrial equipment. The rotation of a toothed metal wheel past an inductive coil generates a signal indicating speed of rotation. During this project interest was shown in the IME principle by a manufacturer of aero engines. The sensor in a aero engine, called a phonic wheel, is of this type. Work must be conducted to determine what factors will need to be considered in building an IME sensor based on a device like this. The effect of the spacing of the inductive coil relative to the teeth of the wheel, the spacing of the teeth on the wheel and the sensitivity of the signal to slight changes in angular position must all be quantified. The model which has been developed for the experimental sensor will then need to be tested and extended if necessary.

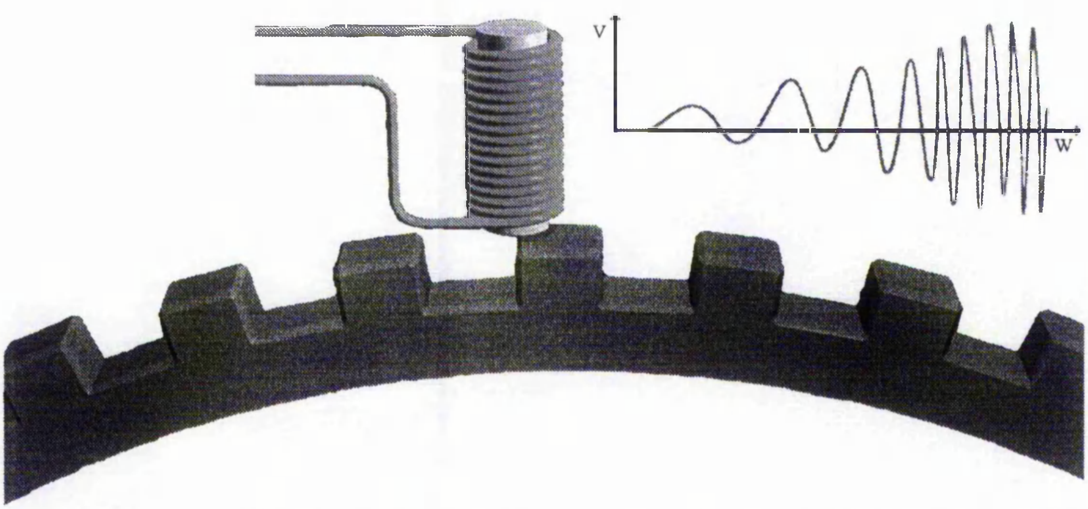


Figure 83. An example of a rotary motion sensor from an industrial machine. The change in the EMF of a coil with the passage of the teeth of a ferrous wheel rotating with the shaft indicates rotational speed.

The development which has the greatest potential for bringing the IME principle into the industrial world is the instrumented bearing. Until recently rolling element bearings were monitored, but had no sensory ability of their own. Their rotational speed would be measured for a machine control system by a shaft encoder and their condition monitored by a separate diagnostic system. The introduction of Anti-lock Braking Systems (ABS) by the automotive industry created the first application for bearings with integrated sensors. Car road wheel speed could be most accurately measured by incorporating an encoder-like electromagnetic sensor into the wheel bearing. Several technologies have been developed to be used by ABS and the sensorised bearing has proven very successful in this area [89]. In the past five years some bearing manufacturers have begun to develop and market sensorised bearings for general applications, particularly in machine tools. Figure 84 shows one example of such a device. This sensorised bearing uses a single Hall effect sensor in the seal attached to the bearing's outer ring and a magnetic multipolar ring fixed to the rotating inner ring to implement an encoder. To achieve the basic functionality of an IME device only the addition of a second sensor would be required. The relative simplicity of the calculations required to process the data from the IME would allow existing micro computer based machine controllers to interface directly to a smart bearing with the addition of minimal hardware and software.

The emphasis of the commercial development of sensorised bearing has primarily been on rotary speed sensing although bearings with internal strain gauges to measure shaft loading have been developed [90]. Shaft loading can be used to monitor tool wear in machine tools and roller tension in a variety of industrial processes.

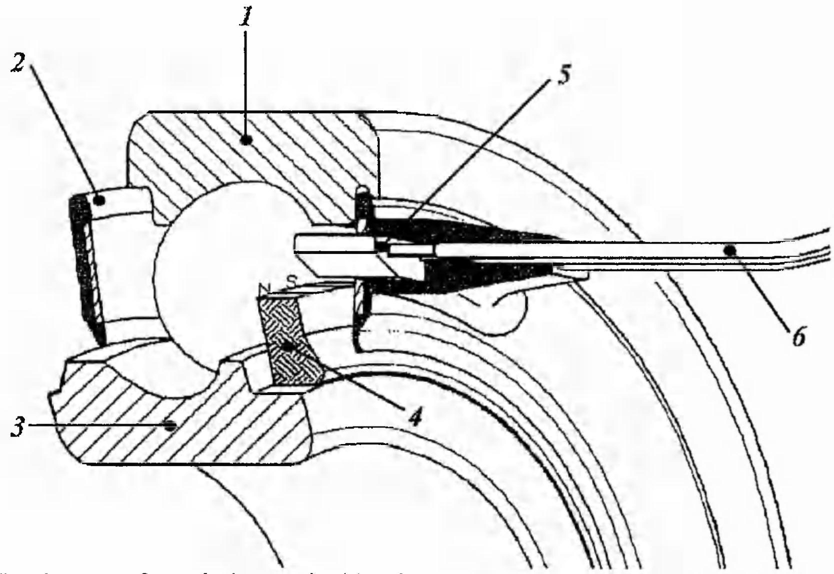


Figure 84. The elements of a typical sensorised bearing manufactured for machine tool speed monitoring by SNR. 1. Fixed outer ring, 2. Seal, 3. Rotating inner ring, 4. Magnetic multipolar ring, 5. Hall effect sensor, 6. Cable.

Based on the work done during this project developing the incremental motion encoder, it is believed that there is a great potential for a new generation of bearings incorporating the IME principle. Such bearings could monitor machine parameters such as angular velocity, angular vibration, torque and shaft loading together with their own condition. A number of conventional sensors and their respective data analysis systems could be replaced with a single unit, an intelligent or 'smart' bearing. This could be integrated into existing machine designs with only minimal modifications for cabling. A bearing of this type could be used for process control and monitoring during normal machine operation as well as for detecting fault conditions. As a system component, a smart bearing would be a significant step forward in the direction of the type of intelligent machine required for use in CIM systems [5].

Figure 85 illustrates a typical industrial machine gearbox. The bearings in this type of environment experience forces due to the meshing of the gears, the equipment being driven as well as their own condition. The site of the bearing is therefore an excellent location for a sensor. The possibility of replacing simple 'dumb' bearings with 'smart' bearings using the IME principle similar to the conceptual example shown, should be investigated. This offers the prospect of industrial collaboration in the future direction of this research as the production of this type of bearing could only be undertaken by a bearing manufacturer. During this project the experimental work undertaken was demonstrated to a bearing company who have expressed interest in its further development.

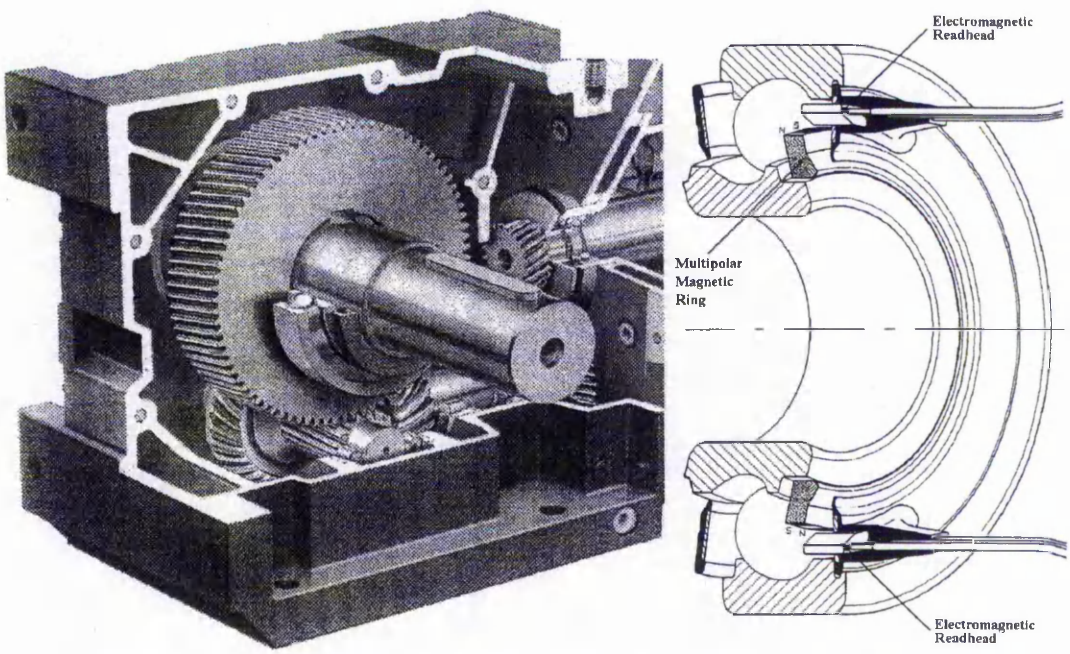


Figure 85. An industrial gearbox (left) and a conceptual illustration of a smart bearing (right) which might be developed in the future to allow such a machine to monitor its own condition.

A possible intermediate step between the current experimental rig and an industrially manufactured smart bearing would be the type of modularised wheel hub sensor bearing produced by SNR for the automotive market. These bearings have a multipolar magnetic ring on their seal which rotates with the inner race. A Hall effect sensor is clipped to the outer race of the bearing on detachable ring. Although the normal practise is to use only one sensor on each bearing, there is no reason why more should not be added to the clip-on ring to provide the sensor configuration of an IME as shown in figure 86.

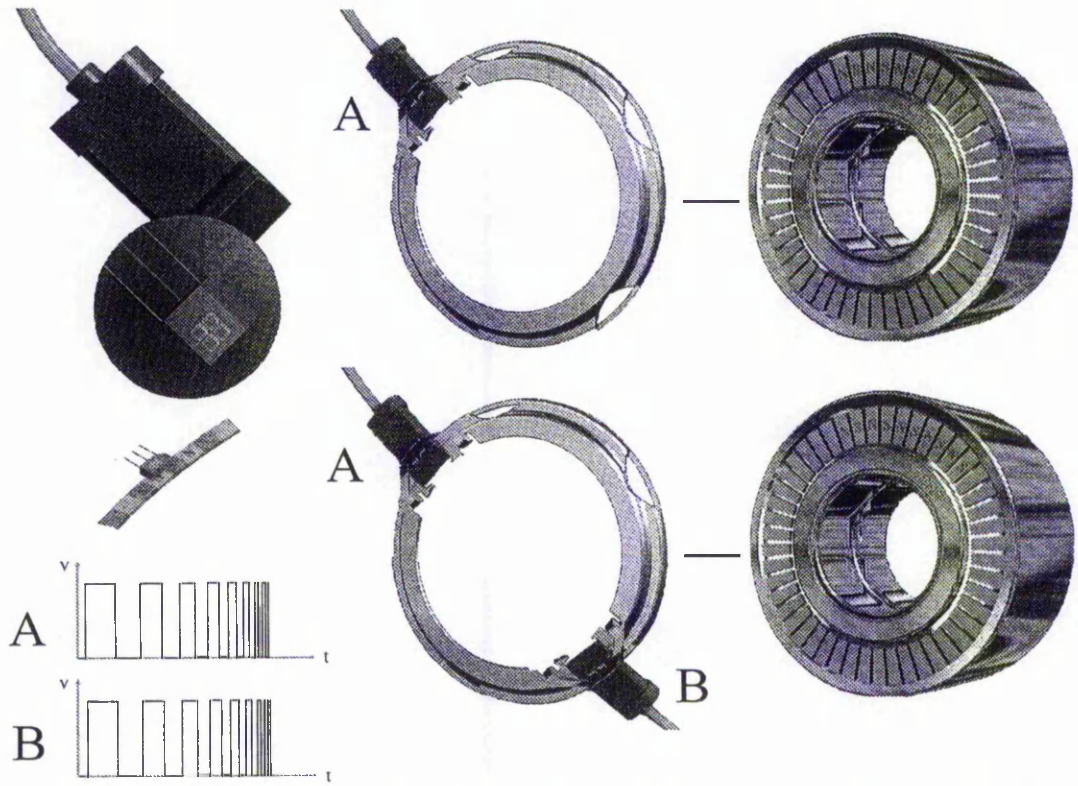


Figure 86. An automotive wheel hub bearing for ABS applications with integral encoder element and detachable Hall effect sensor. An experimental smart bearing could be constructed from this type of device.

With suitable supporting research effort the IME principle and the smart bearing concept could become an industrial reality. The successful exploitation of the results of this project remains as a challenge yet to be met.

References

References

1. ORTON, P. A. - Instrumentation and Control for precision Grinding machines, PhD Thesis, Nottingham Polytechnic, 1990.

2. ORTON, P. A. - Shaft Displacement Measuring System, International patent application number PCT/GB92/00811, World Intellectual Property Organization, International Publication number WO 92/19936 12 November 1992.

3. AYANDOKUN, K. - The realisation of a Shaft Displacement Measuring System for Real-Time Condition Monitoring of Rotating Machinery, MSc Thesis, Nottingham Polytechnic, September 1992.

4. COLLACOTT, R. A. - Vibration Monitoring and Diagnosis, George Godwin Limited, London 1979.

5. DAVIES, A. - The intelligent machine, Manufacturing engineer, p 182 - 185, Aug. 1994.

6. TRANTER, J. T. - The fundamentals of, and the application of computers to, condition monitoring and predictive maintenance, The Institution of Engineers Australia Vibration and Noise Conference, Melbourne, 18 - 20 September 1990.

7. PAZ, N. M. and LEIGH, W. - Maintenance scheduling: Issues, results and research needs, International journal of operational and production management, v 14 n 8 p 47 - 69 1994.

8. SCHEITHE, W. - Better bearing vibration analysis, Hydrocarbon Processing, v 71 n 7 p 57 - 64 1992.

9. TALLIAN, T. E. - Simplified contact fatigue life prediction model-Part I: Review of published models, ASME Journal of Tribology, v 114 p 207 - 213 1992.

10. TALLIAN, T. E. - Simplified contact fatigue life prediction model-Part II: New models, ASME Journal of Tribology, v 114 p 214 - 222 1992.

11. Rolling bearings - Dynamic load ratings and rating life. Draft international standard ISO/DIS 281, ISO, Geneva, Switzerland, 1989.

12. CULVERSON, K. J. - Maintenance analysis the key to successful reliability based maintenance, Proceedings of the 8th international Congress on condition monitoring and diagnostic engineering management (COMADEM'95), Queen's University, Kingston, Ontario, Canada, 1995, v 1 p 181 - 188.

13. NICKERSON, G. W. and HALL, D. L. - Research imperatives for condition-based maintenance, Proceedings of the 8th international Congress on condition monitoring and diagnostic engineering management (COMADEM'95), Queen's University, Kingston, Ontario, Canada, 1995, v 1 p 203 - 210.

14. HANDELIN, G. W. - General Considerations for Vibration Monitoring, Turbomachinery International v 32 n 5 p 28 - 32 1991.

15. HYDE, T. R., MELTON, S. J., HAVELOCK, G. and NICHOLL, R. - Condition monitoring in the UK paper and food industries, Proceedings of the 8th international Congress on condition monitoring and diagnostic engineering management (COMADEM'95), Queen's University, Kingston, Ontario, Canada, 1995, v 1 p 309 - 314.

16. MARTIN, K. F., HOH, S. M. and WILLIAMS, J. H., - Condition monitoring machine tool drives via health indices, IFAC Symp. SAFEPROCESS '91, p 311 - 316, Baden Baden, Sept. 1991.

17. MARTIN, K. F. - A review by discussion of condition monitoring and fault diagnosis in machine tools, Int. J. Mach. Manufact., vol. 34, n. 4, p 527 - 551, 1994.

18. RAMIREZ, C. N. and THORNHILL, R. J., - On estimating rotational speed and torque from the period of an encoder", *ASME Sensors, controls and quality issues in manufacturing*, 1991.

19. KUHNELL, B. T. and STECKI, J. S., - Correlation of vibration, wear debris analysis and oil analysis in rolling element bearing condition monitoring, Condition Monitoring '84 Pineridge Press Ltd., Swansea, United Kingdom, 1984.

20. HUNT, T. M., - What debris monitor for which application, Condition monitoring '94, Pineridge Press Ltd., Swansea, United Kingdom 1994.

21. MUIR, D. E. - In-line oil debris monitoring of rolling element bearings, Proceedings of the 8th international Congress on condition monitoring and diagnostic engineering management (COMADEM'95), Queen's University, Kingston, Ontario, Canada, 1995, v 2 p 767 - 772.

22. ZHONG, L., KEXING, C. and DARONG, C. - An intelligent wear monitoring & diagnosis system based on integrated wear particles analysis, Proceedings of the 8th international Congress on condition monitoring and diagnostic engineering management (COMADEM'95), Queen's University, Kingston, Ontario, Canada, 1995, v 2 p 729 - 733.

23. WANG, T. *et al.* - Comprehensive condition monitoring of gear operation, Journal of Aerospace Power/Hangkong Dongli Xuebao, v 7 n 1 p 27 - 30 Jan 1992.

24. BONGA, M. D. and ROBICHAUD, J. M. - The integration of condition monitoring technologies, Proceedings of the 8th international Congress on condition monitoring and diagnostic engineering management (COMADEM'95), Queen's University, Kingston, Ontario, Canada, 1995, v1 p 323 - 327.

25. MILLER, G. P. and MCCLYMONDS, S. L. - Maintenance cost avoidance through comprehensive condition monitoring, Jnt. ASME/IEEE Power Generation Conference, ASME, Boston, MA, October 21 - 25, 1990.

26. HARRIS, T., MACINTYRE, J. and SMITH, P. - Neural networks and their application to vibration analysis, Proc. of the ASME energy-sources tech conf., v 63, p 65 - 67, 1994.

27 KIRKHAM, C. and HARRIS, T. J. - Development of a hybrid neural network/expert system for machine health monitoring, Proceedings of the 8th international Congress on condition monitoring and diagnostic engineering management (COMADEM'95), Queen's University, Kingston, Ontario, Canada, 1995, v 1 p 55 - 60.

28. LIU, T. I. and MENGEL, J. M. - Detection of Ball Bearing Conditions by an A.I. Approach, Sensors, Controls, and Quality Issues in Manufacturing, American Society of Mechanical Engineers, Production Engineering Division(Publication) PED v 55 ASME New York, p 13 - 21, 1991.

29. LIU, T. I. and IYER, N. R. - On-line recognition of roller bearing states, JAPAN/USA Symposium on Flexible Automation - Volume 1 ASME p 257 - 262 1992.

30. LIU, T. I. and IYER, N. R. - Diagnosis of roller bearing defects using neural networks, The international journal of advanced manufacturing technology v 8 p 210 - 215, 1993.

31. ANVAR, A. M. and MASOOD, S. H. - A Knowledge based system for vibration monitoring in turbo-generator machinery, The Institution of Engineers Australia Vibration and Noise Conference, Melbourne, 18 - 20 September 1990.

32. CHEVALIER, R., RICARD, B. and MOREL, J. - Validation of DIVA a turbine generator diagnosis expert system, Design, Repair and Refurbishment of Steam Turbines ASME pwr div v 13 p 23 - 28 1991.

33. LIDDLE, I. and REILLY, S. - Diagnosing vibration problems with an expert system, Mechanical Engineering p 54 - 55 Apr 1993.

34. DOUCET, R. P. and GODDARD, G. J. - Optimizing mine maintenance through true condition monitoring and total mine systems integration, Proceedings of the 8th international Congress on condition monitoring and diagnostic engineering management (COMADEM'95), Queen's University, Kingston, Ontario, Canada, 1995, v 2 p 425 - 430.

35. SERRIDGE, M. - What makes vibration condition monitoring reliable?, Noise and Vibration Worldwide, p 17 - 24 Sept 1991.

36. APPLICATIONS NOTE - Predictive maintenance through the monitoring and diagnostics of rolling element bearings, Bentley Nevada Corporation, August 1992.

37. RAY, A. G., - Monitoring rolling contact bearings under adverse conditions, Proceedings of the 2nd International Conference on Vibrations in Rotating Machinery, Institution of Mechanical Engineers, p187 - 194, 1 - 4 Sept., 1980.

38. ZENG, Y. and FORSSBERG, E. - Application of digital signal processing and multivariate data analysis to vibration signals from ball mill grinding, Trans. Instn. Min. Metall. (Sect. C: Mineral Process. Extr. Metall.), v 102, 39 - 43, Jan. - Apr., 1993.

39. MAEDEL, P. H. Jr and ESHLEMAN, R. L. - Shock and vibration handbook (Vibration standards), McGraw Hill, 1988.

40. CEMPEL, C. - Vibracoustic condition monitoring Ellis Horwood Limited, 1991.

41. GOODENOUGH, F. - Airbags boom when IC accelerometer sees 50 G, Electronic Design v 39 n 15 Aug 8 1991 7p.

42. GOODENOUGH, F., Heated actuator self-tests 50± accelerometer IC, Electronic Design v 40 n 13 June 25 1992 4p.

43. COLE, J. C. and BRAUN, D. - Applications for a capacitive accelerometer with digital output, Sensors, p 26 - 34, Feb., 1994.

44. SANTILLI, R. - Will microengineering affect your business?, Engineering science and education journal, p 244 - 248, December, 1994.

45. CHUEY, M. D., TONDA, R. D. and SHANBLATT, M. A. - Design of an intelligent bearing vibration monitor, Proceedings of the 1992 International Conference on Industrial Electronics, Control, Instrumentation and Automation, Power Electronics and Motion Control, v 3 p 1518 - 1523.

46. VAN EK, G., HENRY, T. A., and NIBLETT, M. K. - Design functionality for a maintenance field information system (MFIS), Proceedings of the 8th international Congress on condition monitoring and diagnostic engineering management (COMADEM'95), Queen's University, Kingston, Ontario, Canada, 1995, v 2 p 555 - 561.

47. BERRY, J. E. - How to track rolling element bearing health with vibration signature analysis, Sound and Vibration, v 25 n 11 p 24 - 35 1991.

48. SZROM, D. B. - Low Speed Bearing Analysis, Proceedings of the 10th annual meeting of the Vibration Institute, v 24 - 26 p 183 - 188 1988.

49. SCHLOSS, F. - Accelerometer noise, Sound and Vibration, vol 27 n 3 p 22 - 23 1993.

50. HAWMAN, M. W. and GALINAITIS, W. S. - Acoustic emission monitoring of rolling element bearings, Proceedings of the 1988 IEEE Ultrasonics Symposium p 885 - 889.

51. BERRYMAN, F., MICHIE, P., SMULDERS, A. and VERMEIREN, K. - Condition Monitoring - A new approach: A method of monitoring machines using a high frequency acoustic emission technique, Maintenance, p 3 - 9 July 1992.

52. SKVORTSOV, A. V. - Suitability of tool materials for eddy-current tool state monitoring, Soviet Engineering Research v 10 n 8 p 106 - 108 1990.

53. JORDAN, M. A. - What are orbit plots, anyway ?, Orbit, Dec. 1993.

54. PHILIPS, G. J. - Rotating machinery rolling element bearing performance monitoring using the fibre optic method, Reliability Stress Analysis and failure prevention methods in mechanical design published by ASME, p 37 - 43 Aug. 1980.

55. HARKER, R. G. and HANSEN, J. S. - Rolling element bearing monitoring using high gain eddy current transducers, ASME Paper 84-GT-25, October 1984.

56. SPENCER, D. B. and HANSEN, S. J. - A better way to monitor bearings, Hydrocarbon processing, January 1985.

57. KIM, P. Y. - A review of rolling element health monitoring (III): preliminary test results on eddy current proximity transducer technique, Institute of Mechanical Engineers Vibrations in rotating machinery conference C311/84 v 5 p 119 - 125 1984.

58. BRAUN, S. and DATNER, B. - Analysis of Roller/Ball bearing vibrations, Journal of Mechanical Design v 101 p 118 - 125 1979.

59. SU, Y.-T., LIN, M.-H. and LEE, M.-S. - The effects of surface irregularities on roller bearing vibrations, Journal of Sound and Vibration, v 165, n 3, p 455 - 466, 1993.

60. MCFADDEN, P. D. and SMITH, J. D. - Vibration monitoring of rolling element bearings by the high frequency resonance technique, Tribology International, v 17 n 1 p 3 - 10 Feb 1984.

61. MATHEW, J. and ALFREDSON, R. J. - The condition monitoring of rolling element bearings using vibration analysis, Journal of Vibration, Acoustic, Stress and Reliability in Design v 106 p 447 - 453 1984.

62. DYER, D. and STEWART, R. M. - Detection of rolling element bearing damage by statistical vibration analysis Journal of Mechanical Design v 100 p 229 - 235 1978.

63. TAYLOR, J. I. - Identification of bearing defects by spectral analysis, Journal of Mechanical Design v 102 p 199 - 204 1980.

64. MCFADDEN, P. D. and SMITH, J. D. - Model for the vibration produced by a single point defect in a rolling element bearing, Journal of sound and vibration, v 96 n 1 p 69 - 82, 1984.

65. MCFADDEN, P. D. and SMITH, J. D. - The vibration produced by multiple point defects in a rolling element bearing, Journal of Sound and Vibration, v 98(2) p 263 - 273 1985.

66. ERNST, A. - Digital linear and angular metrology verlag moderne industrie, 1989.

67. Incremental angle encoders, Dr Johannes Heidenhain GmbH. Publication No. 20873621.4/91.H, April 91.

68. ANON, Fundamentals of the electronic counters. Application note 200, p21 - 24, Hewlett Packard 1978.

69. ROUTSON, R., WANG, P., STARKEY, J. and DAVIES, P. - Torsional vibration measurements: phase 1 final report. PRF# 670-1288-1751, Purdue University School of Mechanical Engineering, West Lafayette, Indiana, March 1991.

70. WANG, P. *et al.* - A torsional vibration measurement system, IEEE Transactions on instrumentation and measurement, v 41, n 6, p 803 - 807, Dec, 1992.

71. LENAGHAN, J. T. AND KHALIL, A. I., - Condition/performance monitoring of internal combustion engine by continuous measurement of crank shaft twist angle, Proceeding of the COMADEM'90 conference, Brunel University, UK, p 114 - 119, 1990.

72. CITRON, C. J., O'HIGGINS and CHEN, L. Y., - Cylinder by cylinder engine pressure and pressure torque waveform determination utilizing speed fluctuations, SAE Paper No. 890486, p 131 - 145, 1989.

73. ANON, Optoelectronics Designer's Catalog Hewlett Packard 1991, p 1-22 - 1-30.

74. KHAN, A. F. *et al.* - Predicting the Remaining Life of Rolling Element Bearings, Proceeding of the Vibrations in Machinery Conference, ImechE, 1992.

75. FORD, R. A. J. - Rig for demonstrating vibration-based condition monitoring of rolling bearings, Australian Vibration and Noise Conference 1990, Institution of Engineers, Australia n 90 pt 9 p 37 - 40, Published by IE, Aust, Barton, Aust.

76. SIMMS, D. A., - PC based PCM telemetry data reduction system software, Instrumentation in the aerospace industry. Proceedings of the ISA Aerospace Instrumentation Symposium, v 36 p 263 - 273 1990.

77. YOERGER, E. J., *et al.* - Acoustic transient data acquisition and processing on a PC system, Conference Proceedings - Oceans '90 p 140 - 143 1990.

78. RAJA et al. - A SIMD multiple DSP microprocessor system for image processing, Microprocessors and microsystems, v 15 n 9 p 493 - 501 1991.

79. LIN, L., KABAY, S., and JONES, N. B. - DMA-based communications between PC and DSP, Microprocessors and microsystems, v 15 n 3 p 137 - 142 1991.

80. STEIGER, W. - Smooth continuous path motion generation for stepping motor driven manipulators, PhD Thesis, The Nottingham Trent University, 1994.

81. MACMANUS, L., ORTON, P. A., & THOMAS, P. D. - Digital data acquisition and control interface for DSP using a precision reciprocal timing architecture, Proceeding of the International Conference on Signal Processing and Technology, Boston, USA, 1995.

82. AYANDOKUN, K., ORTON P. A., SHERKAT, N., and THOMAS, P. D. - An optical rotary motion sensor for the real-time condition monitoring of rotating machinery, Condition Monitoring 94, Mervin H. Jones, p 657 - 668, Pineridge, Swansea, 1994.

83. OSMAN, H., and Martin, H. R. - Vibration propagation between surfaces in contact, Proceedings of the 8th international Congress on condition monitoring and diagnostic engineering management (COMADEM'95), Queen's University, Kingston, Ontario, Canada, 1995, v 2 p 663- 671.

84. FUKANGA, K - Introduction to statistical pattern recognition, 2nd edition, Academic Press Inc., 1990.

85. THACKER, N. A., AHERNE, F. J. and ROCKETT, P. I. - The Bhattacharyya metric as an absolute frequency measure for frequency coded data, Proceedings of IAPR Statistical Techniques in Pattern Recognition, Prague, Czech Republic, June 1997

86. AYANDOKUN, K., ORTON P. A., SHERKAT, N., THOMAS, P. D. and POLIAKOFF J. F. - The Incremental Motion Encoder: a sensor for the integrated condition monitoring of rolling element bearings in machine tools, Machine tool, In-line, and Robot sensors and controls, SPIE proceeding vol. 2595, 1996.

87. AYANDOKUN, K., ORTON P. A., SHERKAT, N., THOMAS, P. D. - Tracking the development of rolling element bearing faults with the optical incremental motion encoder, Proceedings of the 8th international Congress on condition monitoring and

diagnostic engineering management (COMADEM'95), Queen's University, Kingston, Ontario, Canada, 1995, v 2 p 649- 655.

88. AYANDOKUN, K., ORTON P. A., SHERKAT, N., THOMAS, P. D. - Smart bearings: Developing a new technique for the condition monitoring of rotating machinery, Proceeding of INES'97 IEEE international conference on intelligent engineering systems 1997, Budapest, Hungary, Sept. 15 - 17, 1997 p 505 - 511.

89. BAYER, O. - A new generation: ABS capable wheel bearings, SAE International Congress and Exposition, Detroit, Michigan, Feb. 25 - Mar. 1, 1991.

90. COURTWRIGHT, M. L. - Instrumented bearings get rolling with ABS, *Machine Design*, v 64 n 1 p 84 - 88, 1992.

Appendix A

Incremental motion encoder patent application

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : G01D 5/34	A1	(11) International Publication Number: WO 92/19936 (43) International Publication Date: 12 November 1992 (12.11.92)

(21) International Application Number: PCT/GB92/00811

(22) International Filing Date: 5 May 1992 (05.05.92)

(30) Priority data:
9109501.8 2 May 1991 (02.05.91) GB

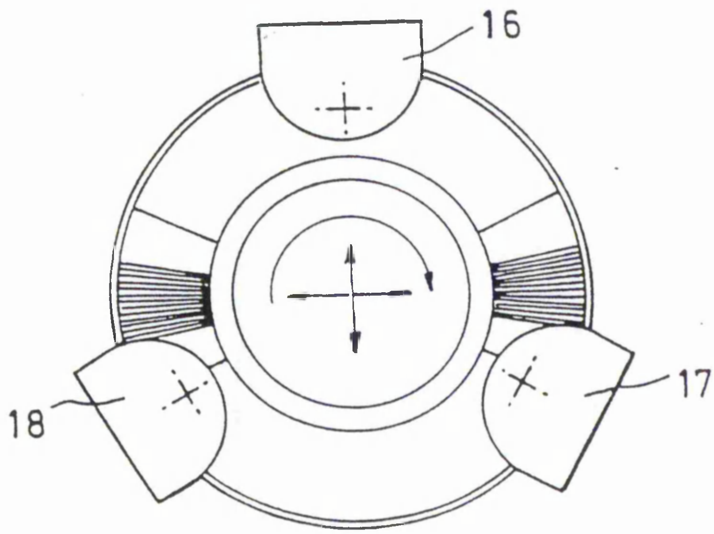
(71)(72) Applicant and Inventor: ORTON, Paul, Anthony [GB/GB]; 33 Farm Road, Beeston, Nottingham NG9 5BZ (GB).

(74) Agent: MESSULAM, Alec, Moses; A. Messulam & Co., 24 Broadway, Leigh on Sea, Essex SS9 1BN (GB).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), MC (European patent), NL (European patent), SE (European patent), US.

Published
With international search report.

(54) Title: SHAFT DISPLACEMENT MEASURING SYSTEM



(57) Abstract

A known method of measuring angular displacement of a shaft (11) using an optical grating mounted for rotation with the shaft is applied at two or more angularly spaced locations (16, 17 and 18) and the resulting measurements are processed to produce a measurement of relative displacement of the shaft in a direction transverse to the shaft. Two locations are needed to measure displacement in a given direction. Three or more locations are necessary to provide full information about the displacement of the shaft.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FI	Finland	ML	Mali
AU	Australia	FR	France	MN	Mongolia
BB	Barbados	GA	Gabon	MR	Mauritania
BE	Belgium	GB	United Kingdom	MW	Malawi
BF	Burkina Faso	GN	Guinea	NL	Netherlands
BG	Bulgaria	GR	Greece	NO	Norway
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IE	Ireland	RO	Romania
CA	Canada	IT	Italy	RU	Russian Federation
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland			SN	Senegal
CI	Côte d'Ivoire	KR	Republic of Korea	SU	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
CS	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		
ES	Spain	MG	Madagascar		

Title

Shaft Displacement Measuring System

Field of the invention

This invention relates to a system for measuring relative displacements of a rotating shaft in a direction transverse to the axis of the shaft.

Background of the invention

Displacements of a shaft axis occur at frequencies comparable to or greater than the frequency of rotation of the shaft in many machines and apparatus. A shaft which supports a grinding wheel in a grinding machine is subject to such displacement both as a result of imbalance of the rotating masses and as a result of external forces applied to the grinding wheel when it engages a workpiece.

One known method of measuring displacement uses a probe to mechanically follow the shaft and a transducer to measure the displacement of the probe. For complete information two probes at right angles are required.

The present invention uses a known method of measuring angular displacement of a shaft. The method is applied at two or more angularly spaced locations and the resulting measurements are processed to produce a measurement of relative displacement of the shaft in the required transverse direction. Two locations are needed to measure displacement in a given direction. Three or more locations are necessary to provide full information about the displacement of the shaft.

Summary of the invention

According to the invention a displacement measuring system for indicating relative displacement of a shaft transverse to its axis, comprises:

a) a first detector at a first station for producing electrical signals in a first channel representative of angular displacement of the shaft as measured at the first station;

b) a second detector at a second station angularly spaced from the first station for producing electrical signals in a second channel representative of angular displacement of the shaft as measured at the second station;

c) means for processing information from the first and second channels to generate an indication of relative displacement between the detectors and the shaft in a direction transverse to the axis of the shaft.

It is an important advantage of the invention that the shaft displacement information is obtained from angular position detecting devices so that full information as to shaft angular motion and transverse displacement can be derived without using two separate measuring systems.

Brief description of the drawings

The invention will now be described with reference to specific embodiments shown in the accompanying drawings in which:

Figure 1 shows a known form of optical shaft encoder;

Figure 2 shows an embodiment of the invention having two measuring heads spaced at 180° approximately and able to

(

measure shaft displacement in a direction transverse to the diameter joining the measuring heads;

Figure 3 shows a second embodiment in which three measuring heads spaced at angular intervals of approximately 60° are used for measuring shaft displacements in any direction transverse to the axis of the shaft;

Figure 4 shows how shaft displacement is derived from the measurements made in the embodiment of Figure 2;

Figure 5 shows how shaft displacement in a plane transverse to the axis of the shaft is derived from the measurements made in the three head embodiment of Figure 3;

Figure 6 is a block diagram of the circuit through which the output signals of the three measuring heads in the embodiment of Figure 3 are processed to provide signal suitable for computer analysis; and

Figures 7 to 10 show applications of the shaft displacement measuring system of the invention.

Detailed description of the preferred embodiments

The shaft displacement measuring system of the invention is an instrument for measuring the relative displacement of a rotating shaft in directions transverse to the axis of rotation of the shaft. In its simplest form (Figure 2) the shaft displacement measuring system is capable only of measuring displacements in a single direction.

The shaft displacement measuring system uses optical shaft encoders of a known kind which produce as output an electrical signal comprising a train of pulses of frequency proportional to the angular velocity of the shaft. As shown in Figure 1, an encoder of this kind comprises a transparent disc 10 arranged for rotation with the shaft 11 from which measurements are required to be taken. A measuring head 12 mounted on a fixed base (not shown) comprises a light source on one side of the disc 10 projecting a beam of light onto a light sensitive semi- conductor device on the other side of the disc.

A radial optical grating on the disc interrupts the beam as the shaft rotates. The output signal from the measuring head 12 comprises a train of pulses, each pulse corresponding to the passage of a space in the grating past the beam.

Normally such encoders are used to measure angular position and speed only. Motion of the disc relative to the shaft axis is considered a source of error when estimating angular position. For this reason optical shaft encoders have their own precision bearings independent of the bearings of the shaft to be measured and the disc is driven by a bellows coupling which isolates the encoder from shaft vibrations.

In the shaft displacement measuring system of the invention the disc 10 is fixed to the shaft 11 and two diametrically opposed measuring heads 14 and 15 are fixed to the base of the apparatus in which the shaft 11 is journalled.

The output signals from measuring heads 14 and 15 are processed in a manner to be described in more detail below in relation to the embodiment of Figure 3 to develop a succession of pairs of angular position measurements. These angular position measurements are processed in a data processor (not shown) which may be a general purpose computer which derives the displacement of the shaft from the relationship illustrated in Figure 4. If l is the distance between the two measuring stations and θ is the

(

difference between the angular position of the shaft measured at the first station and the angular position of the shaft measured at the second station, the displacement of the shaft from a line joining the two measuring stations is given by:'

$$\text{Displacement} = (\pi - \theta) * 1/4 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

The operation of the embodiment of Figure 2 can be understood by reference to Figure 4. Assume initially that the shaft centre-line is in line with the two measuring heads. In this case the angular position measurement derived from the signal from the first station will be exactly πradians different from the angular position measurement derived from the signal from the second station. Using the formula given above the data processor will indicate zero displacement. If the shaft remains in the same position the processor will receive a succession of pairs of angular position measurements with a π radians difference.

Upon displacement of the shaft to the left as viewed in Figure 4 the difference between the pairs of angular position measurements will decrease and the displacement of the shaft will be given by formula (1) above.

It will be appreciated that the two head system of Figure 2 can only measure displacement perpendicular to the diameter on which the heads are positioned, that is horizontally as shown in Figure 4. To measure vertical displacement, two additional heads at opposite ends of a horizontal diameter could be used. It is more economical to use three heads 16, 17 and 18 spaced apart 60° as shown in Figure 3.

The time of passage of individual grating lines does not enable angular position to be measured with sufficient accuracy for many applications. A system which provides for interpolation between successive grating lines and measurement of angular position to an accuracy much greater than the angle between adjacent grating lines will now be described with reference to the embodiment shown in Figures 3, 5 and 6.

Referring to Figure 6, an interface circuit for connecting the output of measuring heads A, B, and C, indicated respectively by reference numerals 16, 17 and 18 in Figure 3 to a general purpose computer (not shown), comprises three channels A, B and C plus an additional channel D for collecting a once per revolution synchronisation signal. The synchronisation signal may be produced by the passage of a single grating line on the disc 10 past by an additional head (not shown) at a different distance from the shaft axis than the main measuring heads A, B and C. The signal in channel D thus provides a reference pulse corresponding to a specific angular position of the shaft. This pulse on line D sets a flip-flop used for resetting a counter 21 and for sending an interrupt via OR gate 22 and interrupt line 23 to the computer.

Each of the channels A, B and C comprises a Schmitt Trigger 24 which converts the pulses from the respective photo-sensitive semiconductor device into a square wave on line 25 connected to a latch 26. Each latch 26 is connected by a bus 29 to a FIFO (first in first out) buffer 27.

The latches 26 are connected by a timer bus 28 to the synchronous counter 21. The counter 21 counts pulses from a precision clock circuit 30.

All three FIFO buffers are connected to the data bus 31 and to the Address Decoder 32 of the computer.

Line 25C is connected to one input of the OR circuit 22 so that in addition to the interrupt sent when the synchronising signal on line D is received, an interrupt is also sent to the computer once for each grating line. These interrupts enable the computer to request the output of the buffers at intervals corresponding to the period between the passage of grating lines past the measuring heads by sending signals to the buffers on lines 33 via the address decoder 32.

In operation, the clock 30 runs continuously and each of its output pulses increments the counter 21. When the synchronising pulse on channel D sets the flip-flop 20, the counter 21 is set to zero.

After a delay dependent upon the angular position of the disc as seen at the three measuring points, pulses arrive at the latches 26 from channels A, B and C. In each channel the count taken by the latch from the timer bus is transferred to the FIFO buffer at the instant the leading edge of the pulse on line 25 reaches the latch.

Sets of three counts are taken from the buffers 27 via data bus 31 under control of address decoder 32 for storage in table form in the computer for subsequent processing as described below. The interface circuit enable the computer to collect reading from which the shaft displacement will be calculated at the rate of one set per grating line.

The computer looks at a set of three readings, one from each channel. Each reading gives the time taken for the shaft to turn to the related angular position. Using an input for shaft angular velocity taken by counting the frequency of interrupts the computer converts the stored readings into angle information.

In Figure 5, A, B, and C indicate the positions of the three measuring heads. These are spaced apart a distance l at the corners of an equilateral triangle. The position of the shaft axis relative to the measuring heads is given by its distance m from A and by the $^\wedge\alpha$ between the vertical through A and the line AS joining the shaft axis to A. The computer stores sets of values for $^\wedge A$, $^\wedge B$ and $^\wedge C$ being the angular position of the shaft as measured at A, B, and C expressed in the form of the angle through which the shaft has rotated since detection of the last synchronising pulse. It can be shown that:

$$^\wedge a = {}^\wedge A - {}^\wedge B \quad \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

$$\text{and} \quad ^\wedge b = 2\pi - ({}^\wedge A - {}^\wedge C) \quad \ldots\ldots\ldots\ldots(3)$$

Using the sine formula for the two triangles shown in Figure 5, it can be shown that for small angles α:

$$\alpha = \frac{\sin a \sin (5\pi/6 - b) - \sin b \sin (5\pi/6 - a)}{\sin a \cos (5\pi/6 - b) + \sin b \cos (5\pi/6 - a)} \quad \ldots(4)$$

$$m = \frac{l \sin (5\pi/6 + \alpha - a)}{\sin a} \quad \ldots\ldots.(5)$$

Horizontal Displacement of the shaft from the centre of the measuring heads is approximately:

$$\alpha * m \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6)$$

Figure 7 shows an application of the shaft displacement measuring system of the invention to the measurement of shaft lateral motion relative to its bearings. The shaft 40 is journalled in bearing mounted on a machine base 42 by flexible mountings 43. The disc with optical grating 44 is fixed to the shaft 40 and a platform 45 carrying the measuring heads is fixed to the bearing by solid mountings 46. The output from the computer gives information as to the vibration of the shaft relative to its bearing as it rotates.

(

The set up shown in Figure 8 differs from that shown in Figure 7 in that the platform 45 is fixed to the base 42 by solid mountings 47 and so that the computer output indicates movement of the shaft and its bearing relative to the machine structure 42.

Figures 9 (side view) and Figure 10 (end view) show how the shaft displacement measuring system can be used to gauge a component 49 fixed to the shaft 40. In this case, the measuring head platform is mounted for sliding movement relative to the machine base 42 on vertical guides 48 and a surface following probe 52 fixed to the platform ensures that the platform follows the contour of the component 49. Since only vertical movement is involved, two measuring heads 50 and 51 are sufficient. It should be noted that in this example the shaft remains fixed and the shaft displacement measuring system is measuring the movement of the platform relative to the shaft.

The shaft displacement measuring system may also be used for the control of automatic bearings such as hydrostatic bearings in which fluid pressure in angularly spaced bearing pockets is increased or decrease to maintain shaft clearance from the bearings within predetermined limits. The shaft displacement measuring system provides information which can be used to controlled fluid pressures within the system to maintain the required clearances.

Similarly, in an active magnet bearing, the output of the shaft displacement measuring system can be used to control current to electromagnetic bearings to maintain required air gaps between a shaft and the electromagnets with varying loads.

The disc 10 may be replaced by a drum with a reflective grating formed on its outer surface and the light sources and detectors outside the drum. Alternatively the drum could be transparent with the light sources on the inside and the detectors on the outside, or vice versa.

(

CLAIMS

1. A displacement measuring system for indicating relative displacement of a shaft transverse to its axis, comprising:

a) a first detector at a first station for producing electrical signals in a first channel representative of angular displacement of the shaft as measured at the first station;

b) a second detector at a second station angularly spaced from the first station for producing electrical signals in a second channel representative of angular displacement of the shaft as measured at the second station;

c) means for processing information from the first and second channels to generate an indication of relative displacement between the detectors and the shaft in a direction transverse to the axis of the shaft.

2. A displacement measuring system as claimed in claim 1 including a disk bearing an optical grating mounted for rotation with the shaft, each detector comprising a photosensitive device adapted to produce electrical signals corresponding to the passage of lines of the grating past the respective station.

3. A displacement measuring system as claimed in claim 1 or claim 2 including a third detector at a third station angularly spaced from the first and second stations for producing electrical signals representative of angular displacement of the shaft as measured at the third station, in which the processing means uses information from all three channels to generate an indication of the position of the shaft in a plane transverse to its axis.

4. A displacement measuring system as claimed in claim 3 in which the three detectors are spaced apart approximately 120°.

5. A displacement measuring system as claimed in claim 1 or claim 2 in which the two detectors are at opposite sides of the shaft spaced about 180° apart and the processor is adapted to indicate displacement of the shaft in a direction at right angles to a line joining the detectors.

6. A displacement measuring system as claimed in claim 5 including third and fourth detectors at third and fourth stations angularly spaced approximately 90° from the first and second stations for producing electrical signals representative of angular displacement of the shaft as measured at the third and fourth stations, in which the processing means uses information from the third and fourth channels to generate an indication of the position of the shaft in a direction at right angles to a line joining the detectors.

7. A shaft angular position, angular velocity and displacement measuring system comprising a system as claimed in any one of the preceding claims including means for developing angular position and angular velocity information from the output of one of the detectors.

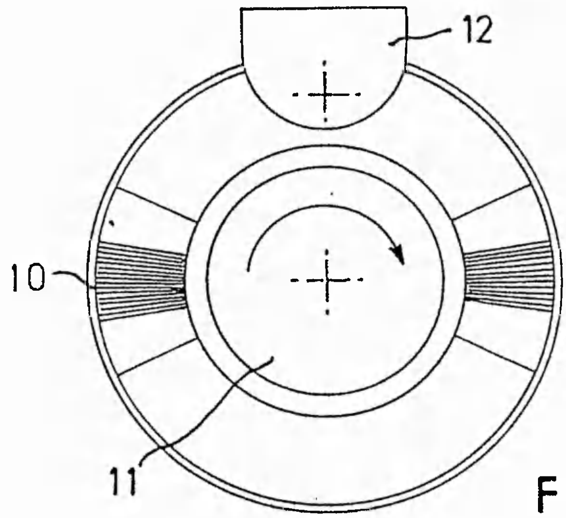


FIG.1.

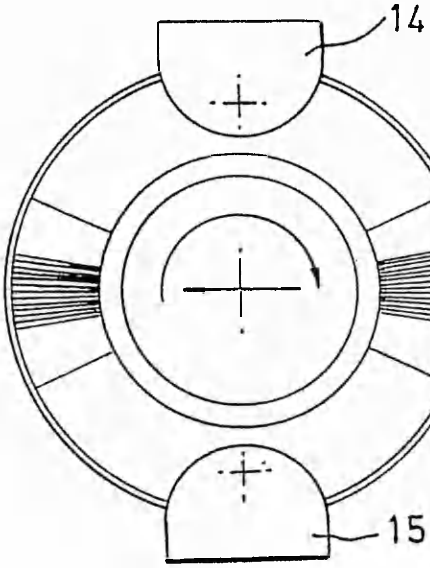


FIG.2.

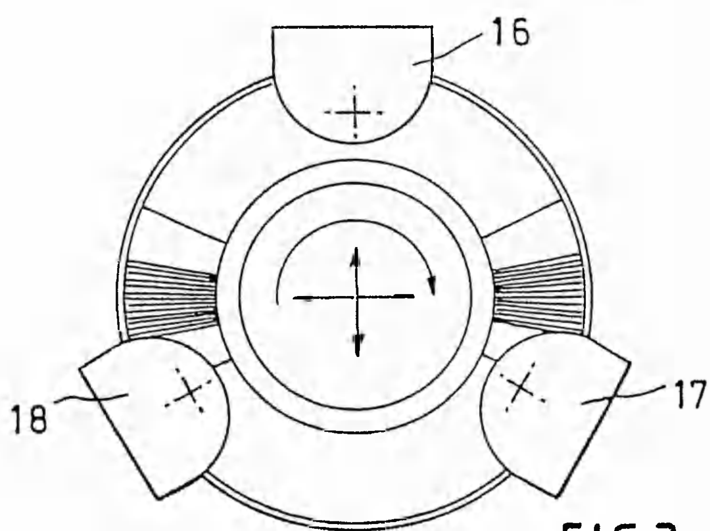


FIG.3.

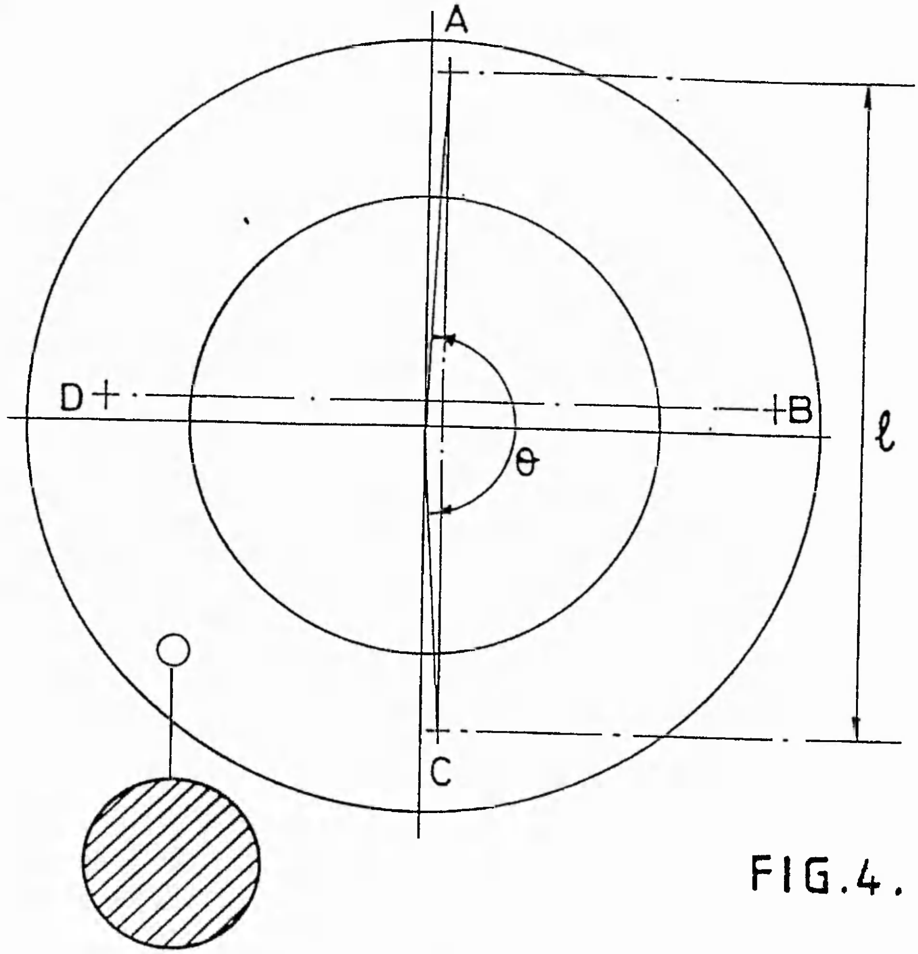


FIG.4.

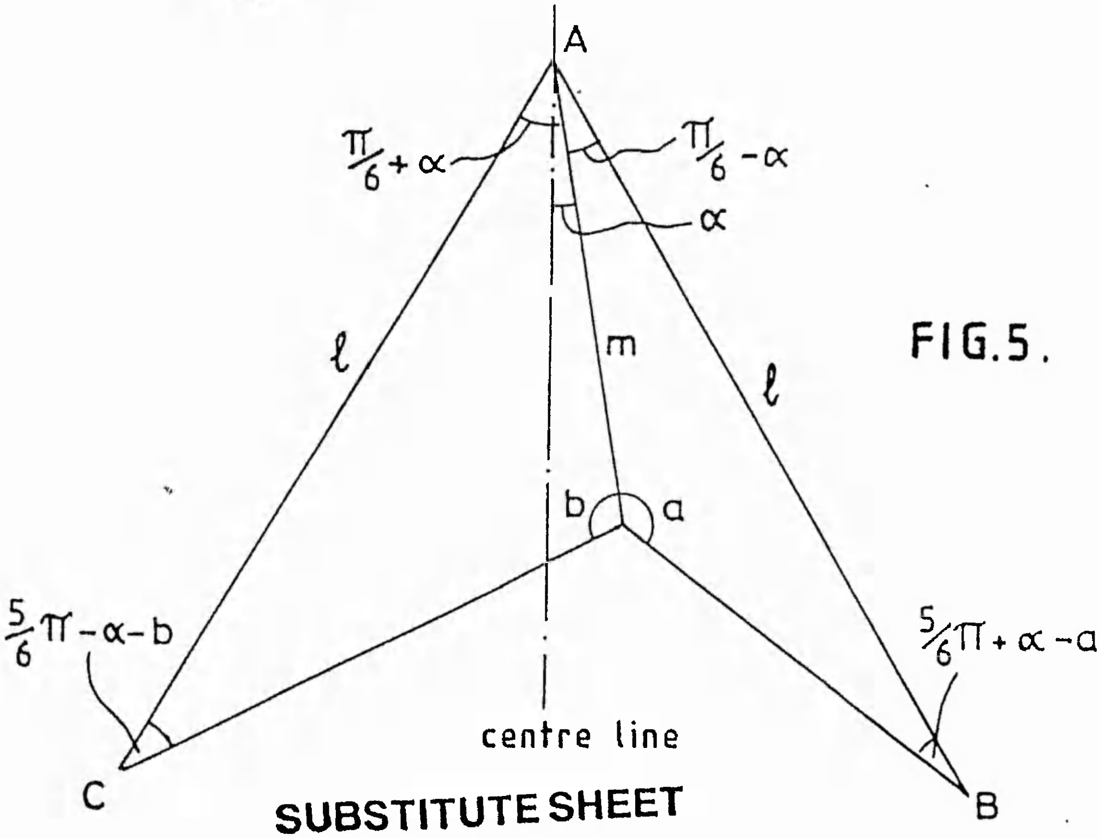
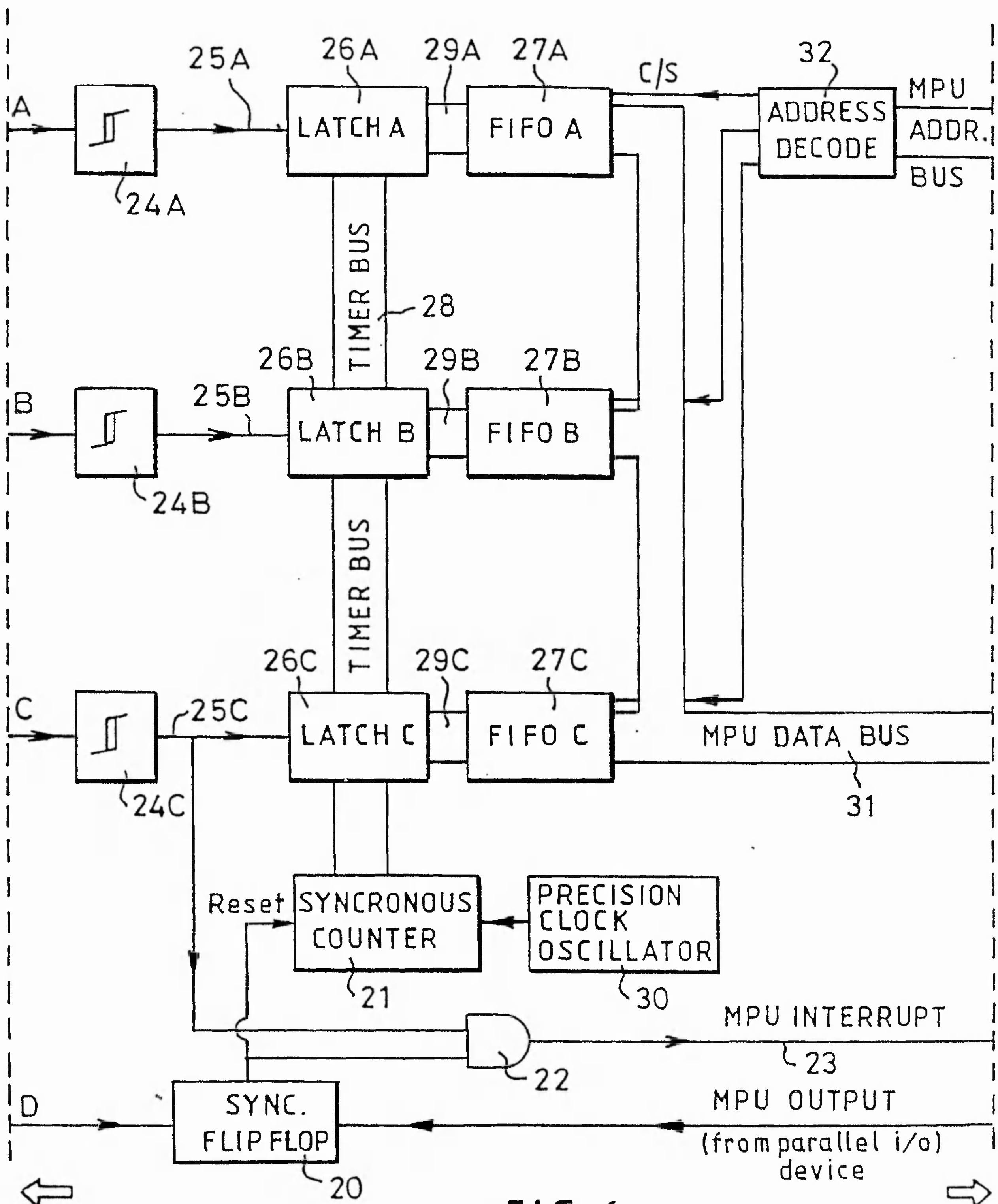


FIG.5.

SUBSTITUTE SHEET



A
B
C
D
24A
24B
24C
25A
25B
25C
26A
26B
26C
29A
29B
29C
27A
27B
27C
LATCH A
LATCH B
LATCH C
FIFO A
FIFO B
FIFO C
TIMER BUS
TIMER BUS
28
C/S
32
ADDRESS DECODE
MPU ADDR. BUS
MPU DATA BUS
31
Reset
SYNCRONOUS COUNTER
21
PRECISION CLOCK OSCILLATOR
30
MPU INTERRUPT
23
22
SYNC. FLIP FLOP
20
MPU OUTPUT
(from parallel i/o) device

FIG.6.

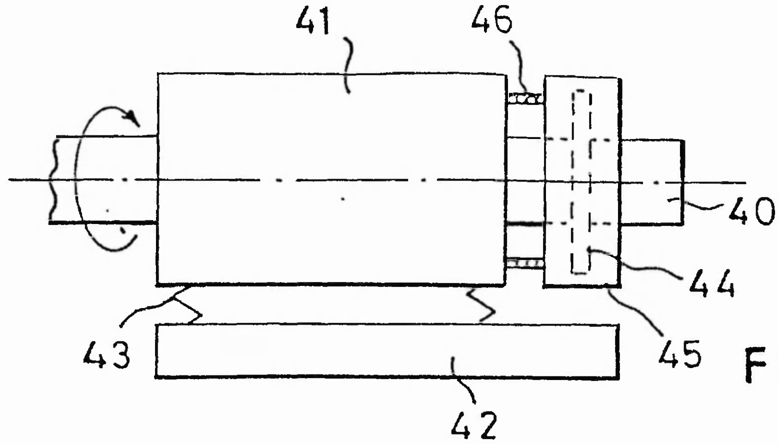


FIG.7.

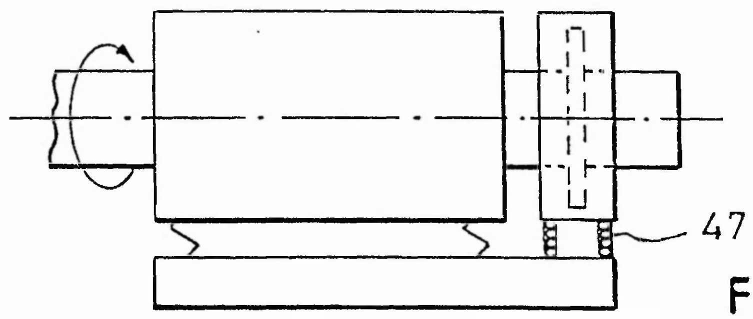


FIG.8.

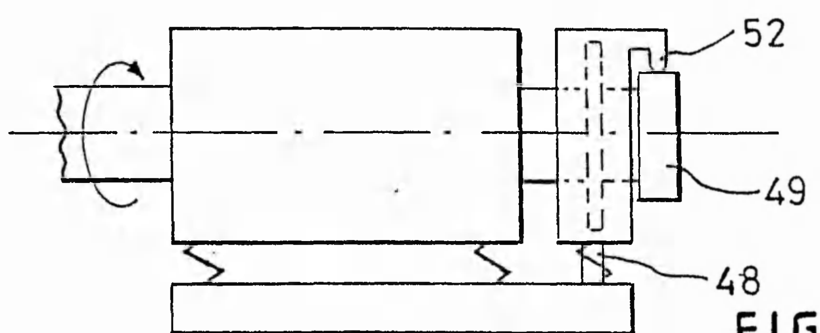


FIG.9.

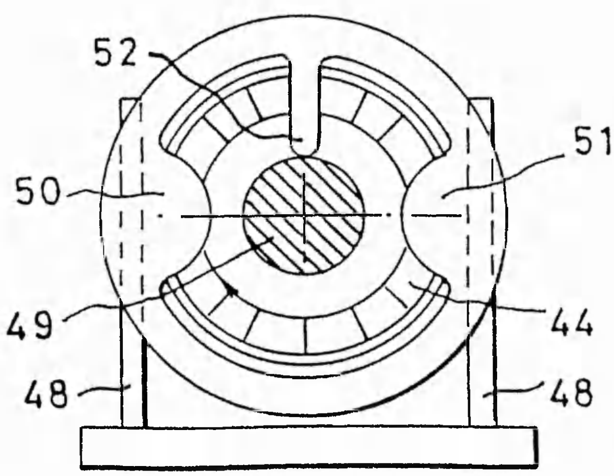


FIG.10.

INTERNATIONAL SEARCH REPORT

International Application No PCT/GB 92/00811

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[6]

According to International Patent Classification (IPC) or to both National Classification and IPC

Int.Cl. 5 G01D5/34

II. FIELDS SEARCHED

Minimum Documentation Searched[7]

Classification System	Classification Symbols
Int.Cl. 5	G01D

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched[8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT[9]

Category[o]	Citation of Document,[11] with indication, where appropriate, of the relevant passages[12]	Relevant to Claim No.[13]
X	GB,A,2 067 282 (DR. JOHANNES HEIDENHAIN GMBH) 22 July 1981 see the whole document ---	1-3,5,6
X	GB,A,2 146 765 (VEB CARL ZEISS JENA) 24 April 1985 see the whole document ---	1-3,5,6

[o] Special categories of cited documents :[10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
12 AUGUST 1992	27. 08. 92
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	BATTESON A.

1

Form PCT/ISA/210 (second sheet) (January 1985)

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. GB 92008[illegible] SA 5[illegible] .5

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information. 12/08/92

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
GB-A-2067282	22-07-81	DE-A-	2952106	02-07-81
		AT-B-	385124	25-02-88
		CH-A-	649629	31-05-85
		FR-A,B	2473171	10-07-81
		JP-B-	3018130	11-03-91
		JP-A-	56098604	08-08-81
		US-A-	4403859	13-09-83
GB-A-2146765	24-04-85	DE-A-	3427047	21-03-85
		NL-A-	8402650	01-04-85

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

Appendix B

The development of the data capture interface

B. The development of the data capture interface

The data capture interface in the experimental system is composed of two parts. The timer counter circuit board, to which the encoder output square wave signals are connected, and the embedded computer hardware by means of which timer counter values are taken from the circuit and stored before transfer to the processing computer. This part of the system thus acts as a link between the outside world, represented by the encoder motion, and the internal computations of the system. The creation of this component in the experimental system accounted for the largest part of the time and effort spent on the project in its first year and proved the most difficult in practical terms to complete. This was mainly due to the requirement for the design and implementation of novel hardware. Not only did the timer counter circuit board have to be developed, but a number of other standard components had to be utilised to create a working system. This appendix describes the development of the data capture interface from the original design to the hardware configuration successfully used in the experimental system. It explains why certain design ideas were implemented and others were not.

B.1 The data capture interface design

At the start of this project the form of the data capture interface was based on the design outlined in the IME patent application [2]. This called for a timer counter circuit board that generated interrupts to a microprocessor. This would collect the timer counter data for storage and subsequent processing. Small details of this design had been slightly amended during the initial requirements analysis stage of the preceding MSc project by the removal of some unnecessary elements. The block diagram of the data capture interface at the start of the project is reproduced in figure B-1.

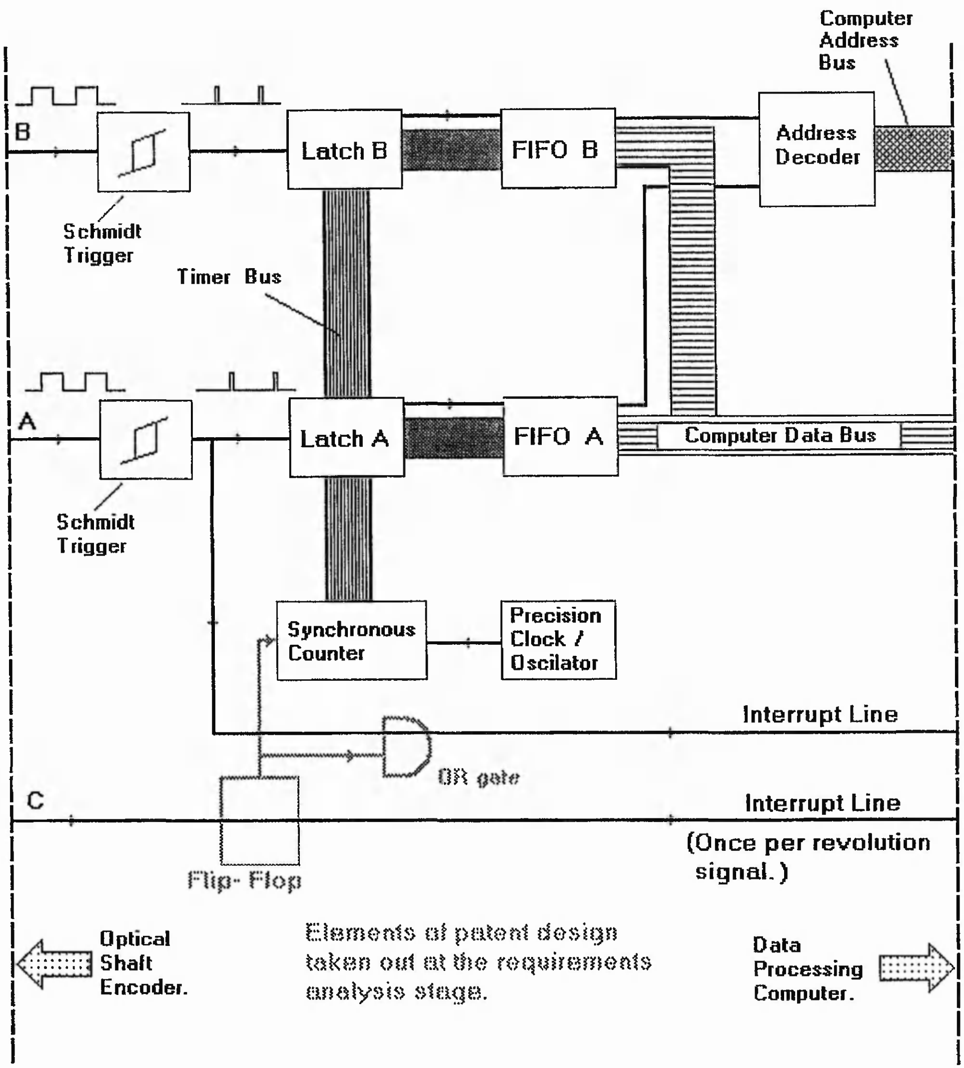


Figure B-1. The first design of the IME data capture interface based on the IME patent application [2].

The design in the patent application was derived from the design of the data logger previously used by Orton for precision angular velocity measurement with a single read head encoder [1]. For the multiple readheads of the IME it was assumed that it would be possible to simply extend the circuit that had been used for the data logger by duplicating its single set of components. There had been no attempt to make an application specific analysis of the unique requirements of the IME. Of greater importance to the actual operation of this design was the poor analysis of the electronic design requirements. This resulted in a circuit design which was faulty from the beginning. The author's inexperience in dealing with electronic circuits and components meant that this fact was slow to be realised and subsequently rectified.

In figure B-1 the address and data buses are shown are those of a VME based mono-board MC68000 computer. The FIFOs shown were two 68230 parallel interface timer chips. These were FIFOs only in the sense that they were capable of storing two input data values before being read by the computer. It is important to note that this use of FIFOs is only required by this design because of the way it was intended that data should be gathered sequentially from each channel by a single interrupt service routine sensitive to the arrival of data on only one channel.

The development of the data capture interface from its beginnings with this design to the conceptual computer architecture has been driven by the need to analyse the system from first principles in order to correctly design a hardware configuration to adequately meet the operating requirements of the experimental IME system.

B.2 The refinement of the timer counter circuit

This section outlines the main problem areas that were discovered with the timer counter circuit's initial design and construction. It describes the solutions to each problem that were implemented to create the circuit used in the experimental system. The fine detail of all the iterative stages that led to each solution is not presented.

After the initial design of the data capture interface was implemented and the system tested, it became clear that there were many errors in the timer counter data captured. Many displacement values generated from the raw data were obviously significantly out of range of the general trend.

The method adopted to trace the sources of these errors in the data capture interface was the same throughout the period of its testing and development. The output signal generated by one read head of the IME was split between the two, later four, inputs of the circuit board. This method allowed for a pair of initially synchronised signals to be propagated through the circuit. The relative arrival time of the signal's square wave edges is the critical parameter the circuit is required to record. The differences in the relative arrival time of signal edges at various points in the circuit were monitored by a logic analyser and used to trace the sources of the errors in the data collected by the computer system. If all had been working correctly then the raw counter data collected by the computer system would consist of a series of identical 16-bit values. A simulated fragment of the data collected at the start of the investigation is shown in table B-1 with examples of the types of errors commonly seen in the data at that time. In the actual data they were much more widely distributed.

Error 1 was the most obvious. Instead of being identical, pairs of values on the two channels showed a variable difference. Errors of type 2 occurred where the lower byte of a value recorded for either channel would be zero. At certain places on channel 1 it was clear that a value had not been recorded, leading to a doubling in the difference between successive values and the displacement of the two channels with respect to one another. This was an error of type 3. On channel 2 it was observed that some data values would be recorded twice giving an error of type 4. This would also lead to the displacement of the two data records with respect to each other. Errors of type 5 occurred where a value was recorded that was out of sync with the incoming edges, again the data would be displaced.

Table B-1

Angular position	channel 1	channel 2
...	...	...
543	18F4	18F4
544	1BE7 error 1	1C40 error 1
545	1F8C	1F33
546	2200 error 2	22D8
547	2624 ⇓ error 3	2624
548	2CBC ⇑ error 3	2970
549	3008	2CBC error 4
550	3354	2CBC error 4
551	34CE error 5	3008
552	36A0	3354
...	...	...

Figure B-2 shows a simplified diagram of the counter timer circuit at the start of the project. In the original, data logger circuit, built with a single input channel, there was no need to take relative propagation times into consideration. The version built for this project used the same design, but with two channels. The inherent problem with this design was the fact that the path lengths between the input connectors and the latches had different relative propagation times. This resulted in the errors of type 1 described above. Signal edges which were synchronous at the inputs to the circuit would be skewed by a variable amount relative to each other by their passage across all the elements in the circuit prior to reaching their respective latches. This was recorded as the variable differences between timer values.

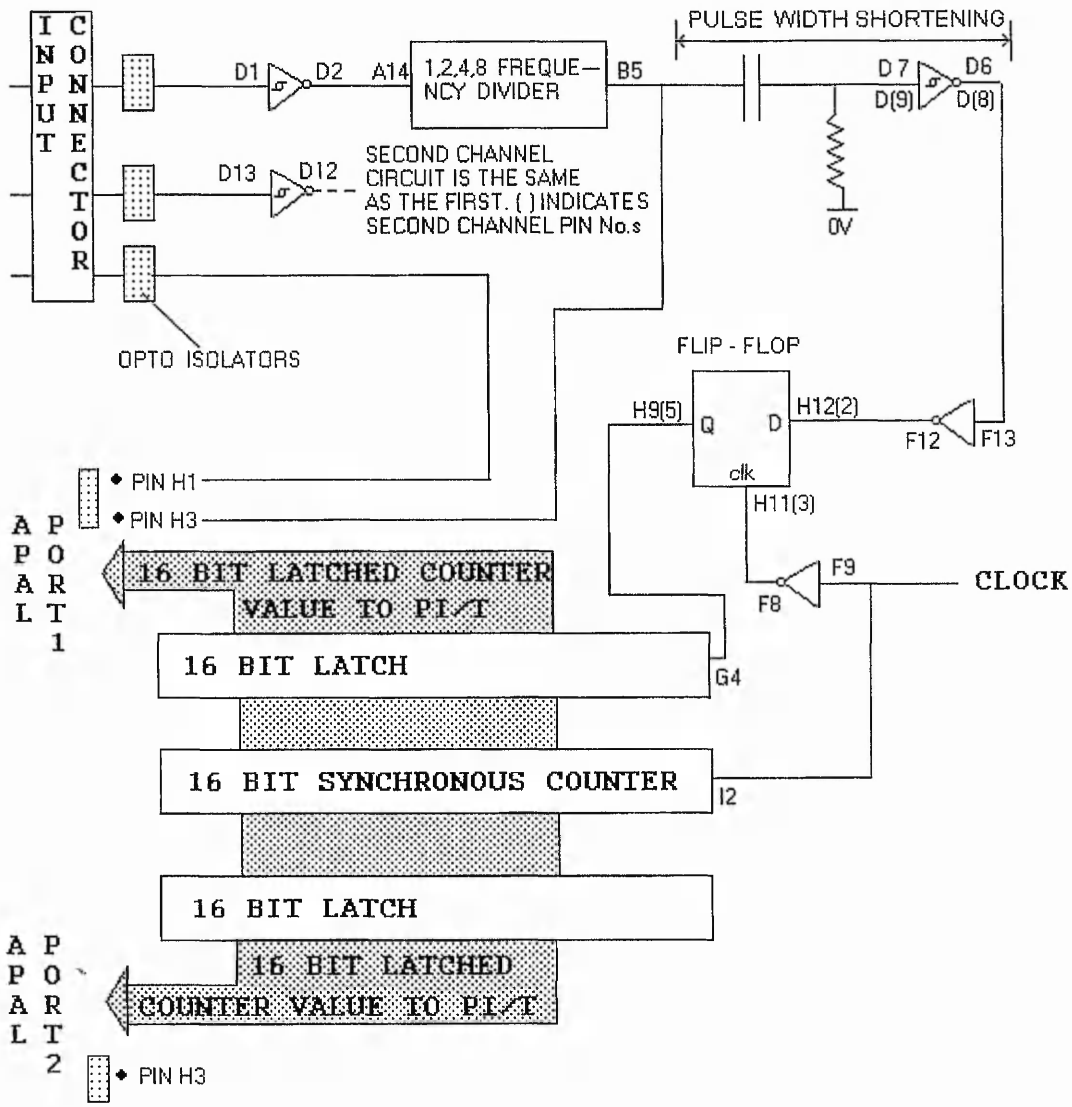


Figure B-2. Simplified diagram of the timer counter circuit at the start of the project.

Initially the main problem area was the pulse width shortening part of the circuit. This produced a short pulse from the wide square wave of the encoder. The width of this short pulse and the relationship of its edges to those of the original wave were determined by the exact value of the capacitor in the circuit. As no two capacitors are exactly the same this produced the largest single component of time difference between edges on the two channels. Decreasing the value of the capacitors, so the shortened pulse width was as close to the value of the time difference as possible, was the first remedy tried. This had the effect of reducing the skew, but when the speed of the clock in the circuit was increased from 10 MHz to 20 MHz the error doubled.

Gradually the circuit was redesigned from first principles with its form determined by its function. The circuit is required to latch a count value on the falling edge of an input square wave, then signal this to an external computer. The square wave edges from the encoder did not need any further shaping and the pulse shortening was an unnecessary complication. The same was true of the frequency divider which again came from the original single channel circuit. The schmidtt triggers and opto-isolators on the inputs were also removed as there was a significant degree of variable delay contributing to the skew as the two signals propagated across these components. The opto-isolators on the signal outputs were also removed. The circuit was streamlined until the input signals drove the latches directly. The skew causing error 1 was then eliminated. The difference between values was reduced to 0 +/- 1 clock cycle. Figure B-3 shows a simplified diagram of this circuit. Appendix C contains detailed schematics of the timer counter circuit.

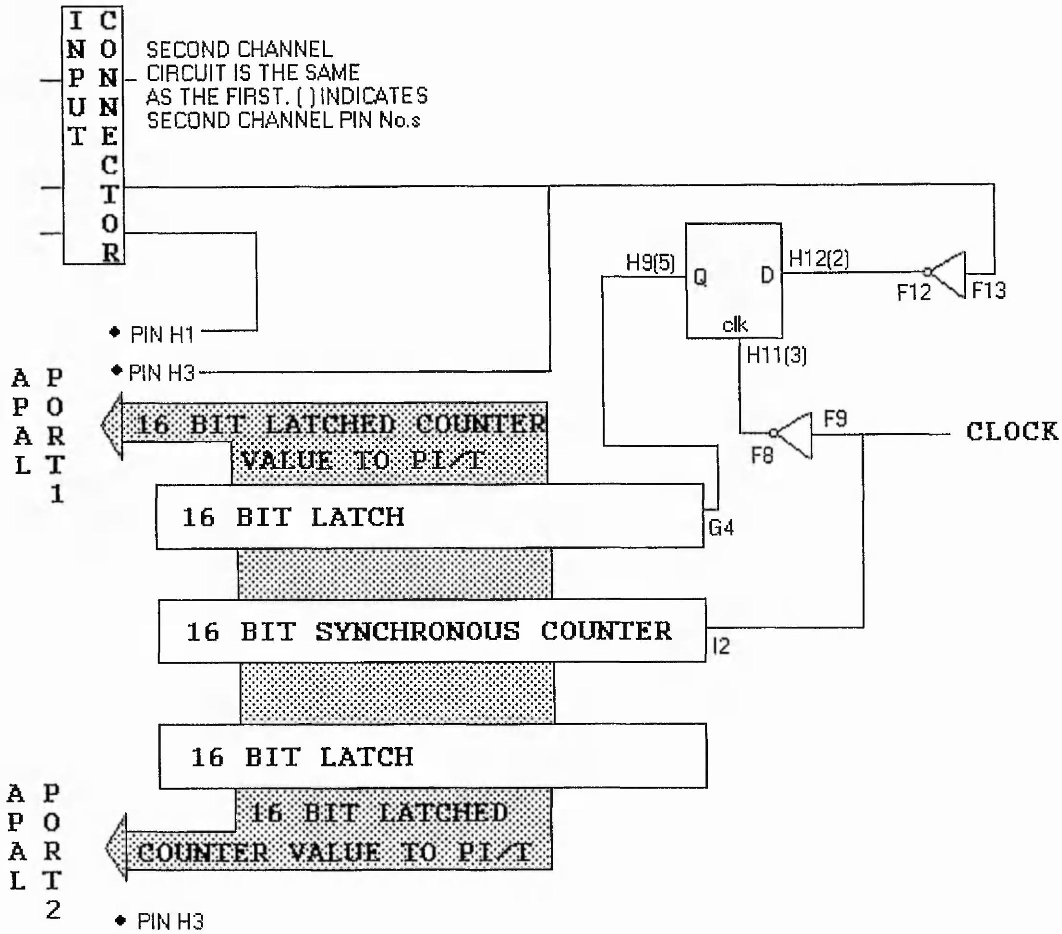


Figure B-3. Simplified diagram of the current interface circuit.

The errors of the second type were traced to the way the data capture interface design attempted to use the 68230s as FIFOs. As shown in figure B-3 the input signal is split by the circuit with one path going directly to the H3 latching pin of the 68230 and the other going to the latch on the circuit board. It was found that the edge that went directly to the 68230 arrived there before the latch on the circuit had become stable with the timer value. The bits in the lower byte of the counter were changing too fast for the set up and hold time of the 68230s. This resulted in the loss of the lower byte of the counter value recorded by the system. The initial solution to this, within the context of the original design, was to place 400ns delay lines in the circuit at just before the output leading to the H3 pin. These allowed the latch time to become stable before the 68230 latched the data itself. This problem only occurred because the 68230 hardware was being used to hold the data prior to being read by an Interrupt Service Routine (ISR), as dictated by the original design. This was a design error as the latch on the timer counter board was effectively not being used. The implementation of the conceptual computer architecture corrected this by polling the immediate value on the 68230 input. In this case the latches were able to go stable in the time taken for the processor to respond to the arrival of the latching edge. The delay lines were then removed.

The third type of error, where a data value was entirely missing from the data record, was found to be due to the interrupt latency of the data collection ISR. A data value was missed when the interrupt took longer to complete than the interval between incoming edges. The degree to which this happened was obviously dependent on the rotational speed of the encoder and the execution time of the ISR code. The interrupt latency was measured with a logic analyser. Calculations showed that, as implemented, the system was not capable of correctly recording data faster than 3000 Hz from two input channels. This was 300 rpm measured with a 1000 line encoder. Since the ISR was required to read the data sequentially from each channel, increasing the number of channels would reduce to performance of the system still further. The discovery of this problem and the subsequent examination of its cause lead to the complete redesign of the data capture interface. The conceptual computer architecture completely eliminated the use of interrupts.

The fourth and fifth type of error, where data values were either recorded twice or an intermediate error value was recorded, resulted from the level of electrical noise on the timer counter circuit board itself. The circuit was built on a wire wrapped prototype board and a degree of error due to noise was to be expected due to this implementation. However the actual number of errors has been and remains very high. Figure B-4 shows a negative edge of the encoder signal measured by the logic analyser at the connector to the H3 latching pin of the 68230 interface. This clearly shows the reason for errors of type four, the fall of the voltage level from high to low occurs with a large oscillation which goes back up enough to register a positive level before going negative again. These two negative edges give the double reading. This effect, called ground bounce, is the result of the lack of a sufficient ground level on the board.

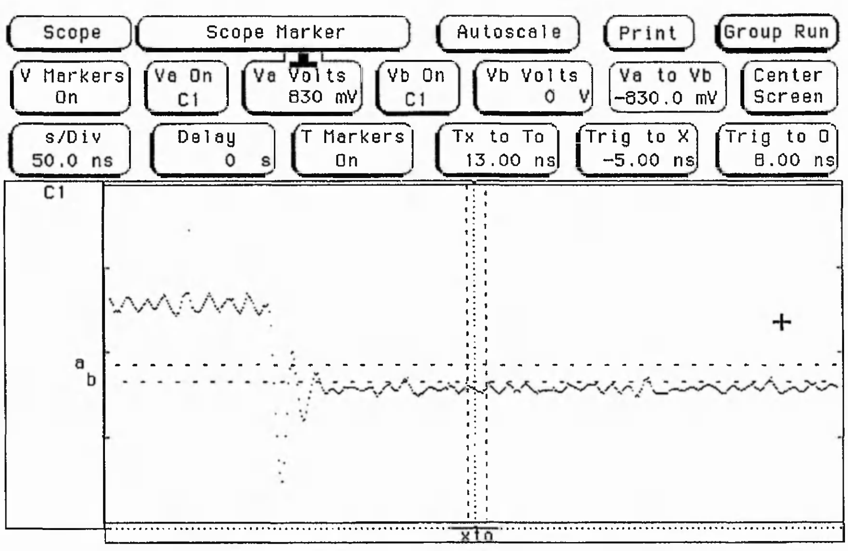


Figure B-4. Ground bounce on the falling edge of the encoder signal.

Looking at the ground level of the signal in figure B-4 it is clear that there is a separate noise signal superimposed on the encoder square wave signal. Figure B-5 is a close up of the 0V level of this signal as measured by the logic analyser. Close examination has shown that this noise signal is composed of various frequencies, due to crosstalk effects from the clock, the power supply to the board, the other read head signals and most significantly, the unused outputs of each read head. Because the components from all of these sources interfere, the noise signal has a variable magnitude. When maximum constructive interference of these noise signals occurs in the ground level of the encoder signal, the peak value of the noise may exceed the threshold for a logic 1 at the 68230's latching pin. This results in the computer being signalled to collect data out of step with the actual encoder signal, leading to an error of the fifth type. In figure B-5 the logic analyser has triggered on an example of such a noise peak.

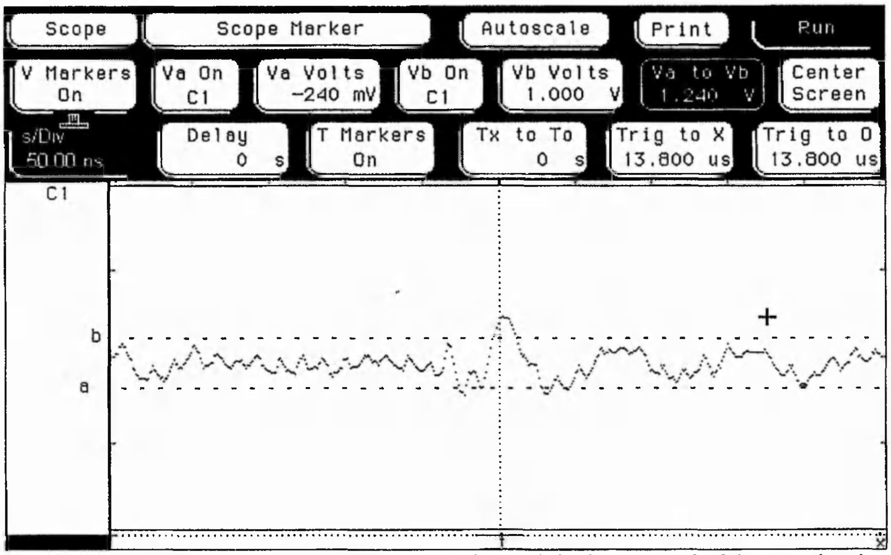


Figure B-5. An example of an error generating noise peak in the 0V level of the encoder signal.

B.3 Timing error analysis of the timer counter circuit

As well as detecting the obvious errors that were leading to gross distortions of the calculated values, it was important to be sure that the circuit was performing as designed and that the values recorded related to the timing interval of the encoder signal. This investigation also allowed the determination of the source of the 0 +/-1 error counter value error. A single encoder read head signal was split between the four inputs of the circuit board and its propagation along the four paths monitored with a logic analyser.

Figure B-6 is a screen shot taken from the logic analyser used to monitor the propagation of the signal through the circuit. Eight probes were used, their signals designated TCOUNT0 to TCOUNT7 in figure B-6. These were placed at various points in the circuit and used to measure a number of timing parameters that were dependent on the path lengths in the circuit. The critical section of the circuit as far as timing and relative propagation is concerned is from the input connector via the flip flops and line drivers to the latching pin, 13, of the latches. For four signals, arriving in step at the

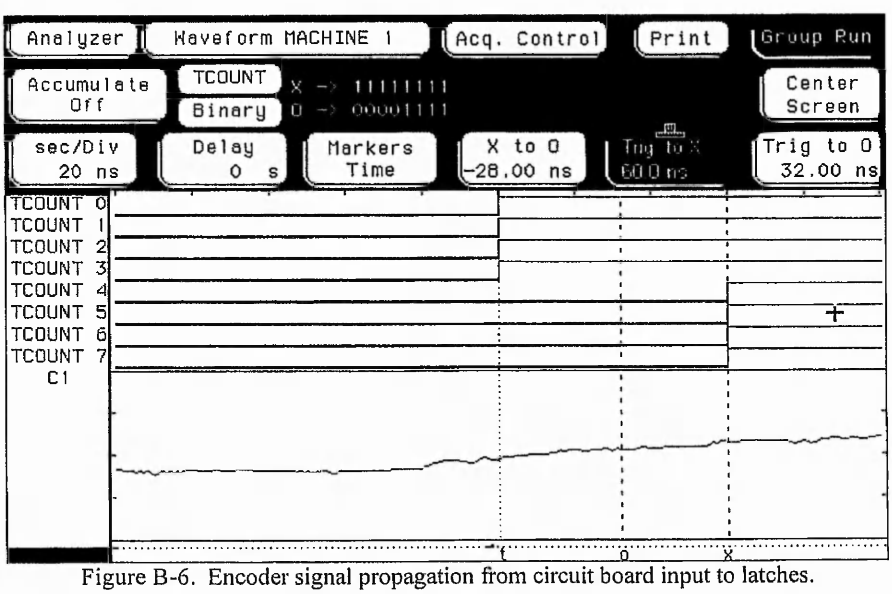


Figure B-6. Encoder signal propagation from circuit board input to latches.

input connector, to latch the same count value, the latching pulses must arrive at pin 13 within one clock cycle of each other.

The ideal situation is shown in figure B-6. T0..3 are the input connector pins, T4..7 the latching pins of the four input channels. The edges are in step in both locations. In this instance the count value recorded on each latch would be exactly the same.

Figure B-7 shows the same signals, but in this case a delay of 4ns is recorded in the arrival of the latching pulse on Ch 2 at pin 13 of the channel 2 latch, T5. Small delays of this magnitude are to be expected due to rise time variations and subsequent delays in triggering. Since the flip-flop is clocked by the same clock as the counter it can happen that the propagation of the signal across the flip-flop is held slightly by the change of the clock, in conjunction with the propagation delay of the clock edge to the various flip-flops. The maximum value of an error due to this source is the length of one clock cycle, assuming in this worst case that the signals are perfectly in step at the input, but

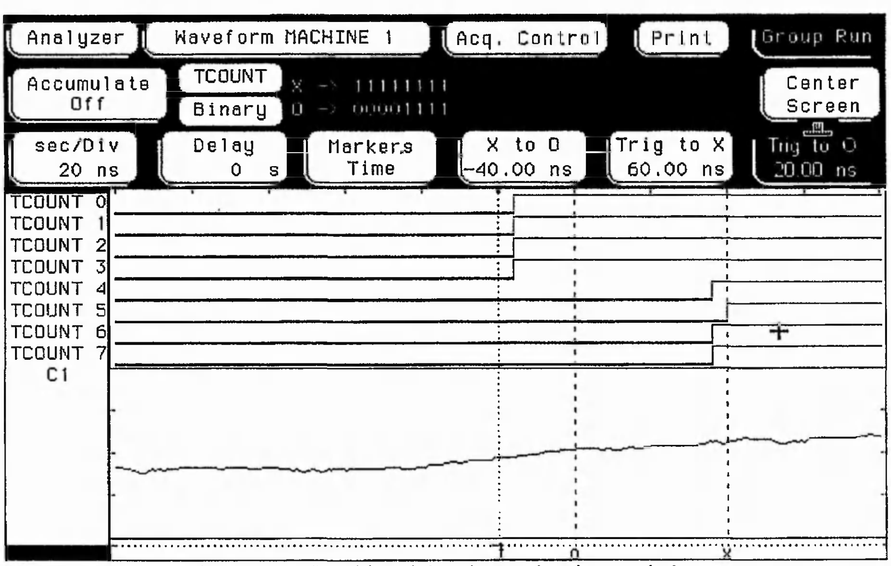


Figure B-7. Latching delay due to rise time variation.

due to a slight skew in propagation reach their respective flip-flops far enough apart that one flip-flop is gated open while the others are gated shut or vice versa. The edges will emerge with a time difference of one clock cycle between them.

An occurrence of this situation is shown in figure B-8. The signals are in step at the inputs, T0..3, but the signal edge on channel 2 reaches the latching pin, T5, ahead of the other three signals. This delay, indicated by the X to O time of 40 ns, is the length of a clock cycle. The value latched on channel 2 from the counter will therefore differ by 1 from the count latched on the other channels. The accuracy of latching is therefore 0 +/- 1 clock cycle, independent of clock speed.

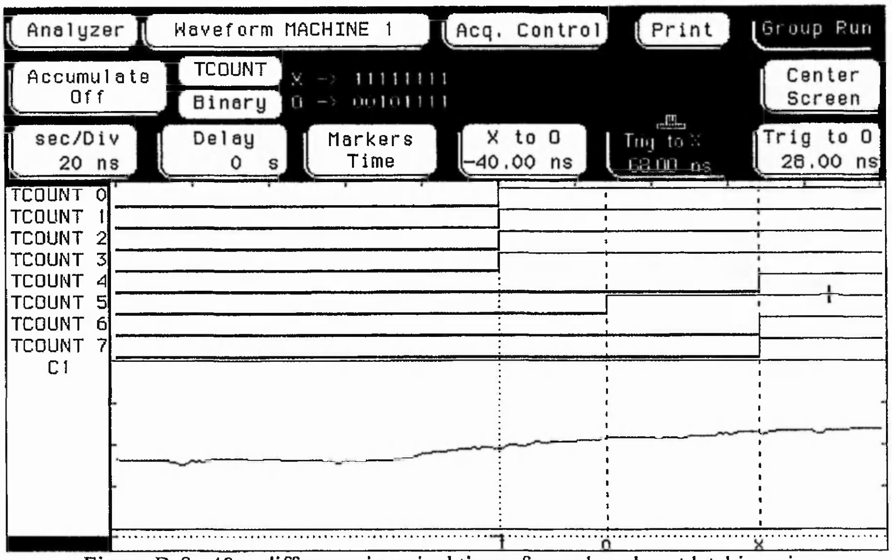


Figure B-8. 40ns difference in arrival time of encoder edge at latching pin.

Clock frequency obviously plays an important part in the relative timings themselves, not just of the signal from the flip-flop to the latch. If the clock frequency were to change by a certain factor then for the same input signal frequency, the measured time intervals on all channels would change by that factor too. In the current circuit the clock frequency, and thus the time between edges due to this error, is only marginally variable. Using the measure facility of a digital oscilloscope the period of the clock was measured as ~41.7 +/- 1 ns. The frequency is thus ~23.8 - 24.1 MHz. The value 23.9 MHz is used in all software calculations.

Signal propagation had to be checked across all the components of the circuit to be certain of the possible errors that might be induced into the data by each part of the circuit. Figure B-9 shows two of the signals being correctly gated across the flip-flop. T0 is the channel 1 IME input on pin 12 of the flip-flop, T1 is the flip-flop output pin 9, T2 is the clocking signal. T3..5 are the same signal for IME channel 2, T6 is The clock output pin, T7 is the clock input to the counter. In this example the two input edges are step at Tx and the output is low. The high level on the inputs is seen on the outputs 8 ns after the positive edge of the clock goes high. This delay represents the propagation time across of the signal across the flip-flop, which is positive edge triggered.

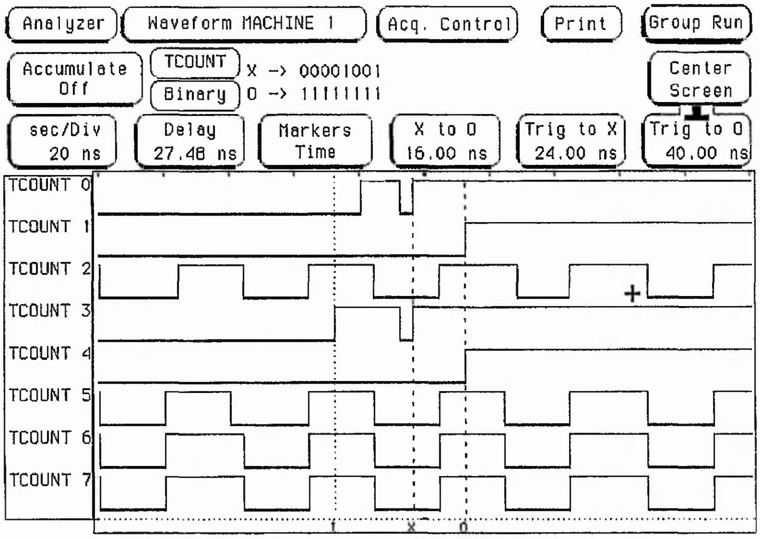


Figure B-9. Flip-flop signal propagation.

Figure B-10 shows the same signals, but in this case the channel 1 IME signal, T0, arrives at the flip-flop at the same instant, to within the resolution of the logic analyser, as the clock signal input to the flip-flop, T2, the output signal, T1 goes high 4 ns later. On channel 2 the signal arrives 4 ns earlier, T3, but this flip-flop does not react as quickly and it takes another clock cycle before the signal passes across the device, T4. This means that a lead of 4ns is transformed into a lag of 44 ns by a flip-flop's slow

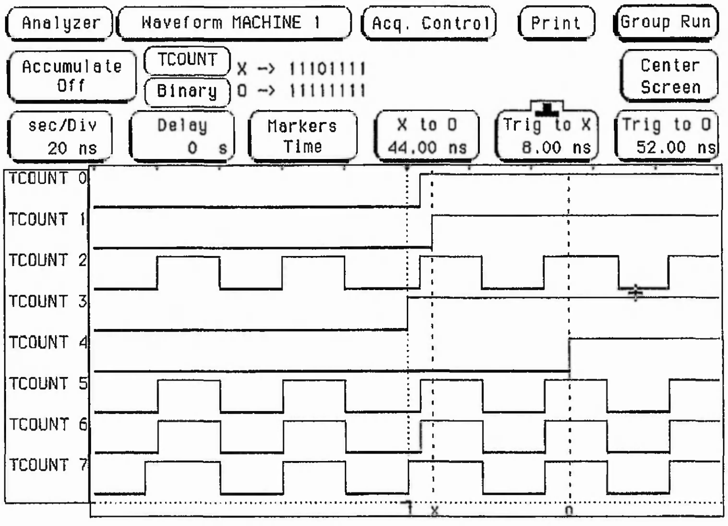


Figure B-10. Error due to propagation delay of the circuit flip-flop.

response to the clock input.

Figure B-11 shows the more normal, predicted problem when two or more OSE signals arrive at their respective flip-flops simultaneously with the clock. The signal on channel 1 goes straight across, the signal on channel 2 is held until the next clock cycle. The delay induced is 40 ns, the clock cycle length.

Both these delays lead to the +/-1 counter value accuracy on the latches.

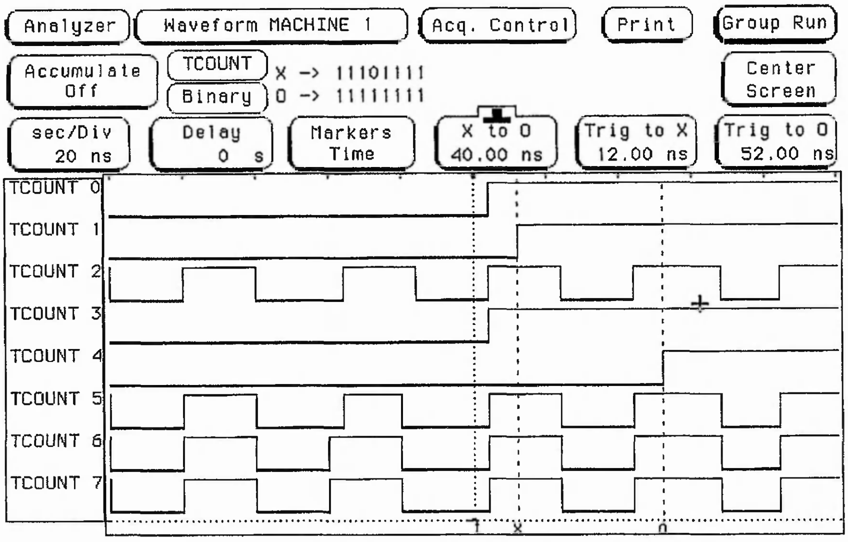


Figure B-11. Second type of error due to propagation across the flip-flop.

Figure B-12 looks at the latches and their relationship to the counter values. T0 is IME channel 1. T1..2, are bits 0..1 on the channel 1 latch outputs. T3..5 are the same signals on channel 2. T6 is the clock input to the counter, T7 is bit 0 of the counter. The Tx to To time measured between the rising edge of bit 0 on the counter and bit 0 of the latch outputs is 12 ns. This is the propagation delay across the latch chips. This obviously imposes a limit on the frequency of the clock that can be used to drive the counter. The maximum clock pulse width must be just over half the propagation delay on the latch since the pulse width of bit 0 of the counter is twice that of the clock. For the devices currently used this would be approximately 83 MHz. Since, as described above, the maximum delay between signals is 40 ns, and the pulse width of counter bit 0 is 80 ns, the error between the two latched values is never greater than 1.

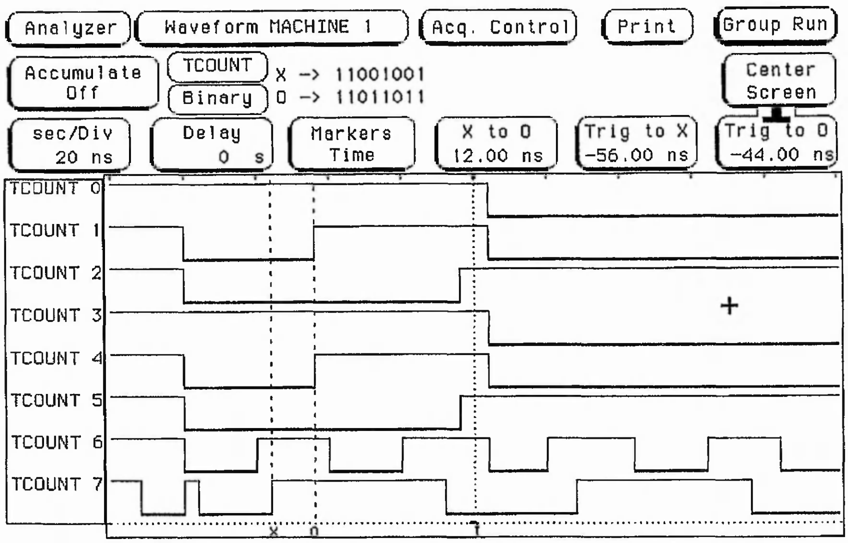


Figure B-12. The relationship between the latches and the counter values.

False edges are seen on counter bit 0 where the level rises and goes high for 4ns. These errors have no effect on the data as they are not of sufficient duration relative to the propagation delay if they occur in the middle of a low signal level. They can cause errors when the false edge extends a counter pulse. Figure B-13 shows how complex it has been observed relationships between waveforms at the counter and latches can become. First there is the 40 ns delay between the channel 1 and 2 inputs, T0 and T3. This results in the count value recorded on channel 1 being 1 less than that on channel 2. Here bit 0 of the latch, T4, has gone high by the time that T3 goes low to latch the value. Now also note that on T7, counter bit 0, there is a 4ns high level error which coincides with the falling edge of bit 2 of the latch outputs. This has the effect of extending the signal on T1 by 4ns. Obviously this is another source of error as in the wrong place it could mean a random one bit error where the bit could in any position. Detection of this type of error is problematic since it will have the same effect as noise when the bit is in the low nibble of the low byte, i.e. the latched error value will be in the same range as the correct value, the previous will be lower and the subsequent higher.

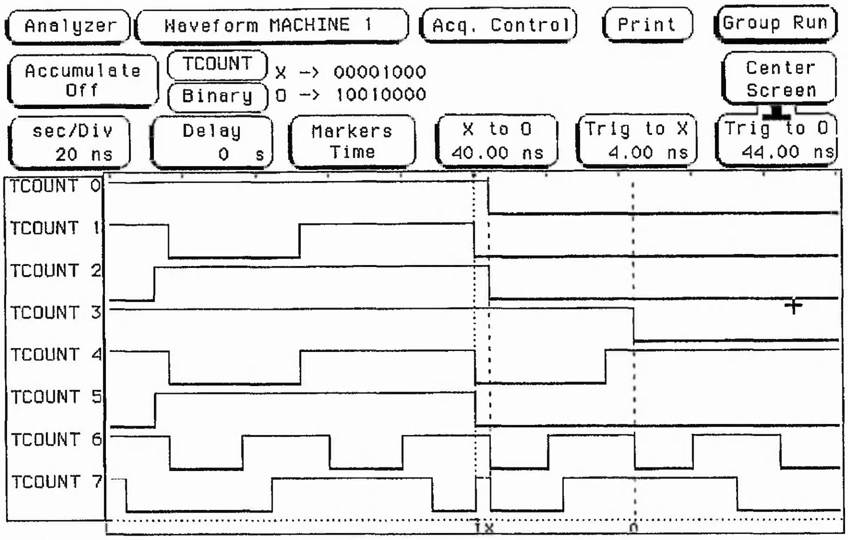


Figure B-13. The relationship between the signals at the counter and the latches.

The error is less likely to effect the higher bits in the latch as they change less frequently, of course if edges did coincide between the latching signal and the extended counter pulse then the value would be massively out of range. This was not observed in the data examined for errors.

Figure B-14 shows the shape of the shape of the signal on bit 0 of the counter. The small peak highlighted is obviously the source of the problem and seems to be part of the general shape of the wave. An increase in the variable magnitude of this peak above the logic 1 threshold causes the error. Figure B-15 shows the signal propagation across the channel 1 latch. The latch outputs are the odd T numbers, and inputs are the even T numbers. T0 is the clock signal, T15 the latching pulse. The error pulse was not seen. Propagation time across the latch in this case was 16 ns. Looking at the latches of the other channels gave a similar picture. The propagation time observed was variable between 8 & 16 ns. This has implications for the data collection method. If the data collection computer or device reacts to the encoder pulse faster than the propagation time to and across the latch, the collected data will be indeterminate as the latch will not have become stable. This was the case with the first data capture interface design and gave rise to the second type of error described in the previous section.

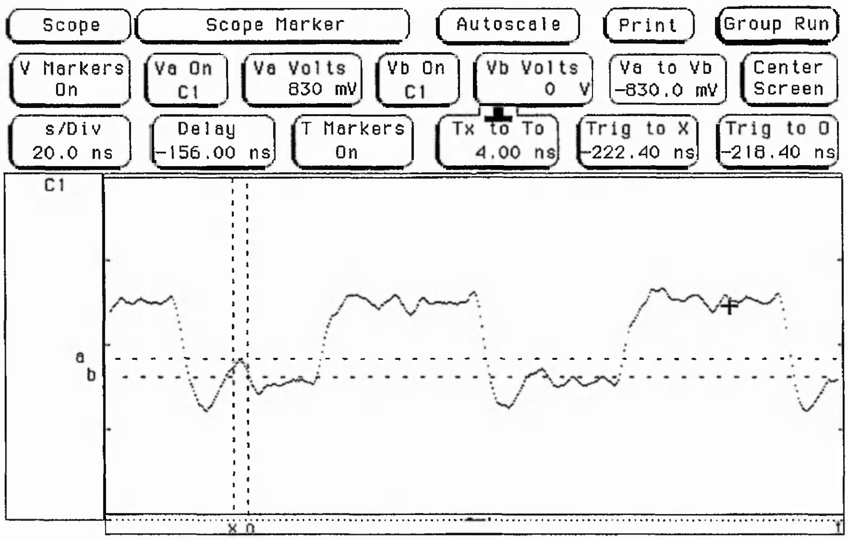


Figure B-14. The shape of the signal on bit 0 of the counter

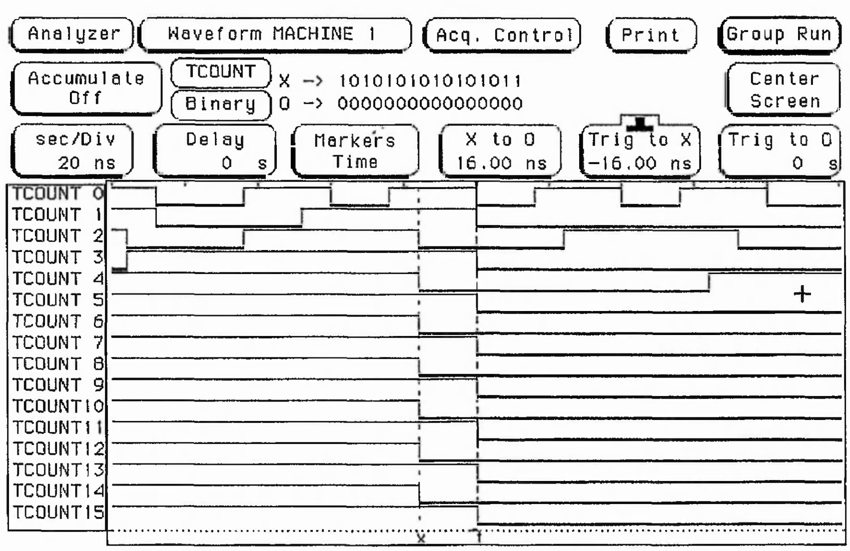


Figure B-15. Signal propagation across the channel 1 latch.

The complete 16-bit latch for each channel is made up of four discrete 4-bit devices. The propagation time of the encoder signal to each of these components is another possible source of error. The signal should reach all the latches simultaneously, particularly as there are two latch pins on each of the four devices in each latch. Figure B-16 shows this signal and the relationship of its arrival on each latch pin. T0 is pin 1 of the low nibble, T1 pin 2, T2 pin 1 of the next nibble and so on. The reason for the 4 ns delay between T0, T7 and the other signals can be seen in figure B-17. The falling edge which triggers the latch takes 13ns to fall. The various latches in the worst case will each choose a separate point on this edge to trigger. The error due to this will not be more than one counter cycle.

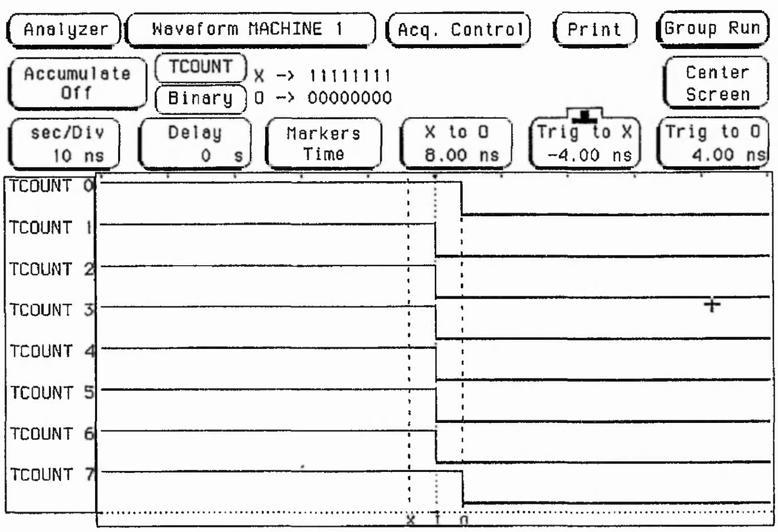


Figure B-16. Latching signal propagation across latch enable pins

At the end of the development of the interface circuit the whole of the signal path had been investigated with the logic analyser. The sources of all the errors seen in the data record had found and accounted for. The +/- 1 clock cycle error which remained had also been determined as part of the basic mechanism of the circuit related to the various component's propagation delays.

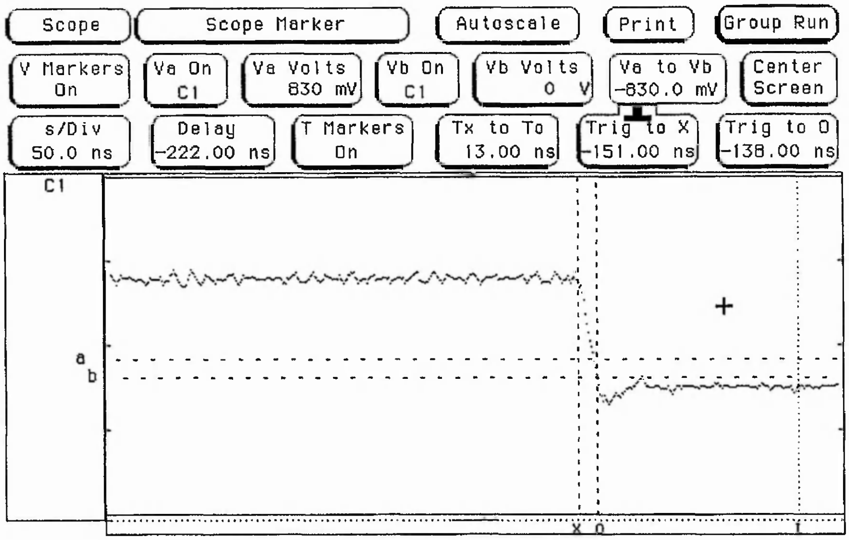


Figure B-17. The 13ns taken for negative edge of the encoder signal to fall to zero

Appendix C

Data capture interface circuit schematics

C. Data capture interface circuit schematics

This appendix contains a block diagram and circuit schematics for the data capture interface circuit board.

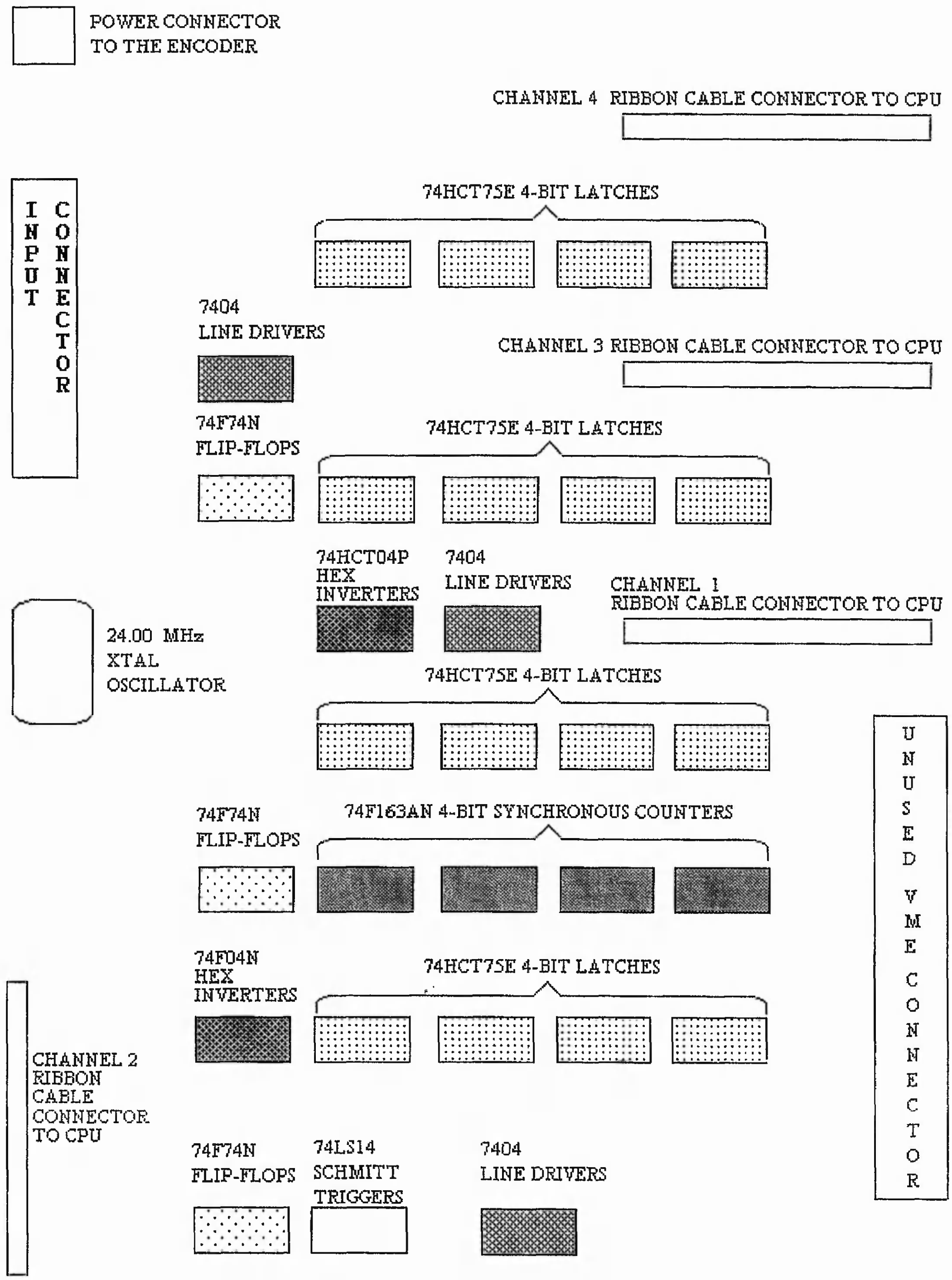
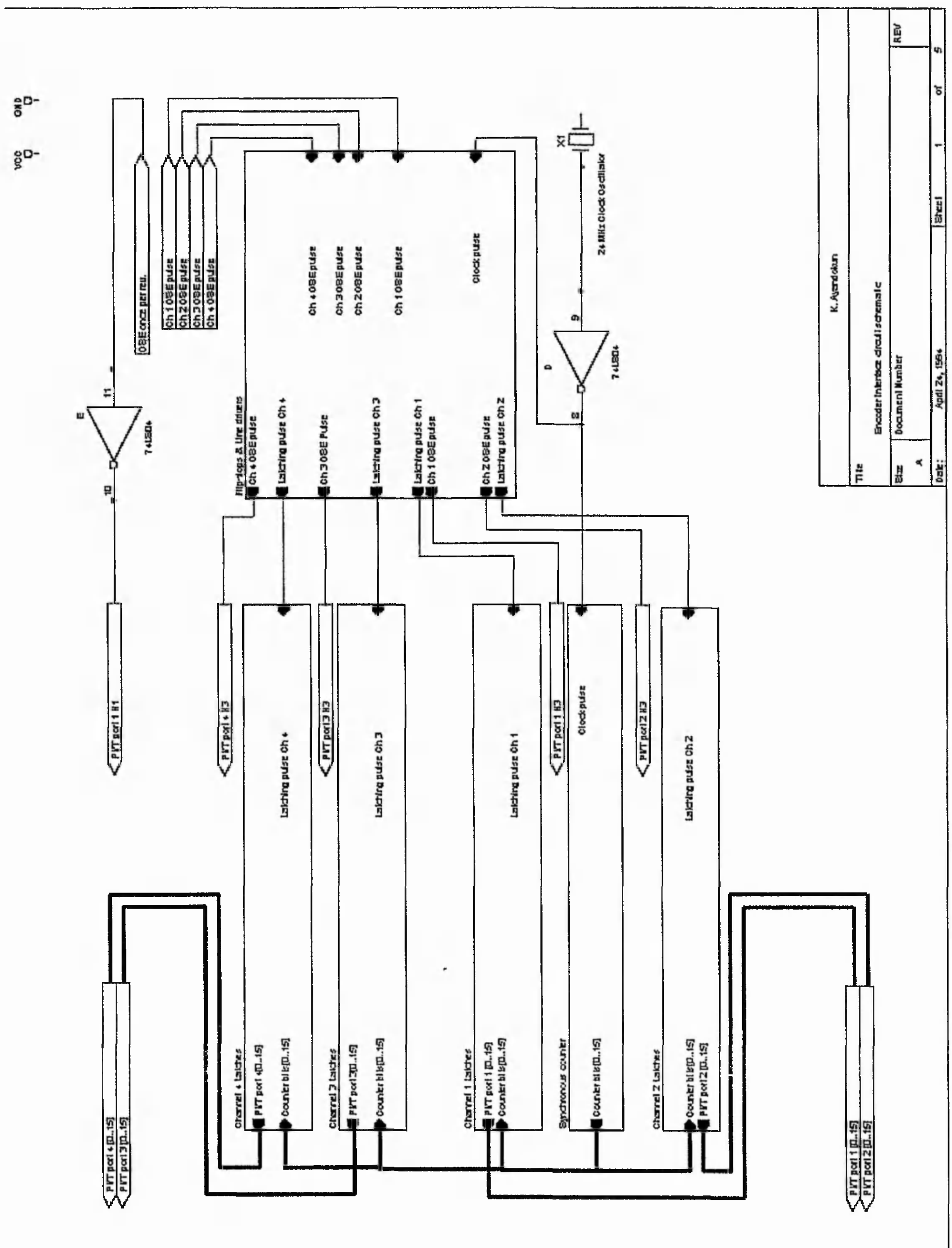
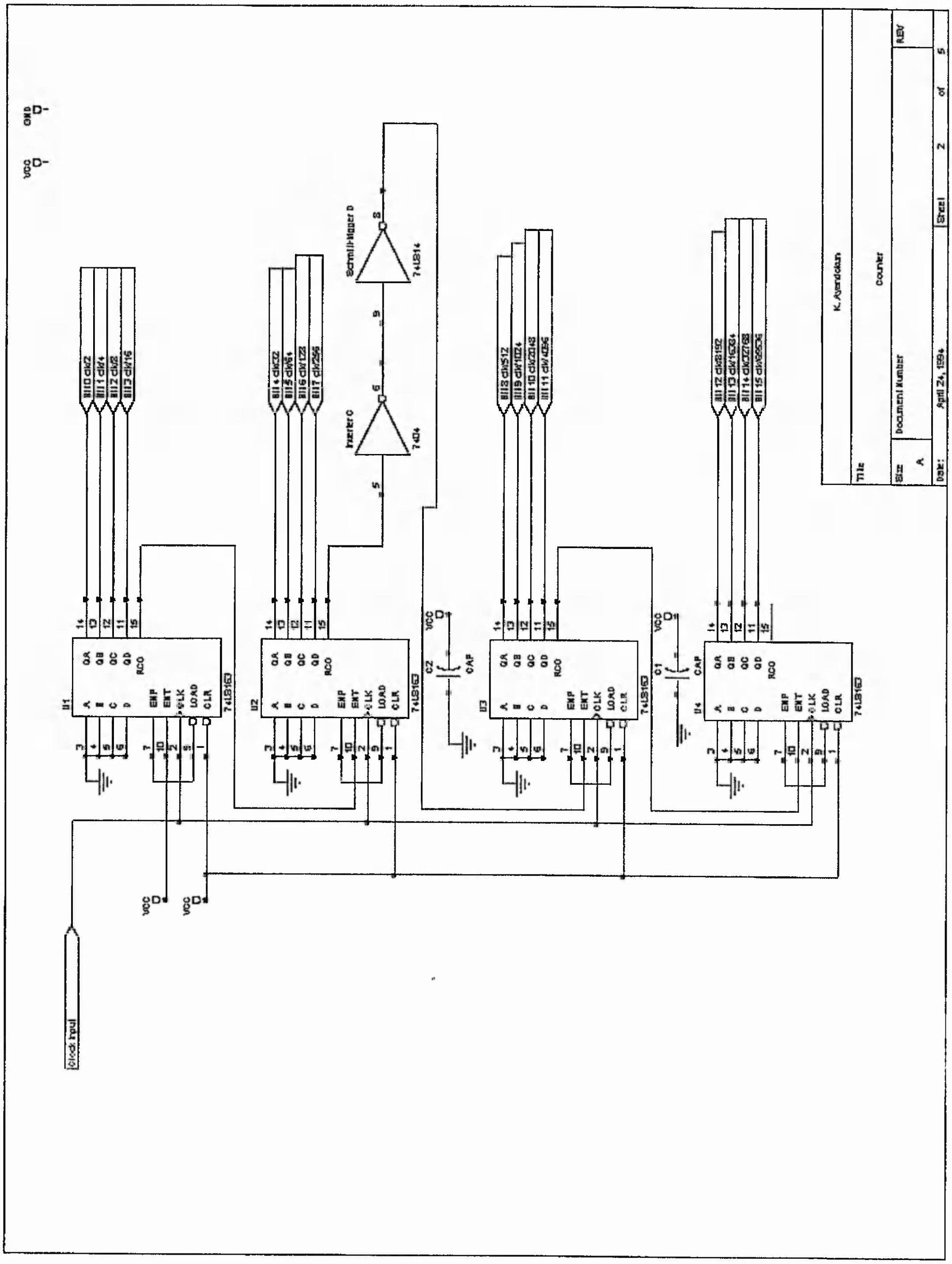


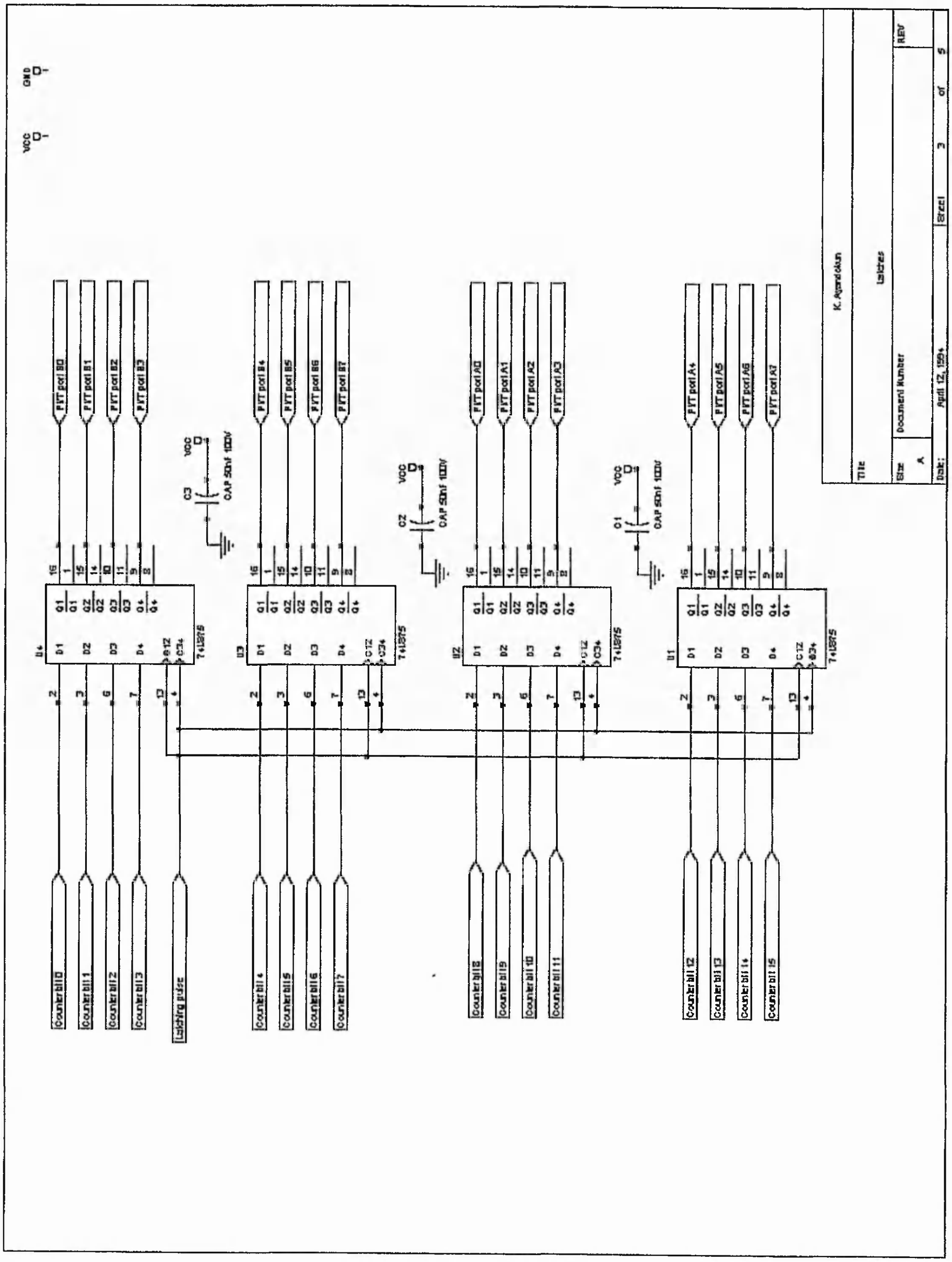
Figure C-1 Interface circuit board bloc diagram

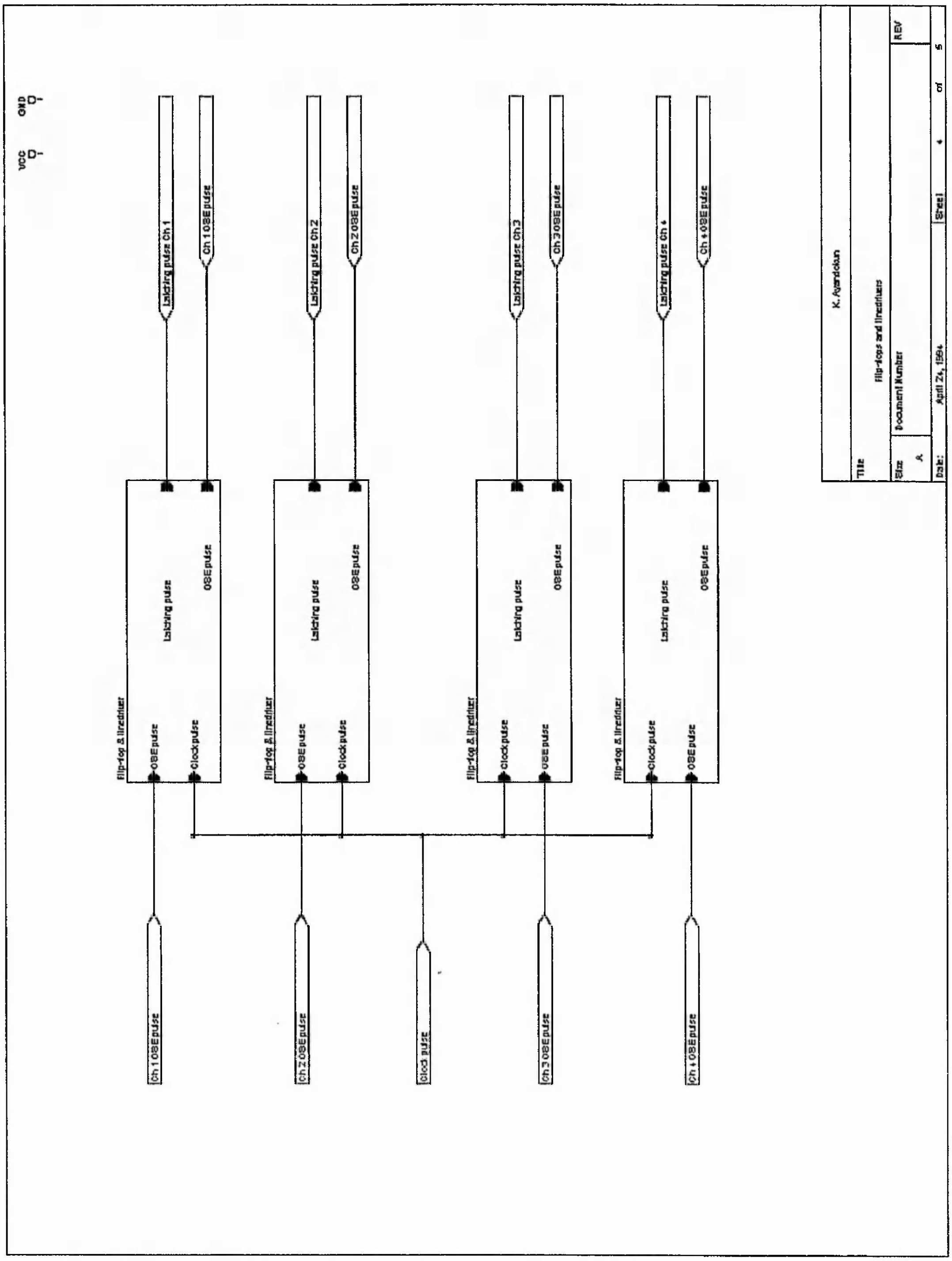


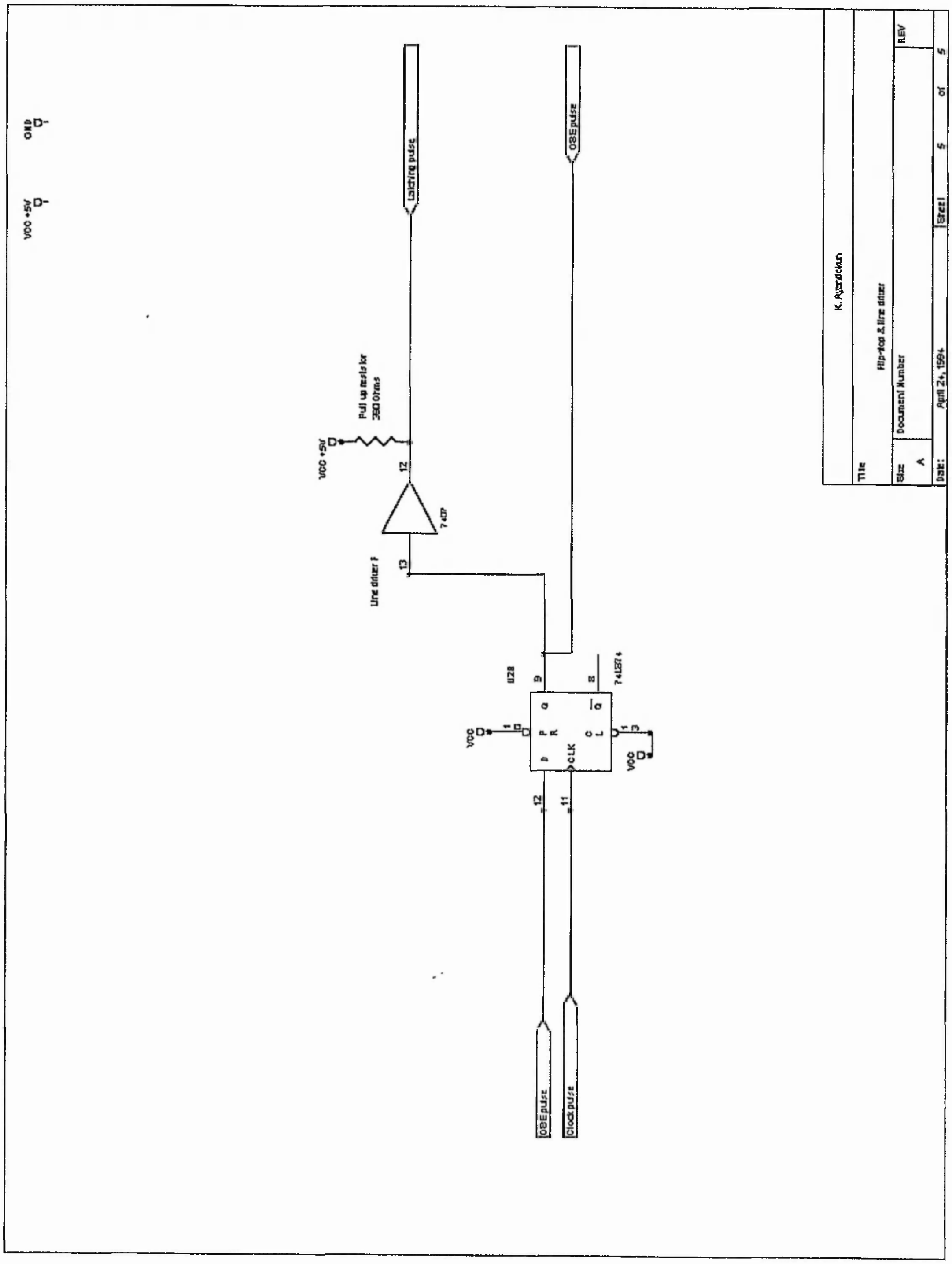
VCC
GND
OBE once per rev.
Ch 1 OBE pulse
Ch 2 OBE pulse
Ch 3 OBE pulse
Ch 4 OBE pulse
74LS04
24 MHz Clock Oscillator
X1
Flip-flops & Line drivers
Clock pulse
Synchronous counter
Counter bits[0..15]
Channel 1 Latches
Channel 2 Latches
Channel 3 Latches
Channel 4 Latches
Latching pulse Ch 1
Latching pulse Ch 2
Latching pulse Ch 3
Latching pulse Ch 4
PVT port 1 [0..15]
PVT port 2 [0..15]
PVT port 3 [0..15]
PVT port 4 [0..15]
Title
Encoder interface circuit schematic
Size
A
Document Number
REV
Date:
Sheet 1 of 5



Clock Input
74LS163
Inverter C
7404
Schmitt-trigger D
74LS14
Title
Counter
Document Number
Sheet 2 of 5
April 24, 1994




VCC
GND
Latching pulse Ch 1
Ch 1 OBE pulse
Latching pulse Ch 2
Ch 2 OBE pulse
Latching pulse Ch 3
Ch 3 OBE pulse
Latching pulse Ch 4
Ch 4 OBE pulse
Flip-flop & linedriver
OBE pulse
Clockpulse
Latching pulse
OBE pulse
Ch 1 OBE pulse
Ch 2 OBE pulse
Clock pulse
Ch 3 OBE pulse
Ch 4 OBE pulse
K. Ayandokun
Title
Flip-flops and linedrivers
Size
A
Document Number
REV
Date: April 24, 1994
Sheet 4 of 5


VCC +5V
GND
Latching pulse
OBE pulse
Pull up resistor
390 Ohms
VCC +5V
7407
Line driver F
U28
74LS74
CLK
VCC
VCC
OBE pulse
Clock pulse
Title
Flip-flop & line driver
Size
A
Document Number
REV
Date:
Sheet
5
of
5

Appendix D

A fibre optic rotary motion sensor for the real time condition monitoring of rotating machinery

P. A. Orton, K. Ayandokun, N. Sherkat and P. D. Thomas, in *Fibre Optic Physical Sensors in Manufacturing and Transportation*, John W. Berthold, Richard O. Claus, Michael A. Marcus, Robert S. Rogowski, Editors, Proc. SPIE 2072, pp. 195-207, 1994.

A fibre optic rotary motion sensor for the real-time condition monitoring of rotating machinery

P. A. Orton, K. Ayandokun, N. Sherkat & P. D. Thomas

The Nottingham Trent University, Department of Computing, Burton Street, Nottingham NG1 4BU, United Kingdom,
Tel: +44 (0)602 418418 ext. 2944, Fax: +44 (0)602 486518, Email: kaa@doc.nottingham-trent.ac.uk

ABSTRACT

This paper describes the operating principle and construction of a new type of rotary motion sensor. This novel optical device is based on a patented evolution of the high precision optical shaft encoder. It integrates highly accurate measurement of the angular position and two dimensional centre position of a rotating shaft. The current work described is aimed specifically at using this sensor for monitoring the condition of rolling element bearings. The process by which the output from four stationary, non-contact, fibre optic sensing heads can scan a coded disk coupled to a shaft and derive its centre position to ~30 nanometres at an angular resolution of ~0.0005 min is explained. The design and implementation of the computer architecture to which the sensor is interfaced via noise immune fibre optic links and the algorithms which underlie the processing software are examined. Initial results obtained from development tests carried out with the sensor far are presented. The integration of three dimensions of motion in a single instrument has yielded high resolution measurements of shaft centre movements. These movements are directly related to the condition of the bearing supporting the shaft and so should offer a much clearer picture for condition monitoring than measurements made conventionally by sensors on the machine casing.

1. INTRODUCTION

In modern, highly automated, process industries predictive maintenance has become the prime method of avoiding costly unscheduled production stoppages due to machine breakdown. The basis of a successful predictive maintenance strategy is the condition monitoring of equipment through the selection, measurement, and trending of a critical machine parameter[1]. Monitoring of this parameter allows detection of the gradual deterioration of key components without the physical examination of the individual component. This allows optimal employment of maintenance resources by ensuring that only necessary repairs and replacements are conducted. Downtime of equipment is thus reduced and maximum economic use is gained from components designed with a certain wear lifetime.

There are many parameters which can be measured, dependent on the type of machine and component to be monitored. For rotating machines, from large turbo-generators to small scale machine tools the key components are the rolling element bearings which support shafts and spindles. These bearings consist of an inner race fixed to the shaft, an outer race fixed into the machine housing and a number of rolling elements in the space between the two races. Bearing failure is due to surface defects which appear on one or more of these components due to wear. All bearings excite their supporting structures at certain natural frequencies related to shaft speed and the physical nature of the bearings. Defects on the bearing elements change these frequencies. The parameter by which the condition of these bearings is most commonly monitored is therefore vibration[2]. Vibration can reveal component imbalance, misalignment and bent shafts as well as bearing defects. The term vibration, used in this context, covers the three parameters displacement, velocity and acceleration.

Wear debris analysis is also used to monitor bearing conditions. In this method it is the analysis of the size, shape and number of particles in the bearing lubricating oil that provide an indication of the condition of the rolling contact surfaces of the bearings. Due to the need for oil to be removed for laboratory analysis, collection of data for this method cannot be automated or considered as part of a real-time approach.

Conventionally measurement of the selected vibration parameter is by means of analogue electro-mechanical sensors placed on the machine casing. The choice between acceleration or velocity sensors is made based upon the calculated failure frequencies for the bearings to be monitored. Each parameter works best over a given frequency range[3]. The raw vibration signal from a single point on a machine is not sufficient information on it own. A vibration needs to be measured at several carefully selected points and in several directions so that the combined signals can be analysed into the basic components of the complex raw waveform. If the raw signal does not contain the frequency components that are representative of a fault condition no amount of signal processing or analysis will reveal that condition. Therefore the location and direction of measurement are critical for acceleration and velocity sensors. The signals measured remotely on the machine casing are also subject to attenuation and distortion. This is particularly the case where the machine speed, and hence the vibrational energy due to any bearing defect, is low.

Displacement sensors produce a signal that is proportional to the displacement of a shaft relative to its mountings. It is a realization of a new type of this form of sensor that this paper describes. Proximity measurement of shaft displacement should be preferable to remote velocity or acceleration sensing. Shaft displacement increases with bearing deterioration but vibration velocities measured on the bearing housing may not always be sensitive to this increase. Unfortunately, conventional displacement spectra can miss a great deal of bearing health information, suppressing much of the spectral content relevant to defects at low frequencies. Although there are existing sensors based on capacitive, magnetic and reluctance principles, the type most commonly found in use, the proximity probe, is based on the eddy current principle. The probe sets up an eddy current between itself and the shaft, as the shaft moves relative to the probe, the eddy current changes modulating the output voltage. Fibre optic probes have also been developed in which laser light is shone from the tip of the probe and the gap is measured by the amount of reflected light picked up. Used in this way the proximity probe is only capable of measuring gap, nothing else. Methods of using proximity probes for more complex bearing condition monitoring have been described[4,5] in which the elastic deformation of the outer race is measured and used to detect bearing element flaws.

Our new device, the incremental motion encoder (IME), extends the use of displacement measurement as a condition monitoring technique. It allows displacement to be related to angular position to an unparalleled degree and increases substantially the useful bandwidth over which measurements can be taken. It is also a 100% digital device with no analogue to digital conversion step. This allows noise immune optical fibres to carry its output signal directly from the point of measurement.

2. THE HIGH PRECISION OPTICAL SHAFT ENCODER PRINCIPLE

The incremental rotary optical shaft encoder (OSE) is a device commonly found in industry and is used for the measurement of angular position and angular velocity[6]. In its simplest form, illustrated in figure 1., it consists of a radial optical grating disk coupled to the shaft and a single read head mounted on the shaft housing.

The read head consists of a light source, usually an LED, on one side of the disk projecting a beam of light onto a set of photo-electric cells on the other side. As the shaft turns, the optical grating on the disk interrupts the beam and the output of the photo-cells comprises a train of pulses corresponding to the passage of spaces in the grating past the beam, PC_1 & PC_2. Commonly two photo-cells are arranged in the read head so that their output signals have a 90° phase difference as shown. This phase difference indicates the direction of shaft rotation. A third signal, PC_3 generated once for each rotation of the disk, is used as a position reference. These signals are shaped by Schmitt triggers

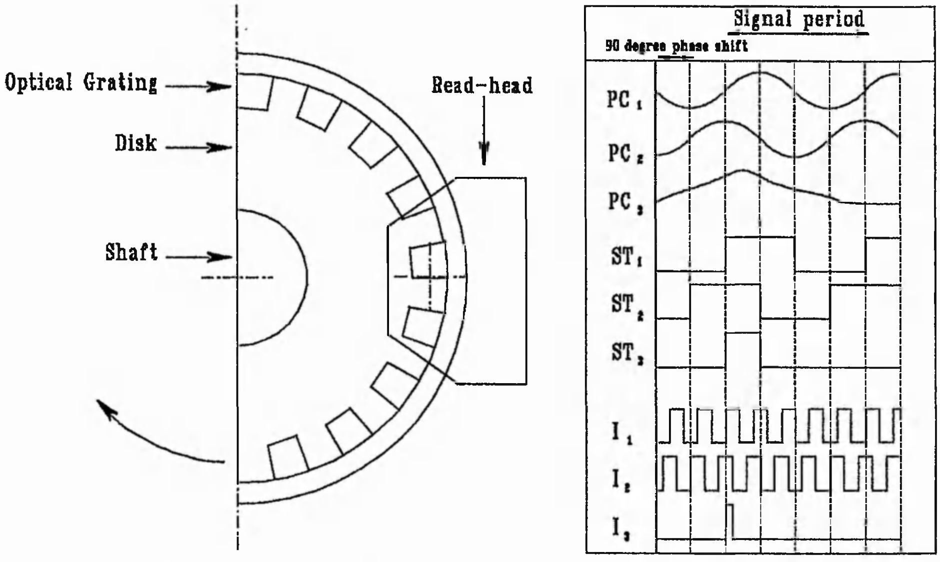


Figure 1. Optical shaft encoder output signal interpolation.

incorporated in the read head to give the square wave signals $ST_{1,2,3}$ which form the output signal of a simple encoder. Angular displacement and velocity are then simply determined by the counting of these pulses.

For high precision angular measurement applications however the measuring step given by the actual grating line spacing is insufficient. High precision optical shaft encoders are therefore based on the principle of interpolation. As it is not technically feasible to use an optical grating with the line spacing necessary for resolutions of 0.001° or better, interpolation is used. This may be an analogue or digital process depending on the accuracy required. Using vectorial addition resistor networks within a read head can produce additional phase shifted signals from the photo-cell outputs. In figure 1. the signals $I_{1,2,3}$ represent the result of 18° phase shifted signals producing a 5x interpolation of the original output. The output signals have five times the frequency of the input signals and twenty edges are produced for each grating spacing. Digital interpolation allows for the highest possible subdivision of the grating spacing. The analogue signals from the read head are converted to digital voltage values, an external or embedded microprocessor then acts as an arc-tangent calculator which uses the result of the quotient PC_1/PC_2 to pick the matching interpolation value from a look-up table. Interpolation of 1024x on a 36000 line grating allows a measuring step of 0.00001° is achieved in this way.

Due to their degree of angular accuracy, high precision encoders are particularly susceptible to shaft axial eccentricity and angular vibration as a source of error. Precise couplings between encoder and shaft are therefore used to limit these errors, but precise mechanical transfer of the motion of the shaft to the encoder without error is not possible. Within the conventional high precision encoder the use of a second read head allows compensation for these errors. Encoders with digital interpolation have four read heads which together allow for a positional correction prior to interpolation. As will be explained the IME uses four read heads in an analogous manner to provide the angular data used to resolve shaft centre position.

3. THE INCREMENTAL MOTION ENCODER PRINCIPLE

Recent work with precision encoders for the control of grinding machines[7] produced the idea of a sensor, the IME, which would use shaft centre position measurements as a vibration monitor. A method has also been developed for producing the high degree of angular accuracy required from a optical grating of relatively low resolution.

3.1 Time base interpolation

The basis of the high angular accuracy of the IME is the use of time base interpolation. Rather than count the edges of a encoder output wave in the time duration of a sample period, figure 2a., the IME uses a precision clock oscillator to generate a time base which is sampled at the equal intervals of angle represented by the period of the encoder measuring step, figure 2b.

Using this method the size of the smallest interpolation step is limited only by the frequency of the time base oscillator, taking the ability to count high frequency edges as a limiting factor. This is however a dynamic interpolation method, the number of interpolation steps N is given by:

$$N = \frac{F_{tb}}{rps \times G} \qquad (1)$$

F_{tb} = frequency of the time base.
rps = rotations per second of the encoder disk.
G = number of grating lines on the encoder disk.

The above equation shows that despite its theoretical accuracy the IME cannot measure point to

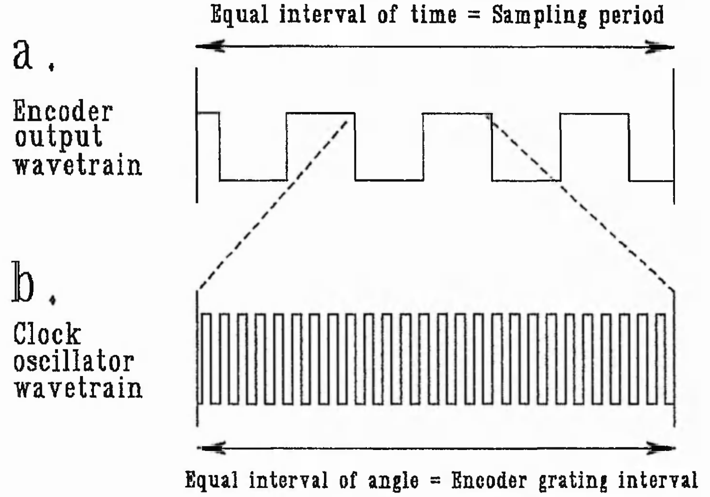


Figure 2. Time base interpolation.

point angular rotation in the manner of a conventional precision encoder with any more than the actual resolution of its disk. With the relatively high speed rotating machines this device is aimed at monitoring, this accuracy is generally sufficient. The high degree of angular accuracy given by the sampled time base value is only available to the system internally to derive the position of the shaft center.

The use of time base interpolation results in the ability of the IME to produce highly accurate measurements of angular velocity precisely related to angular position. Conventionally the angular displacement measured within one sample period will be accurate to +/- one encoder measuring step and the velocity calculated will not bear a relationship to angular position. By using the angular distance of the measuring step, averaged to account for interslot errors, and the time base value recorded over this angular distance, the IME is capable of resolving changes in the angular velocity which occur between measuring steps.

3.2 The IME interface circuit

To produce the time base interpolated angular measurements which can be processed by a computer system a special interface circuit is required to form the link between the computer system and the IME outputs. This circuit is simple in construction and is shown schematically in figure 3. below. A precision clock oscillator (1) drives a free running synchronous counter (2). Four identical latches (3) are joined to the counter by a bus (4). Their activating pins are connected to the square wave outputs of each encoder read head, their output pins are connected to the computer system. The falling edge of the incoming square wave from each of the four encoder read heads latches a count value onto its respective latch representing the exact time it arrived at the latch. The computer system is signaled that data is waiting on the latch by the same falling edge. It can then collect the latched count for storage in a memory table. The difference between successive values gives the time base interpolation count. The current experimental circuit is constructed using a counter and latches which are 16-bits wide and a 24 MHz oscillator.

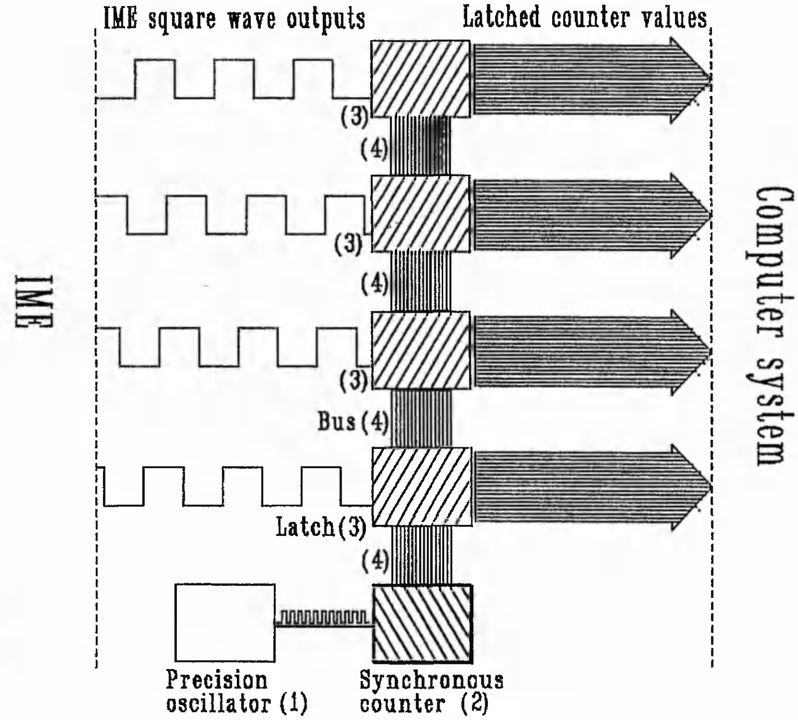


Figure 3. IME interface circuit.

3.3 Method of shaft centre position calculation

Figure 4. below illustrates the way in which the IME calculates the shaft centre position. The device is constructed with four read heads spaced at 90° intervals around the circumference of an optical grating disk, marked A - D. Each diametrically opposed pair of read heads is used to calculate the position of the shaft centre along the perpendicular bisecting the diameter joining them. Consider the read head pair A and B, distance l microns apart. If the centre of the shaft X is in line with the two read heads at the point X' then the difference between the angular position measurement derived at A and that at B, θ, will be π radians. Assume then that X is displaced to the left as shown in the figure 4. In this case the angle θ will no longer be π. The shaft centre displacement from X' to X, D is given by:

$$D = \frac{(\pi - \theta)\, l}{4} \tag{2}$$

This is an approximation which can be made for centre displacement of the order of tens of microns if the disc is of the order of tens of millimeters in diameter.

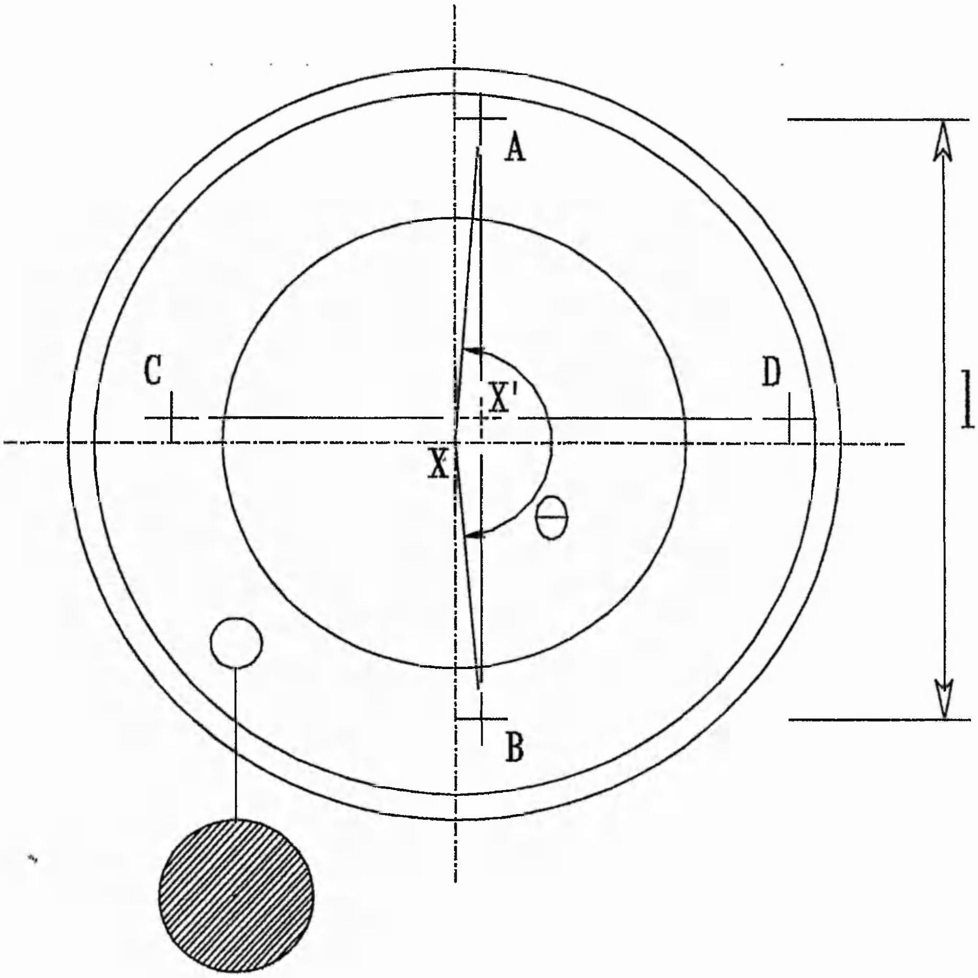


Figure 4. Shaft centre position calculation.

Each pair of diametrically opposed read heads provides the position of the shaft centre in a single dimension. Together the angular measurements from four as shown allow the shaft centre to be located in two dimensions. The error in this calculation due to the radial movement of the sensors can be eliminated by an iterative method once the initial estimate for the displacement of the encoder disc centre has been found. A detailed analysis of the trigonometry shows that three read heads spaced at 120° would be sufficient to allow the disk centre position to be calculated. This method would however require far more computation as a complex sine cosine relationship would then need to be evaluated . Our industrial collaboration with Muirhead Vactric Components Limited, a manufacturer of encoders, has also revealed that the production of accurately placed diametrically opposed read heads is significantly easier. This allows conventional techniques to be used for the production of the read head optics.

3.4 Relative angular position calculation

As described, the computer system receives from each output of the IME interface circuit a series of counter values, one for each line on the encoder grating disk. From these the angular difference between read heads is calculated. The start of data collection from all four reads heads is synchronised by the once per revolution signal of the encoder disk taken from the output of one read head. From this point the successive counter values are stored in a four column table in memory. Each row consists of two pairs of values from the diametrically opposed read heads. The row number indicates the angular position, at intervals of the encoder grating line spacing, from which the readings were taken. At any one of these positions the angular difference between the pairs of read heads should be π radians if the disk is centred. The diameter joining the read heads will be a grating line position which will have latched the two counter values in the table simultaneously at both corresponding interface circuit latches. In this case the numerical difference between pairs of values in each row will be zero. When the disk is off centre the angular difference between the read head pairs will be θ. Counter values at the same angular position, indicated by their row number, in the table will now be different. The grating line passes each read head at a different point in time and therefore latches different counter values. This difference in time is used to find the precise angular difference between the read heads by means of time base interpolation.

Figure 5. shows a graph of angular position against time for read heads A and B with the format of the memory table inset. In this example A is shown to be leading B as would be the case from figure 4.

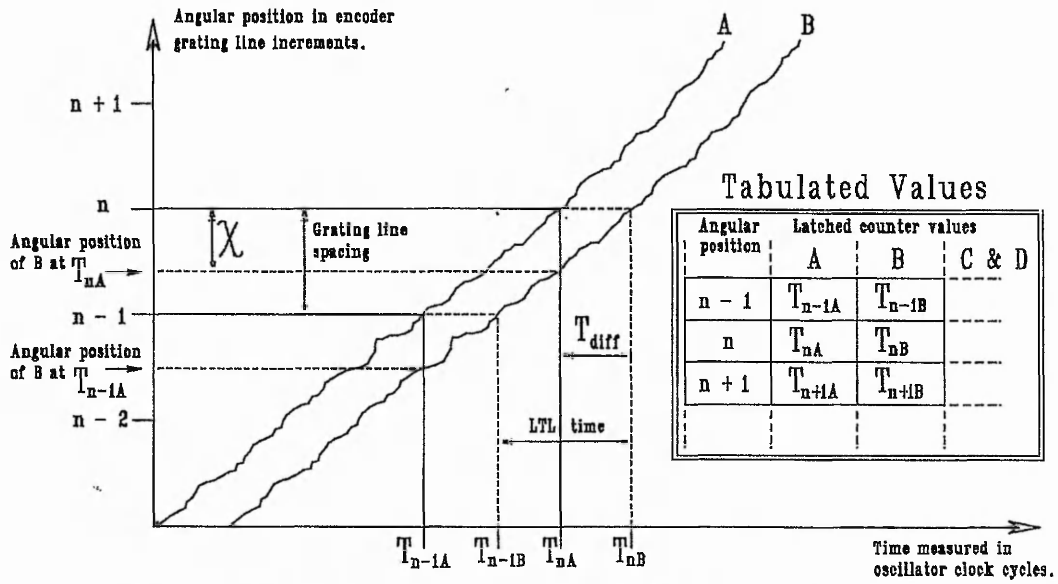


Figure 5. Calculation of relative angular position by time base interpolation.

At angular position **n** time value T_{nA} is less than T_{nB}. To find χ, the difference in angular position at this point, the angular position measured by B at T_{nA} must be found by interpolating between the values **n** and **n - 1** measured by A. Assuming a constant angular velocity over the distance **n - 1** to **n**, the difference in time measured between T_{nA} and T_{nB}, T_{diff}, will bear the same percentage relationship to the line to line time $T_{nB} - T_{n-1B}$ as χ bears to the encoder grating line spacing, **n - n - 1**. This allows χ to be found. θ is then given by the straightforward relationship:

$$\theta = \pi - \chi \qquad (3)$$

The above algorithm, which is used in the current processing software, is a simple interpolation between two points. To increase the margin of accuracy interpolation could be performed between a greater number of points using more complex methods.

4. COMPUTER ARCHITECTURE

4.1 Conceptual architecture

The computer architecture required to support the IME is functionally divided between data collection and data processing, analysis and display. As shown schematically in figure 6. below each output channel of the IME interface circuit is independently interfaced to the computer system which performs the measurement calculations by a data collection unit (DCU).

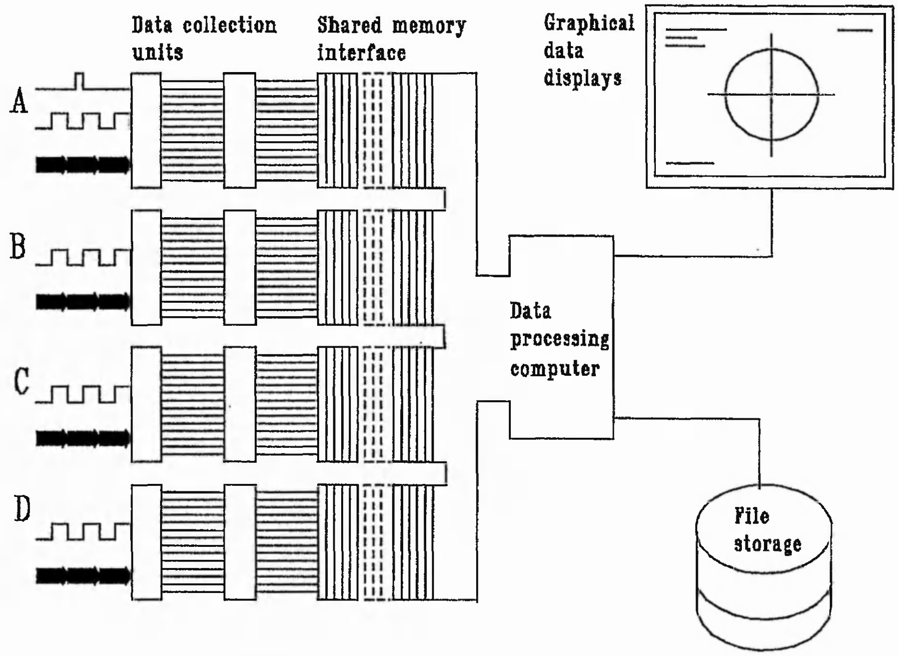


Figure 6. Conceptual computer architecture.

Figure 7. shows the functional components which make up a DCU. A parallel interface (PI) is used to link the latches of the IME interface circuit to the data bus of the DCU. The count values are read from the PI into the dual port FIFO memory of the DCU under the control of the memory management unit (MMU) responding to the falling edges of the IME output. The processing computer system sees the counter values across the shared memory interface as the four column table described above. The start of data collection is synchronized by the once per revolution signal from one read head. Use of this signal allows the data collected to be precisely related to shaft angular position externally. After it has been set by the computer system the DCU which receives this input signals the others via the shared memory interface. The number of counter values required is written to all the MMUs and they fill the data table accordingly. Data collection then stops and the values in the

table are then processed to give values for displacement and angular velocity. The user has control of the collection and display of this data via a graphical user interface (GUI). The system can gather data continuously, averaging the resulting values for display or writing them to disk for historical analysis.

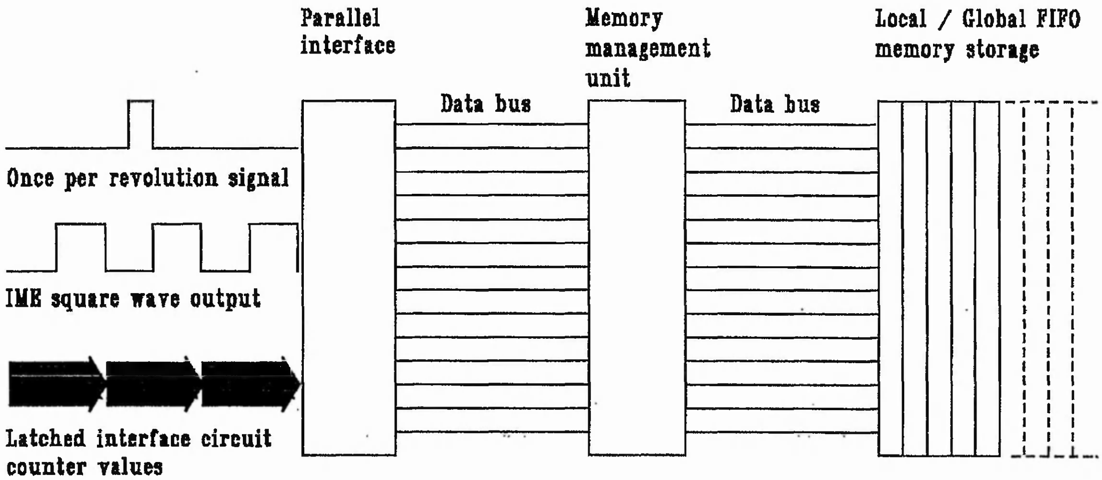


Figure 7. Data collection unit.

4.2 Implementation of experimental architecture

The conceptual architecture and operation outlined above are capable of being implemented in various ways with different physical components. The system, IME interface and processing computer, could be produced entirely as an ASIC device. Embedded with the encoder this would produce a compact sensor. This is a long term production objective of our industrial collaborators. At the initial experimental stage we are using discrete proprietary components. The function of a DCU is essentially that of a DMA controller and a system could be composed around such components and simple dual port FIFO memory chips if it were to be built from scratch. We have implemented the DCUs with four monoboard MC68000 computers joined on a VME bus to global memory. Each board has a parallel interface and local memory with the processor acting as the MMU. The IME interface circuit is made up from FAST TTL components and is housed in the same enclosure as the VME bus. A fifth processor on the VME bus acts as bus controller, reading the data collected out of the global memory and transmitting it to a UNIX workstation. Here the data processing program runs under an Xmotif user interface. The program runs under UNIX mainly as it is still under development. Later, to produce a portable system with which to conduct field trials of the system, we shall port this program to a PC platform.

5. EXPERIMENTAL PROTOTYPE INCREMENTAL MOTION ENCODER

Three different prototype IME devices have been constructed so far in the course of the project. The first, with only two read heads, was an adaptation of a standard production model from Muirhead Vactric. This allowed testing of the basic concept. The second prototype was constructed from standard kit encoder parts. It has four photo-electric read heads and allowed initial testing of the system. During the use of this device it was found that despite its wholly digital nature the sensor was affected by electrical noise. This was due to the sensitivity of the latches on the IME interface to false edges on the encoder outputs. Despite shielding, it was found that joining the encoder to the interface was problematic due to this problem when the system was used near sources of electrical noise. Embedding the system as an ASIC within the encoder as a single shielded unit, thus eliminating the cable link was considered as a possible alternative solution. However, as we will almost certainly develop the processing architecture further, this was not feasible. The use of fibre optics to produce a noise immune link was the next obvious choice. The use of fibre optics also opens the prospect of increasing the number of degrees of freedom that the IME could be able to measure. If two devices were to be coupled to the ends of a single shaft, their two sets of output signals could feed back to an interface circuit with eight latches referring to the same counter time base. Comparing the two sets of angular positions and velocities would allow measurement of the torque along the shaft. This would require a large length of cables and freedom from noise would be critical.

A sideways section diagram of the optical fibre prototype is shown in figure 8. below. It consists of two circular metal plates which hold the fibres relative to the grating disk sandwiched between them. The encoder works on a light in, light out principle with no active devices within the disk housing. All the emitter detector electronics are mounted on remote circuit boards mounted in the VME bus housing. The light source is a 660nm LED and the fibre is a 1mm polymer. The detector is a pin photodiode driving directly into a trans-impedance amplifier followed by a digitizing circuit and a line driver. This gives the square wave output that feeds into the IME interface circuit. The disk is glass with a chrome pattern of 128 counts per turn. As this pattern is coarse it can be resolved by the fibres themselves without the need for relatively expensive lenses or gratings. Addition of these components would allow higher resolutions. Each of the four read heads has a single emitter fibre in the top plate the light from which is picked up by two detector fibres in the base plate. One fibre, outside the circumference of the disk acts as a reference which the output level of the second is compared to by the remote digitizing circuit. This second fibre detects the passage of the grating lines. One read head has a third detector to pick up the once per revolution signal. Figure 9. illustrates this configuration.

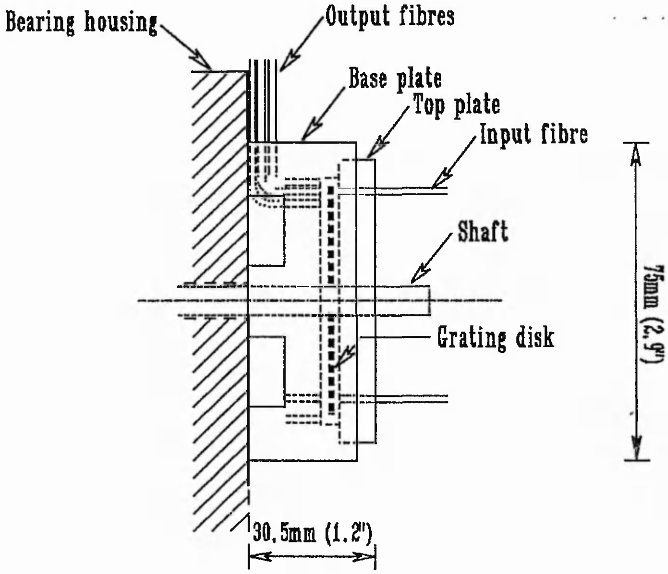


Figure 8. Prototype optical fibre IME.

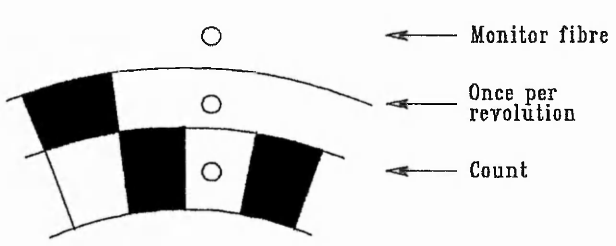


Figure 9. Detector fibre configuration.

6. INITIAL EXPERIMENTAL RESULTS

As the IME is a completely new type of sensor, interpreting the data it produces requires a new approach. The output, of displacement values at angular position increments, needs to be conditioned in a manner not yet fully developed before the conventional frequency analysis techniques can by applied. Full experimental testing of the IME has not yet been carried out. It is intended that a comprehensive set of tests will be undertaken to gain a detailed understanding of the IMEs ability to accurately detect bearing defects. We will then be able to make direct comparison of this method with the abilities of conventional sensors. A test rig has been constructed for this purpose in which different bearings can be mounted onto the shaft carrying the IME. This is illustrated in figure 10. So far only development tests have been carried out during the implementation of the computer system software. These have shown that the device gives results in the manner expected.

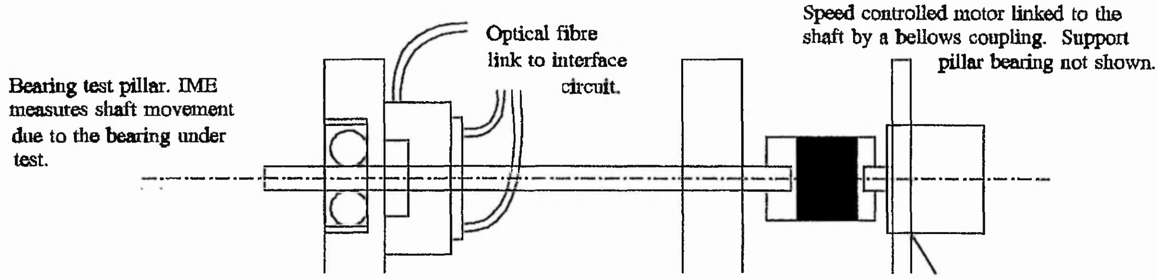


Figure 10. Experimental IME test rig.

The ability of the IME to produce high resolution measurements of angular velocity is shown in figure 11. below. This plot of angular velocity covers a single rotation with a duration of 0.05s. A regular angular vibrational waveform is clearly superimposed on an equally clear sinusoidal wave with a period equal to one rotation. The angular vibration has a constant period of 1/25th of a rotation and an amplitude which varies over the rotation. One of the strengths of the IME is its ability to directly relate the points of greatest variation to the precise angular position at which they occur. This combined with other parameters could be used to determine imbalance and shaft - housing rub.

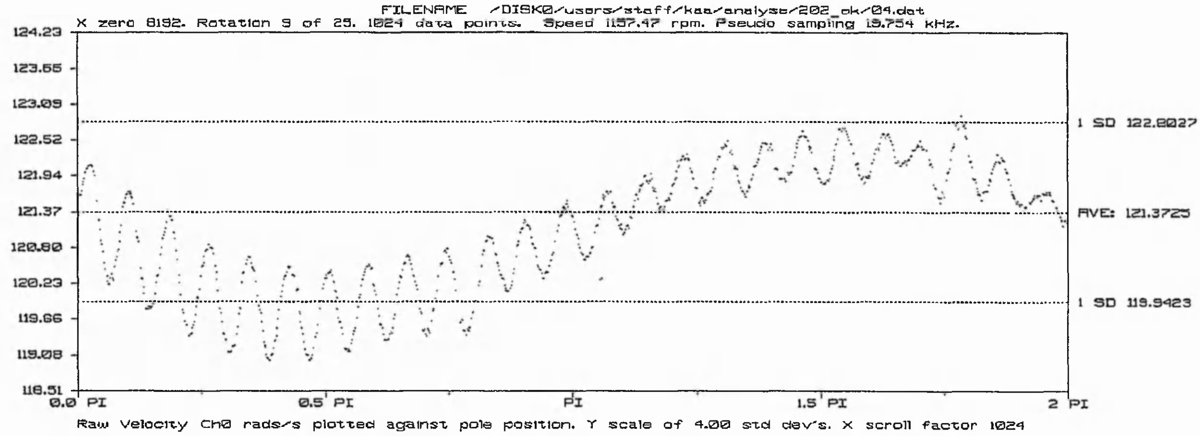


Figure 11. Angular velocity measured over a single rotation.

To test the ability of the system to measure displacement as it is expected to be able to do, a simple test using angular contact bearings was performed. Angular contact bearings are designed with a certain amount of play to accommodate shaft thrust. At one extent inner and outer races clamp tightly around the rotating balls. At the other there is a measure of freedom between the outer race and the ball cage. If the system was to perform as designed, it would have to be able to detect the difference between these two states. This is a simple simulation of an badly mounted shaft or the change between new and worn bearings. The bearings used in the test rig for this experiment were RHP 7202 BETN 40° angular contact ball bearings. They were new with no induced defects.

Figure 12. shows the resulting screen output taken from the system when the bearings were tightly adjusted. The displacement values measured by the two opposing pairs of read heads traces are shown plotted against each other. The data is from twenty five rotations, each of which has been superimposed. This has been judged a better way of estimating average position, at least initially, than a simple average of all the data. Each point represents the position of the nominal shaft centre

as if it were the actual centre of the optical disk. As in this case we are looking at the change in centre position the displacement due to disk eccentricity is not averaged out. The traces show a high degree of coincidence. All are broadly concentric about the same point and the divergence of their diameters is ~1μm.

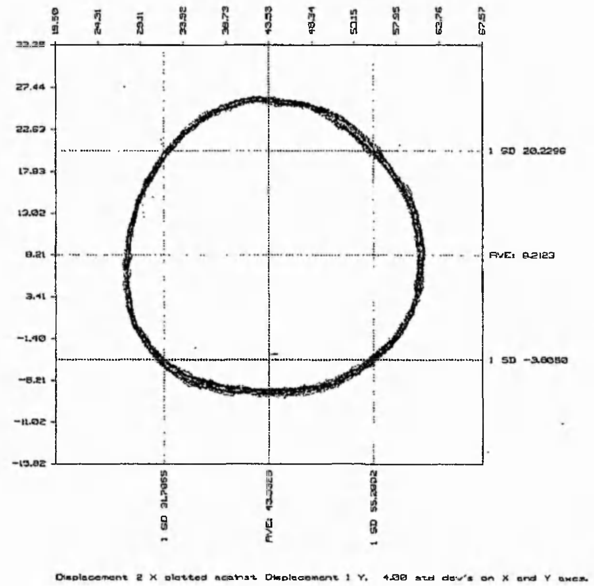


Figure 12. Displacement measured from a tightly adjusted angular contact bearing.

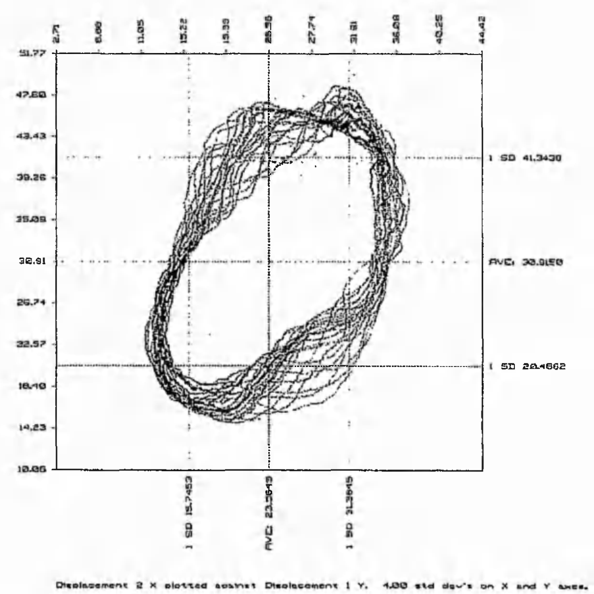


Figure 13. Displacement measured from a loosely adjusted angular contact bearing.

This contrasts sharply with the result of loosening the bearing, measured by the system and shown in figure 13. The increase in eccentricity is obvious, although the actual magnitude of displacement does not appear to increase. The divergence of successive traces produces the braided pattern apparent in the upper left and lower right of the graph. This pattern is due to the cyclic nature of the shaft displacement. The most extreme point in this region of the graph moves from trace to trace, reflecting the movement of the balls between the inner and outer races.

To test the ability of the system to detect bearing defects we have so far only made one simple test, contrasting the results from a new bearing with those from a bearing with an induced defect. The system averages the displacement due to grating disk eccentricity and takes this average from the original value. This gives a measure of the displacement of the shaft due to the movement of the bearings. An SKF NJ202 single row roller bearing was used for this test. Figure 14. shows the graphical display of this data from the bearing without the defect. Figure 15. is taken from a bearing with a ~1mm flat ground onto the inner race. The difference is quite marked. The defect has produced a distinct frequency on top of the original. This is a very encouraging result as it corresponds to what we would expect. Further work is required to produce a detailed analysis of the frequency spectrum of this output.

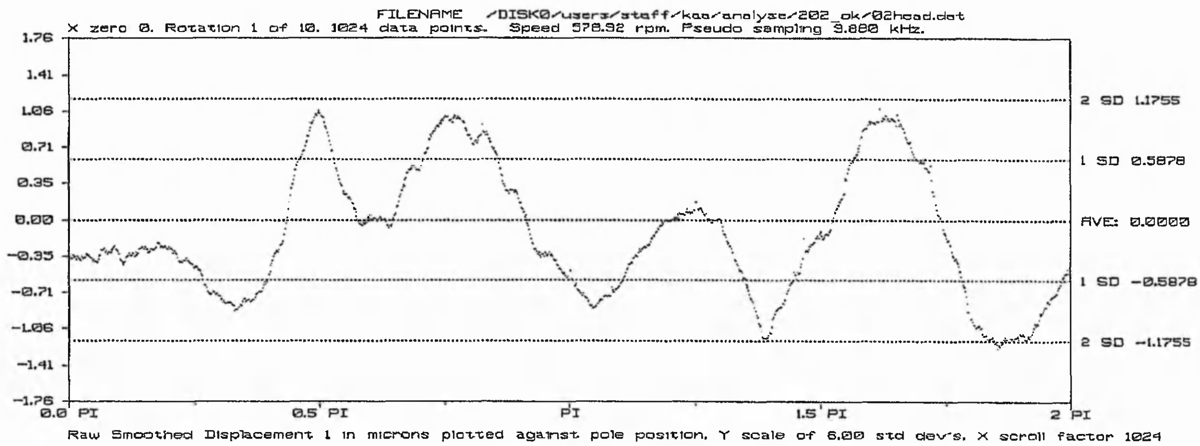


Figure 14. Displacement measured from an unmarked bearing.

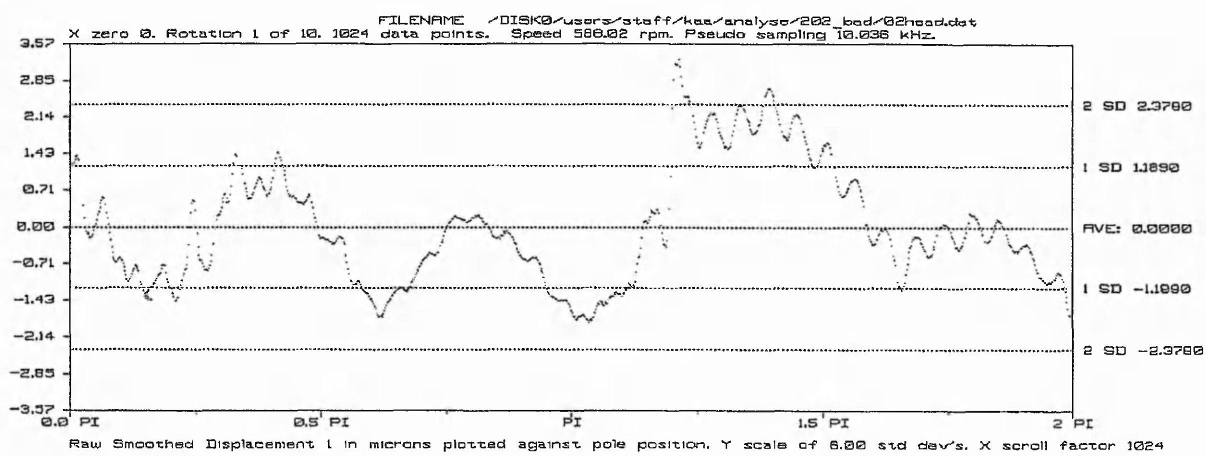


Figure 15. Displacement measured from bearing with defect.

7. CONCLUSIONS

A new type of sensor, the incremental motion encoder, has been developed. The IME is a sensor that is intended to provide more information on bearing condition and shaft rotary movement than any single conventional sensor. Integration of the measurement of angular velocity and shaft centre displacement with precise angular position enables an extension of vibration analysis. Initial results have shown that the sensor is able to make the raw measurements expected.

To produce a viable diagnostic system based on this sensor will require further work based upon detailed analysis of the results of a comprehensive series of tests. This work will require the development of the processing algorithms that take the raw data and produce the interpretable measures of bearing condition. These will be validated by field trials of the sensor in a industrial setting.

8. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of Muirhead Vactric Components Limited, in particular the hard work of Mr. Graham Freakes in support of the project and Mr. Steven Andrews in the design and production of the fibre optic prototype. We also wish to thank Mr. Sanjay Bhubuta of Hewlett-Packard Components Limited for providing kit encoder parts.

This work is funded within the Nottingham Trent University by the Department of Computing.

9. REFERENCES

1. G. P. Miller and S. L. McClymonds, "Maintenance cost avoidance through comprehensive condition monitoring", Jnt. ASME/IEEE Power Generation Conference, ASME, Boston, MA, October 21-25, 1990.

2. J. E. Berry, "How to track rolling element bearing health with vibration signature analysis", *Sound and Vibration*, vol 25, part 11, pp. 24-35, November, 1991.

3. Collacott, Ralph A., Vibration Monitoring and Diagnosis, ch 7, pp 115-156, George Godwin Limited, London, 1979.

4. P.Y. Kim, "A review of rolling element bearing health monitoring (III): preliminary test results on eddy current proximity transducer technique", 3rd International conference on vibrations in rotating machinery, pp 119-125, Mechanical engineering publications, London, 1984.

5. G.J. Philips, "Rotating machinery rolling element bearing performance monitoring using the fibre optic method", Reliability stress analysis and failure prevention methods in mechanical design, published by ASME, Aug. 1980, pp 37-43.

6. A. Ernst, Digital linear and angular metrology, pp. 25-28, verlag moderne industrie, 1989.

7. P.A. Orton, "Instrumentation and control for precision grinding machines," Ph.D Thesis, Nottingham Polytechnic 1990.

Appendix E.

An Optical Rotary Motion Sensor for the Real-Time Condition Monitoring of Rotating Machinery

K. Ayandokun, P. A. Orton, N. Sherkat and P. D. Thomas, Condition Monitoring '94, Mervin H. Jones, ed., pp 657 - 668, Pineridge Press, Swansea, 1994.
ISBN 0-906674-83-2.

An Optical Rotary Motion Sensor for the Real-Time Condition Monitoring of Rotating Machinery

K. Ayandokun, P. A. Orton, N. Sherkat & P. D. Thomas

The Nottingham Trent University, Department of Computing,
Burton Street, Nottingham NG1 4BU, United Kingdom. ©1993
Tel: +44 (0)602 418418 ext. 2944, Fax: +44 (0)602 486518,
Email: kaa@doc.nottingham-trent.ac.uk

ABSTRACT

Condition monitoring of many types of rotating machinery depends largely on an accurate assessment of the state of rolling element bearings. This paper describes a novel optical rotary motion sensor which integrates highly accurate measurement of the angular position and two dimensional centre position of a rotating shaft. These high resolution measurements of shaft centre movements are directly related to the condition of the bearings supporting the shaft and should offer a much clearer picture for condition monitoring than measurements made conventionally by sensors on the machine casing. The paper outlines the principle and construction of the device and the design and implementation of the computer architecture it is interfaced to. The current work being undertaken to interpret the raw data from the sensor's initial development tests is presented. The principle of the sensor has numerous other potential applications in the field of machine control and monitoring. These include feedback for automatically controlled bearings, measurement of shaft loading, torque and rotating tool wear. All using the same integrated sensor package.

1. INTRODUCTION

Monitoring the condition of critical components within industrial plant machinery as the basis of a predictive maintenance strategy has become the prime method of avoiding costly unscheduled production stoppages due to machine breakdown. Detection of the gradual deterioration of key components allows maintenance resources to be optimally employed, downtime of equipment is reduced and maximum economic use is gained from components designed with a certain wear lifetime [1]. Rolling element bearings which support shafts and spindles are the key components within many rotating machines. Their deterioration and failure is due to surface defects which appear on the bearing's internal elements due to progressive frictional wear. All bearings excite their supporting structures at certain natural frequencies related to shaft speed and the physical nature of the bearings. Defects on the bearing elements change these frequencies and therefore the parameter by which the condition of these bearings is most commonly monitored is vibration [2]. Used in this context the term vibration covers the three parameters displacement, velocity and acceleration. The choice between vibration parameters is made based on the calculated failure frequencies for the bearings to be monitored as each parameter works best over given a frequency range [3].

Acceleration and velocity are measured conventionally with analogue electro-mechanical sensors placed remotely on the machine casing. Signals measured in this way are subject to attenuation and distortion, particularly when the machine speed, and hence the vibrational energy due to any bearing defect is low. Consequently the location and direction of measurement are critical if the complex raw waveform, containing the frequencies representative of a fault condition, is to be correctly recorded.

Direct measurement of the displacement of a shaft relative to its mountings should be preferable to remote velocity or acceleration sensing as it eliminates the problems associated with sensor placement. Shaft displacement increases with bearing deterioration but vibration velocities measured on the bearing housing may not always be sensitive to this increase. Unfortunately, conventional displacement spectra, measured with proximity probes, can miss a great deal of bearing health information, as much of the spectral content relevant to bearing defects is suppressed.

Our new device, the Incremental Motion Encoder (IME), extends the use of displacement measurement as a condition monitoring technique. It allows displacement to be related to angular position to an unparalleled degree and increases substantially the useful bandwidth over which measurements can be taken. It is also an optical, non-contact, 100% digital device with no analogue to digital conversion.

2. THE HIGH PRECISION OPTICAL SHAFT ENCODER PRINCIPLE

The incremental rotary Optical Shaft Encoder (OSE) is a device commonly used in industry used for the measurement of angular position and angular velocity. For high precision angular measurement applications which require resolutions of 0.001° or higher, it is not technically feasible to use an optical grating with the necessary line spacing. High precision optical shaft encoders therefore use the principle of interpolation. This is an analogue or digital process, depending on the accuracy required, and produces additional phase shifted signals from the encoder outputs. For example an 18° phase shifted signal produces a 5x interpolation of the original input. The output signals have five times the frequency of the input signals and twenty edges are produced for each actual grating spacing. Digital interpolation allows for the highest possible subdivision of the grating spacing. Interpolation of 1024x on a 36000 line grating allows a measuring step of 0.00001° to be achieved in this way [5].

Due to their degree of angular accuracy, high precision encoders are particularly susceptible to shaft axial eccentricity and angular vibration as a source of error. Precise couplings between encoder and shaft are used to limit these errors, but precise mechanical transfer of the motion of the shaft to the encoder without error is not possible. Use of a second read head within the conventional high precision encoder allows compensation for these errors. Encoders with digital interpolation have four read heads which together allow for a positional correction prior to interpolation. As will be explained the IME uses four read heads in an analogous manner to provide the angular data used to resolve shaft centre position.

3. THE INCREMENTAL MOTION ENCODER PRINCIPLE

Recent work with precision encoders for the control of grinding machines [6] produced the idea of a sensor, the IME, which would use shaft centre position measurements as a vibration monitor [7]. A method has also been developed for producing the high degree of angular accuracy required from an optical grating of relatively low resolution. In its construction, data collection method and calculation algorithms the IME differs markedly from a measurement system with a broadly similar basic principle developed for torsional vibration measurement [8].

3.1 Method of shaft centre position calculation

Figure 1. illustrates the way in which the IME calculates the shaft centre position. The device is constructed with four read heads spaced at 90° intervals around the circumference of an optical grating disk, marked A - D. Each diametrically opposed pair of read heads is used to calculate the position of the shaft centre along the perpendicular bisecting the diameter joining them. Consider the read head pair A and B, distance l microns apart. If the centre of the shaft X were in line with the two read heads at the point X' then the difference between the angular position measurement derived at A and that at B, θ, will be π radians. Assume then that X is displaced to the left as shown in the figure 1. In this case the angle θ will no longer be π. The shaft centre displacement from X' to X, d is given by:

$$d = \frac{(\pi - \theta)\, l}{4} \tag{1}$$

This is an approximation which can be made for centre displacements of the order of tens of microns if the disc is of the order of tens of millimeters in diameter.

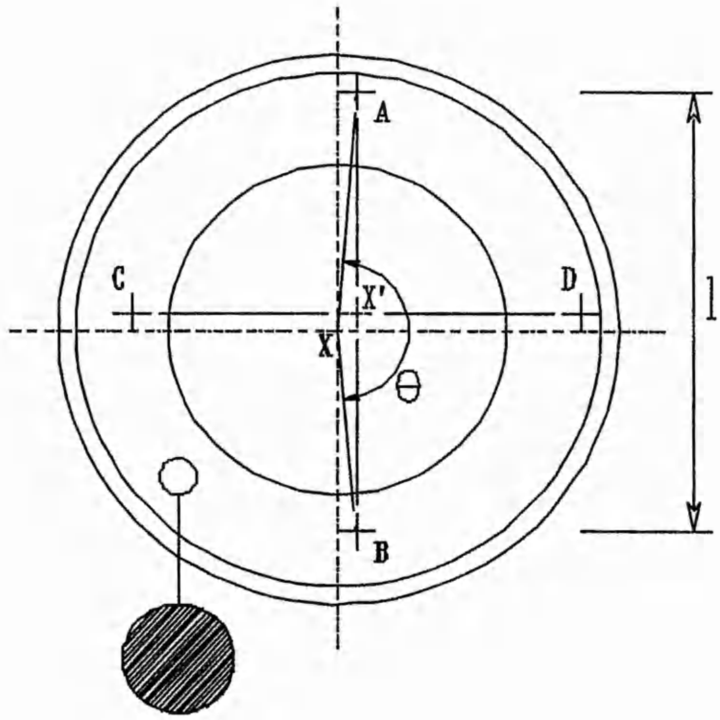


Figure 1. Shaft centre position calculation.

Each pair of diametrically opposed read heads provides the position of the shaft centre in a single dimension. Together the angular measurements from four as shown allow the shaft centre to be located in two dimensions. The error in this calculation due to the radial movement of the sensors is made negligible by an iterative method once the initial estimate for the displacement of the encoder disc centre has been found. A detailed analysis of the trigonometry shows that three read heads spaced at 120° would be sufficient to allow the disk centre position to be calculated. This method however requires far more computation as a complex sine cosine relationship would then need to be evaluated . Our industrial collaboration with a manufacturer of encoders has also revealed that the production of accurately placed diametrically opposed read heads is significantly easier. This allows conventional techniques to be used for the production of the read head optics.

3.2 Time base interpolation

The basis of the high angular accuracy of the IME is the use of time base interpolation. Rather than count the edges of the encoder output wave in the equal time duration of a sample period as conventionally, figure 2a., the IME uses a precision clock oscillator to generate a time base which is sampled at the equal intervals of angle represented by the encoder grating interval, figure 2b.

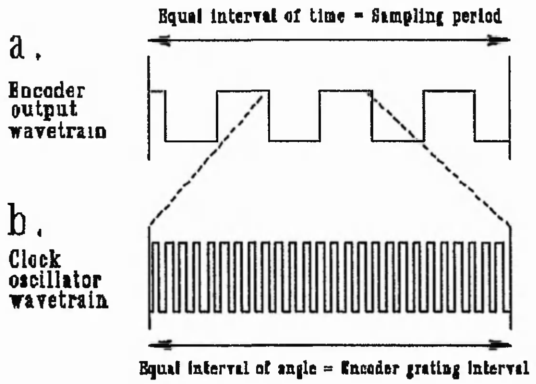


Figure 2. Time base interpolation.

This allows the sub-division of the measuring step to be limited only by the frequency of the time base oscillator, taking the ability to count high frequency edges as a limiting factor. This is however a dynamic interpolation method, the number of interpolation steps N is given by:

$$N = \frac{F_{tb}}{rps \; x \; G} \tag{2}$$

F_{tb} = frequency of the time base, rps = rotations per second of the encoder disk, G = number of grating lines on the encoder disk.

The above equation shows that the IME cannot measure point to point angular rotation in the manner of a conventional precision encoder with any more than the actual resolution of its disk. The high degree of angular accuracy given by the sampled time base value is used by the system internally to derive the position of the shaft center by interpolation between angular values recorded from opposite read heads. A simple linear interpolation algorithm, detailed in [4], is then used to apply equation (1) to calculate displacement from each pair of corresponding angular values.

To produce the time base interpolated angular measurements in a form that can be processed by a computer system a special interface circuit is required to link the computer with the IME outputs. This circuit consists of a precision clock oscillator driving a single free running synchronous counter joined to the input pins of four identical latches by a bus. The activating pins of the latches are connected to the square wave outputs of each encoder read head and their output pins are connected to the computer system. The falling edge of the incoming square wave from each of the IME's four read heads latches a value from the counter onto its respective latch representing the exact time it arrived at that latch. The computer system is signaled that data is waiting on the latch by the same falling edge and the value is then collected and stored in a memory table with a column for each read head. The position in each column gives the angular displacement, the difference between successive values in each column gives the time base interpolation count for that read head. The current experimental circuit is constructed using a counter and latches which are 16-bits wide formed from FAST TTL logic chips, together with a 24 MHz oscillator.

The use of time base interpolation results in the ability of the IME to produce highly accurate measurements of angular velocity precisely related to angular position. Conventionally the angular displacement measured within one sample period will be accurate to +/- one encoder measuring step and the velocity calculated will not bear a relationship to angular position. By using the angular distance of the measuring step, averaged to account for interslot errors, and the time base value recorded over this angular distance, the IME is capable of resolving changes in the angular velocity which occur between measuring steps.

3.3 Prototype IME devices

Two IME sensor prototypes have been constructed to test the principle outlined above. One has been built using optical fibre technology to successfully prove the ability of a sensor of this type to be made as an entirely optical device [4]. The other uses standard commercially available Hewlett Packard kit encoder parts, four read heads and a 1024 line grating disk.

4. COMPUTER ARCHITECTURE

4.1 Conceptual architecture

The computer architecture required to support the IME is functionally divided between data collection and data processing, analysis and display. As shown schematically in figure 3. each output channel of the IME interface circuit is independently interfaced to the computer system which performs the measurement calculations by a Data Collection Unit (DCU).

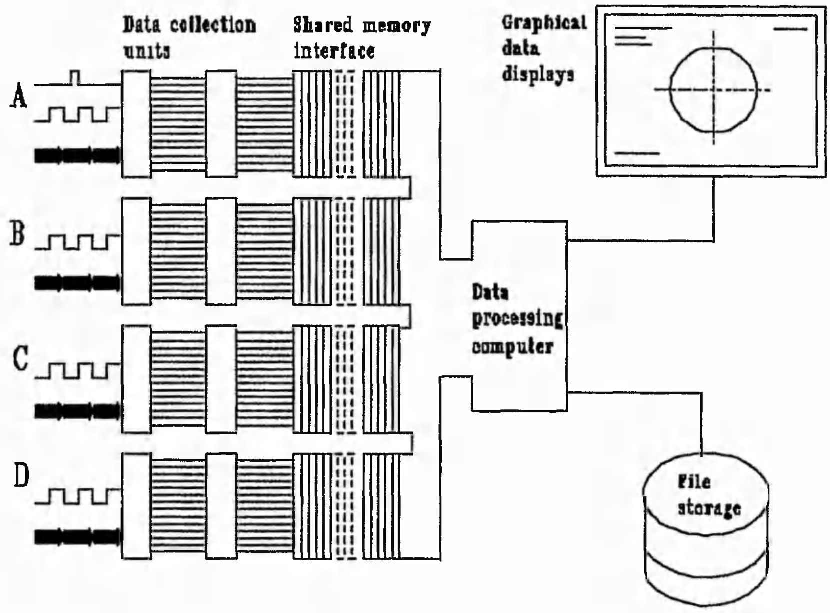


Figure 3. Conceptual computer architecture.

The input to each DCU is composed of the square wave output directly from the IME and the latched counter values from the interface circuit. Each DCU comprises a parallel interface linking the latches of the IME interface circuit to the DCU's data bus and an internal data management unit . The count values are read from the parallel interface into the dual port FIFO memory of the DCU under the control of the internal data management unit responding to the falling edges of the IME output. The processing computer system sees the counter values across the shared memory interface as the four column table described above. The start of data collection is synchronized by the once per revolution signal from one read head, shown received by DCU A. Use of this signal allows the data collected to be precisely related to shaft angular position externally. After this signal the data table is filled with a number of values equal to a whole number of rotations. Data collection then stops and the values in the table are then processed to give values for displacement and angular velocity which may then be analysed with digital signal processing algorithms to detect bearing conditions. This architecture is designed to allow the maximum data collection rate as data must be collected and stored at the rate it is produced by the sensor, which is dependent on shaft speed. Collection and display of data is controlled via a graphical user interface. The system can gather data continuously, averaging the resulting values for display or writing them to disk for historical analysis.

4.2 Implementation of experimental architecture

The conceptual architecture and operation outlined above can be implemented using various components. For our experimental system we have implemented the DCUs with four monoboard 16 MHz MC68000 computers connected via a VME bus to global memory. Each board has a parallel interface and local memory with the processor acting as the data management unit. These allow the system to collect data at shaft speeds up to 8000 rpm from a sensor with a 1024 grating line disk. The IME interface circuit is housed in the same enclosure as the VME bus. A fifth MC68000 on the VME bus acts as bus controller, reading the data collected out of the global memory and transmitting it to a UNIX workstation where the data processing program runs under an XMotif user interface. A long term production objective of our industrial collaborators is the production of the IME interface and processing computer as an ASIC device. Embedded with the encoder this would produce a compact sensor. The function of a DCU is essentially that of a DMA controller and an intermediate system could be built around such components and simple dual port FIFO memory chips on a PC card. With the analysis program running on a PC this would produce a portable system with which to conduct field trials in an industrial setting.

5. EXPERIMENTAL TEST RESULTS

The IME is a new type of sensor whose output needs to be interpreted in a way which is not yet fully developed. We are currently engaged in research and development aimed at determining the IME's ability to detect bearing defects and at the same time assess its other possible uses. The data which we derive directly from the sensor, displacement in two dimensions and angular velocity, is effectively sampled at equal intervals of angle rather than time. It is possible to resample the data using the recorded time base values to derive the displacement at fixed intervals of time, however this would remove the inherent positional information. Consequently we have found it advantageous look at our data in terms of angular position. A test rig has been constructed [4], which allows different bearings to be mounted on a shaft in a test pillar which also supports the IME. We have already been able to see that the sensor monitors sub-micron relative motion of a shaft as the theory predicts. Work still needs to be done to quantify the effect of possible sources of error in the system such as read head misalignment, disk flexure and interslot errors.

The centre of the encoder disk is not expected to be coincident with the actual centre of the shaft rotation. The point tracked by the sensor is the nominal centre of the encoder disk as it orbits around the actual centre of rotation. This orbit will be modified by the motion of the actual centre and is measured relative to the zero position indicated by the intersection of the diameters joining the read heads. Figure 4a. is an output trace from the system with a tightly adjusted new angular contact bearing supporting the shaft in the test pillar rotating at 280 rpm. The displacement values in microns from the two axes of measurement are shown plotted against each other for a single rotation. Each of the 1024 separate points on the trace is derived from the four digitally recorded time base interpolated angular difference values described.

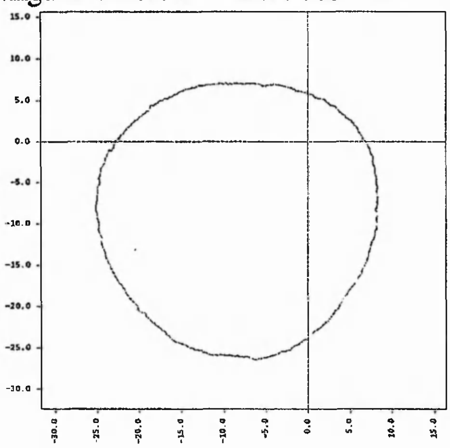


Figure 4a.

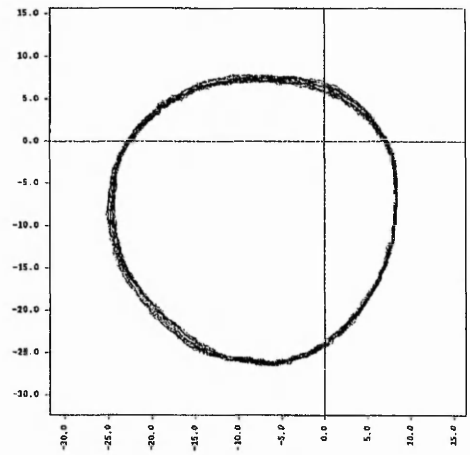


Figure 4b.

In figure 4b. the traces from 25 rotations show a high degree of coincidence. Points from the same angular position plot within ~1μm in each successive trace. This degree of convergence is currently considered by us to be a good macroscopic indicator of relative bearing condition. As wear occurs the difference the sensor will measure between the same point on successive rotations will increase.

By changing the relative position of the inner and outer races of an angular contact bearing it is possible to increase the amount of freedom between the balls and the outer race. We have used this as a way to simulate the possible condition of a badly adjusted or grossly worn bearing. The output from the sensor monitoring this condition has produced a very interesting set of results which may reveal phenomena not previously seen with this detail. Figure 5a. shows the braided net-like pattern produced when 14 successive traces from the loosened bearing rotating at 300 rpm are overlaid. The X and Y axes are displacement in microns as before. Looking closely, it is clear that the pattern is formed from 7 distinct paths traced by the shaft centre on successive rotations, each repeated twice. Our interpretation of this is that each path represents a particular set of positions taken up the bearing's internal elements during each rotation. The coincident parts of all the traces are likely to be the points at which the shaft is naturally constrained by its weight and the influence of the rest of the rig. At higher speeds complex motion within a single rotation can be observed as the inner race and rolling elements supporting the shaft skip and bounce against the outer race, shown in figure 5b, a single rotation monitored at 630 rpm.

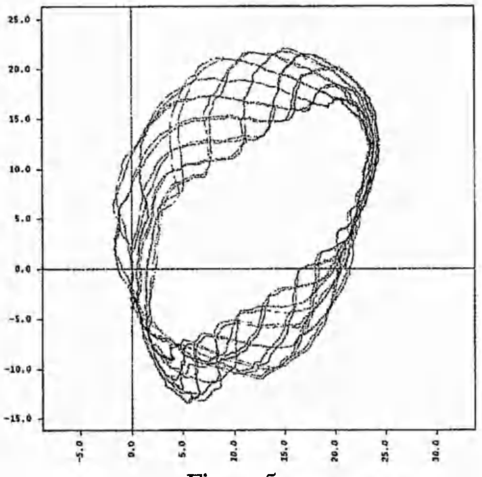


Figure 5a.

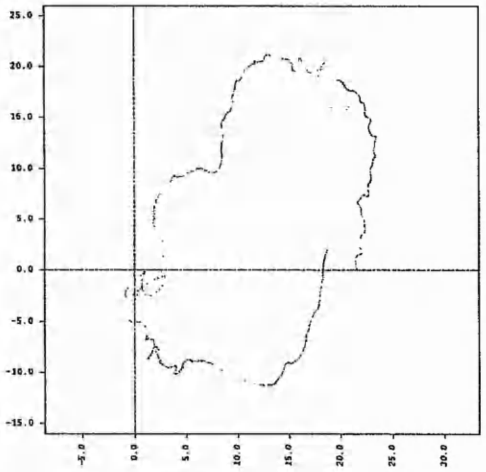


Figure 5b.

Figures 4 and 5. have shown the data from the sensor with the eccentricity of the disk included. To look for bearing damage our current approach is to remove this eccentricity to leave the motion of the shaft about its average centre rather than zero point of the sensor. An SKF NJ202 roller bearing was used in our initial test of the ability of the sensor to detect an induced bearing defect. An ~1mm flat was ground onto the inner race. Figure 6a. shows a trace recorded from the damaged bearing rotating at 2700 rpm, axes units are as before. The nature of the fault on the bearing is not apparent when the data is viewed like this. By averaging the displacement at each of the 1024 angular positions over a series of 25 rotations an approximation can be made for the path of the disk centre with disk eccentricity removed, figure 6b.

We have started developing the analysis of this varying displacement signal by using conventional signal processing techniques to look for the frequency components related to bearing damage [2]. Figure 7. shows the frequency spectrum resulting when a fast fourier transform is applied to the Y axis displacement signal after the average has been subtracted. The RMS power of the displacement is shown plotted against the frequency in cycles per revolution. The sensor has picked up the inner race defect frequency quite clearly. Further work in this area needs to be done to interpret all the peaks in the spectrum and determine their sources. We specifically want to determine whether our sensor will be sensitive to all bearing defect frequencies and what extra information we can obtain by applying filters to the raw signal.

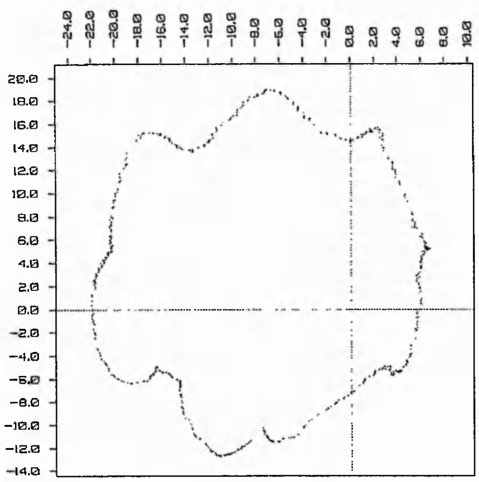


Figure 6a.

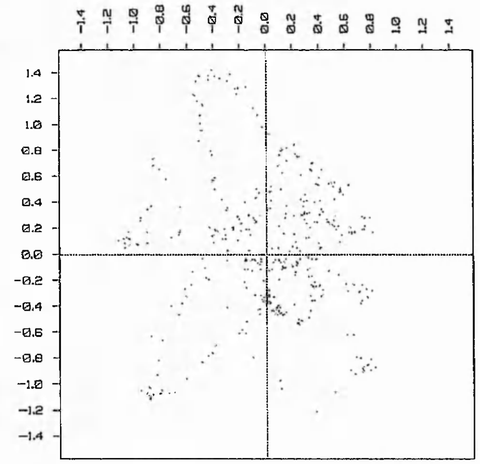


Figure 6b.

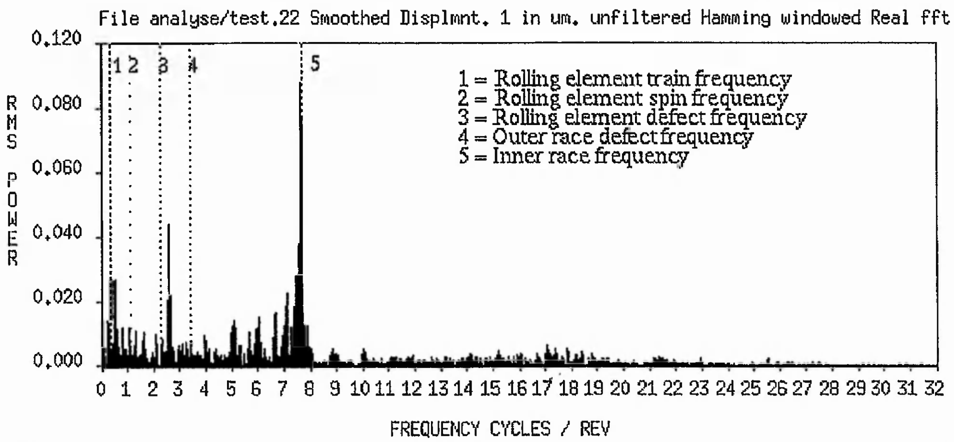


Figure 7.

6. CONCLUSIONS

The Incremental Motion Encoder is intended to provide more information on bearing condition and multi-dimensional shaft movement than any single conventional sensor. Initial results have shown the sensor is able to monitor shaft motion as predicted theoretically by its design. It has also demonstrated an ability to detect one type of bearing defect with minimal signal conditioning. We are confident that further work will develop the processing algorithms required to produce a viable diagnostic system, able to produce interpretable measures of bearing condition, based on this sensor.

8. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of Muirhead Vactric Components Limited, in particular the hard work of Mr. Graham Freakes in support of the project and Mr. Steven Andrews in the design and production of the fibre optic prototype. We also wish to thank Mr. Sanjay Bhubuta of Hewlett-Packard Components Limited for providing kit encoder parts.

This work is funded within the Nottingham Trent University by the Department of Computing.

9. REFERENCES

1. G. P. MILLER and S. L. MCCLYMONDS - Maintenance cost avoidance through comprehensive condition monitoring, Jnt. ASME/IEEE Power Generation Conference, ASME, Boston, MA, October 21-25, 1990.

2. BERRY, J. E. - How to track rolling element bearing health with vibration signature analysis, Sound and Vibration, vol 25, part 11, pp. 24-35, November, 1991.

3. COLLACOTT, RALPH A. - Vibration Monitoring and Diagnosis, ch 7, pp 115-156, George Godwin Limited, London, 1979.

4. ORTON, P.A. *et al.* - A fibre optic rotary motion sensor for the real-time condition monitoring of rotating machinery, Proc. SPIE Symposium on Optical tools for manufacturing and advanced automation, vol. 2072, September 1993, Boston.

5. ERNST, A. - Digital linear and angular metrology, pp. 25-28, verlag moderne industrie, 1989.

6. Orton, P.A. - Instrumentation and control for precision grinding machines, Ph.D Thesis, Nottingham Polytechnic, 1990.

7. Orton, P.A. - Shaft Displacement Measuring System, International patent application number PCT/GB92/00811, World Intellectual Property Organization, International Publication number WO 92/19936 12 November 1992.

8. Wang, P. *et al.*, - A Torsional Vibration Measurement System. IEEE Transactions on instrumentation and measurement, vol. 41, no. 6, pp. 803 - 807, Dec 1992.

Appendix F

Detecting rolling element bearing defects with the optical incremental motion encoder

K. Ayandokun, P. A. Orton, N. Sherkat & P. D. Thomas, in *Industrial Optical Sensors for Metrology*, Kevin G. Harding, H. Philip Stahl, Editors, Proc. SPIE 2349, pp. 53-64, 1995.

Detecting rolling element bearing defects with the optical incremental motion encoder.

K. Ayandokun, P. A. Orton, N. Sherkat & P. D. Thomas.

The Nottingham Trent University, Department of Computing, Burton Street, Nottingham NG1 4BU, United Kingdom,
Tel: +44 (0)115 941 8418 ext. 4193, Fax: +44 (0)115 948 6518, Email: kaa@doc.ntu.ac.uk

ABSTRACT

This paper describes the experimental use of a new type of optical sensor to detect rolling element bearing defects for the purpose of condition monitoring. The incremental motion encoder (IME) is a patented development of the industrially proven technology of the optical shaft encoder. It combines the measurement of angular position with that of two dimensional shaft centre position to sub-micron accuracy. The operational principle of the IME is explained. Its high resolution measurements allow shaft motion directly related to the condition of the supporting bearing to be sensed. This presents potential advantages over current bearing condition monitoring sensors which monitor transmitted vibration in the machine structure. A model explaining the effect of a bearing inner race defect on the IME signal is presented. Based on this model the statistical parameter kurtosis is proposed as a method of detecting such defects from the IME signal. The results of experiments comparing the IME signal from a new undamaged roller bearing with the signal from the same bearing with an induced defect are presented. These confirm the application of the model to defect detection.

Keywords: incremental motion encoder, rolling element bearings, predictive maintenance, condition monitoring, kurtosis

1. INTRODUCTION

As industrial plant machinery has grown increasingly more complex and automated, predicting the failure of key machine components has become a necessary method for avoiding costly breakdowns and maintaining safety. In order to be effective this approach, known as predictive maintenance, must be based on efficient analysis of accurate assessments of the condition of the key components[1]. The performance of many rotating machines relies on the rolling element bearings supporting shafts, spindles, gears and other components. The condition of these bearings is of critical importance in machines of this type. Due to various operating factors such as lubricant contamination, shaft misalignment or improper loading, bearing service life may vary considerably from manufacturers estimates, potentially leading to unexpected failures[2]. It has been estimated that only 10 to 20% of bearings reach their design life[3]. The monitoring of bearing condition is therefore extremely important to prevent otherwise unpredictable machine breakdowns.

Bearing condition is most commonly monitored by measuring machine vibration[4,3], although analysis of lubricating fluids for wear products is also an important technique[5,6]. Vibration has the advantage of giving warnings about the presence of other types of machine defects, such as misalignment or imbalance, as well as bearing damage. Alternative monitoring methods, if used in isolation, limit the type of faults that can be detected therefore failure can result from an undetected fault. Newly installed bearings excite their supporting structures at certain natural frequencies related to shaft speed and the normal contact of rolling elements and races. As the bearing ages, the vibration increases due to gradual breakdown of the contact surfaces. The presence of surface defects in the bearings results in the production of impulsive vibration at frequencies determined by the location of the defect and the shaft speed[3]. These impulses excite the bearing rings and the bearing housing, propagating to the machine casing where the vibration can be measured by an analogue electro-magnetic sensor, most commonly a piezoelectric accelerometer.

The vibration measured by a surface mounted sensor is determined by two factors. The bearing condition itself and the mobility of the transmission path between the point where the impulsive forces act and the sensor location[7]. Ideal conditions for monitoring exist where the sensor can be mounted with a direct path between itself and the bearing, in the centre of the load zone and with the nominally loaded, properly lubricated bearing rotating above a minimum speed of 1000 rpm[8]. Where these conditions are not met, which is the case in most complex machines, particularly those with a high density of vibration sources, proper diagnosis of individual components may be extremely difficult. Despite its good performance over most of the frequency spectrum of interest for condition monitoring, the piezoelectric accelerometer is unable to monitor the frequencies in the range commonly produced by low speed machines. Low rpm, commonly combined with massive structure, means that little energy is transmitted to the machine casing resulting in low levels of vibration[9]. These low level, low frequency measurements are limited by the electromechanical efficiency of the piezoelectrical material used in accelerometers[10]. The noise level produced by an accelerometer and its preamplifier overwhelm the low signal due to the actual vibration.

Rather than monitor bearing condition remotely from the machine casing, there are good arguments why relative shaft displacement measurement should be preferable[11]. Principly, relative displacement between a shaft and its mounting is directly related to the condition of the bearing supporting the shaft. Currently shaft displacement monitoring is almost exclusively limited to fluid film journal bearings.

The use of the incremental motion encoder (IME) as a condition monitoring device extends the use of relative displacement measurement as a condition monitoring technique to rolling element bearings. It has the advantage of sensing the mechanical motion caused by the presence of a defect directly allowing monitoring of low speed machines. It can also directly monitor the state of single bearings in complex machines.

2. THE INCREMENTAL MOTION ENCODER PRINCIPLE

The operating principle of the IME has been described in detail previously[14]. This section presents a summary to set the model that has been developed into context.

The IME is based on the combination of a well understood principle used in high precision optical shaft encoders and a method for deriving a high degree of angular accuracy from a relatively low resolution encoder disk[12]. When using any optical shaft encoder, shaft centre movement is recognised as a possible source of error in the angular position measurement. When the disk is of low angular resolution, this can be minimised by the use of a flexible coupling to limit the transmission of shaft axial movement to the encoder disk. High precision encoders are highly susceptible to this source of error and make a correction for the shaft's axial eccentricity prior to their calculation of angular position[13]. This correction is achieved by determining the angular difference indicated between diametrically opposed points on the grating disk. For this purpose two, or for the highest accuracy four, read heads are used. The measure of shaft centre position derived in this way is treated as an error and is discarded after correction has been made for it.

The construction of the IME does not attempt to isolate the grating disk from the motion of the shaft. On the contrary, it has the disk directly mounted on the shaft to sense shaft motion fully. Four read heads mounted on the machine structure at 90° intervals around the disk allow the position of the shaft centre to be calculated in the plane of the disk. Figure 1 schematically illustrates this construction and the geometry which allows the centre position calculation to be made. When the centre line of the disk coincides with the line joining the read head pair A - B, distance L apart, the horizontal displacement, d, is zero and the difference between the angular position measured at A compared with that at B is equal to π. When the horizontal displacement increases, the value of the angular difference between A and B will decrease to θ. The relationship between θ and the horizontal displacement is given by equation 1.

$$d = \frac{(\pi - \theta)L}{4} \qquad (1)$$

This is an approximation which can be made for centre displacements of the order of tens of microns if the disk is of the order of tens of millimeters in diameter. It is possible to use three read heads spaced at 120° degree intervals. However this requires the evaluation of a more complex trigonometric relationship.

The basic measurement made by the IME is of the time of passage of grating lines past each read head. When the disk is centred relative to the read heads, each pair of diametrically opposed read heads will record the passage of a grating line at the same instant as the lines are ruled in pairs forming a diameter across the disk. As the disk moves with shaft displacement a difference will be measured between the time values recorded for grating lines at each read head of the measuring pair. This time difference will be proportional to the angular difference θ. To make these time measurements the IME uses a counter circuit driven by a precision clock oscillator. This counter sets up a timebase to which the relative time of passage of each grating line at each read head can be precisely related. The timebase has the effect of sub-dividing the angular measuring step of

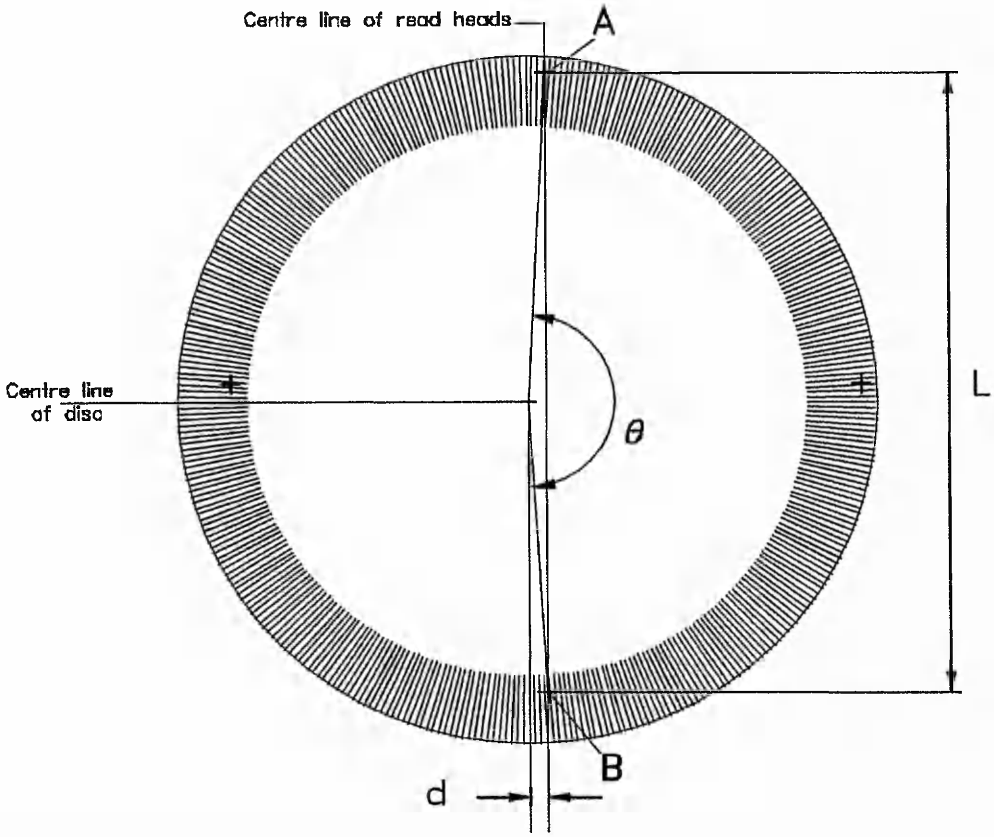


Figure 1. Shaft centre position calculation.

the encoder disk by the number of clock cycles recorded in the interval between each grating line. This effectively increases the angular resolution of the disk by a factor limited principly by the speed of the clock oscillator. This angular resolution determines the accuracy with which θ can be measured and hence the accuracy to which the displacement can be calculated. The angular difference is calculated from comparison and interpolation between the timer counter values latched for each read head pair[14]. The software required to do this calculation and then to apply equation 1 runs on the computer system to which the IME sensor is interfaced.

The IME is able to measure angular velocity highly accurately due to the use of the timer counter since it records to high resolution the time taken by the disk to rotate through the angular distance between grating lines. Changes in angular velocity which occur between measuring steps can therefore be estimated. The fact that each angular velocity value is also precisely related to an angular position in the disk's rotation allows measurement of torsional vibration more accurately than conventional encoders.

There are many potential applications for IME sensors, which are capable of measuring all the parameters of a conventional shaft encoder and giving shaft centre position as well in the same integrated package. Possible uses include monitoring of shaft loading and imbalance and active vibration control for automatically controlled magnetic or fluid film bearing. Two devices with a common timebase spaced along a shaft would allow for transmission torque measurement.

The recent development of bearings with integrated encoder sensors[15] offers the possibility, combined with the IME principle, for self monitoring bearings able to sense a number of machine parameters.

3. THE EXPERIMENTAL IME SENSOR

Sensors based on the IME principle can be implemented using a number of different technologies capable of giving a square wave output from a rotating marked surface. As well as opto-electrical read heads and a grating disk, optical sensors using a reflective drum on the shaft could also be used. For harsher environments, where contamination would limit the effectiveness of an optical sensor, electromagnetic sensing elements can be used. In many cases rotating machines have a sensor for measuring angular motion which produces the type of signal to which the IME principle can be applied with the addition of extra read heads.

Within the scope of this project the concentration has been on the use of optical devices. This has been partly due to the previous experience of the authors with devices of this type and also the practical considerations of engineering a prototype device.

Two IME sensors have been constructed so far as experimental prototypes. One was built using fibre optic technology to successfully prove the ability of a sensor of this type to be made as an entirely optical device[14]. The other was built using standard commercially available kit encoder parts from Hewlett Packard, four opto-electrical read heads and a 1024 line grating disk. This is the device which is being used for current experimental work because of its higher resolution.

An experimental rig has been constructed to allow either encoder, or any other compatible with the rig, to be mounted on a shaft and to have its responses to shaft movements monitored. The rig provides the facility for the bearings supporting the shaft to be interchanged so that the sensor can be used to detect changes in the condition of various test bearings.

In previous descriptions of the IME's application to determining bearing condition we have not been able to explain or predict the nature of the sensor signal's response to a defect. Although shaft motion was clearly detectable from the sensor output[14], frequency analysis of the signal[16], although initially promising, was found to be inconclusive. In order to develop a method of interpreting the sensor signal a model for the effect of a defect on the IME signal was required.

4. A MODEL TO EXPLAIN THE EFFECT OF A BEARING INNER RACE DEFECT ON THE IME SIGNAL

4.1 The nature of the IME signal

The nature of the signal recorded by an IME device may be explained with respect to its construction with reference to figure 2. When an encoder disk is mounted on a shaft the disk centre, at X, is not expected to be coincident with the actual centre of the shaft's rotation, at R. The point tracked by the sensor is the centre of the encoder disk which will describe an orbit around the actual centre of rotation, R. This orbit will be modified by the motion of the centre of rotation, Δx and Δy. The coupled Δx, Δy motion of the disk centre is recorded by the IME as displacement measured relative to the zero position indicated by the intersection of the diameters joining the read head pairs A - B and C - D at X'. The distance between the centre of rotation and the centre of the disk, the eccentricity *e*, is determined by the radius of the disk centre orbit in this example of simple motion.

The displacement measured from each axis of the sensor produces a sinusoidal waveform. There is a 90° degree phase difference between the two signals. Figure 3. shows the output from the sensor monitoring a single rotation of a new SKF NJ202 roller bearing. The traces show displacement in microns on the y axis and angular position from 0 to 2π on the x axis. The top trace is derived from the data recorded in the vertical axis of measurement by read head pair C - D, the lower trace is from the horizontal axis measured by A - B. Each trace is made up of 1024 individual displacement data points calculated from the two sets of timer counter values recorded from each read head pair, one per disk grating line, or pole. The angular interval between each data point is equal, the time interval is variable, determined by rotational speed. This is a principle difference between the data presented by the IME and that from conventional analogue sensors where data is sampled at equal intervals of time with no angular reference. Since the timer counter values provide a complete time record it is possible to resample the IME data to derive the displacement at equal intervals of time. However this would remove the inherent angular position information which allows the exact comparison of features in successive rotations. Consequently the data is generally presented in terms of equal intervals of angular displacement. The start of data collection is synchronised to one angular position on the grating disk. This allows features of individual rotations to be related to each other and a known position on the machine.

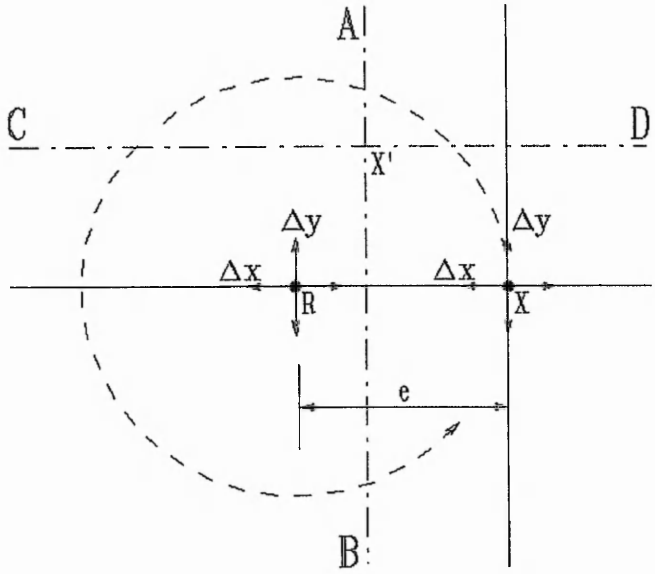


Figure 2. The relationship between the encoder disk centre, the centre of shaft rotation and the signal produced.

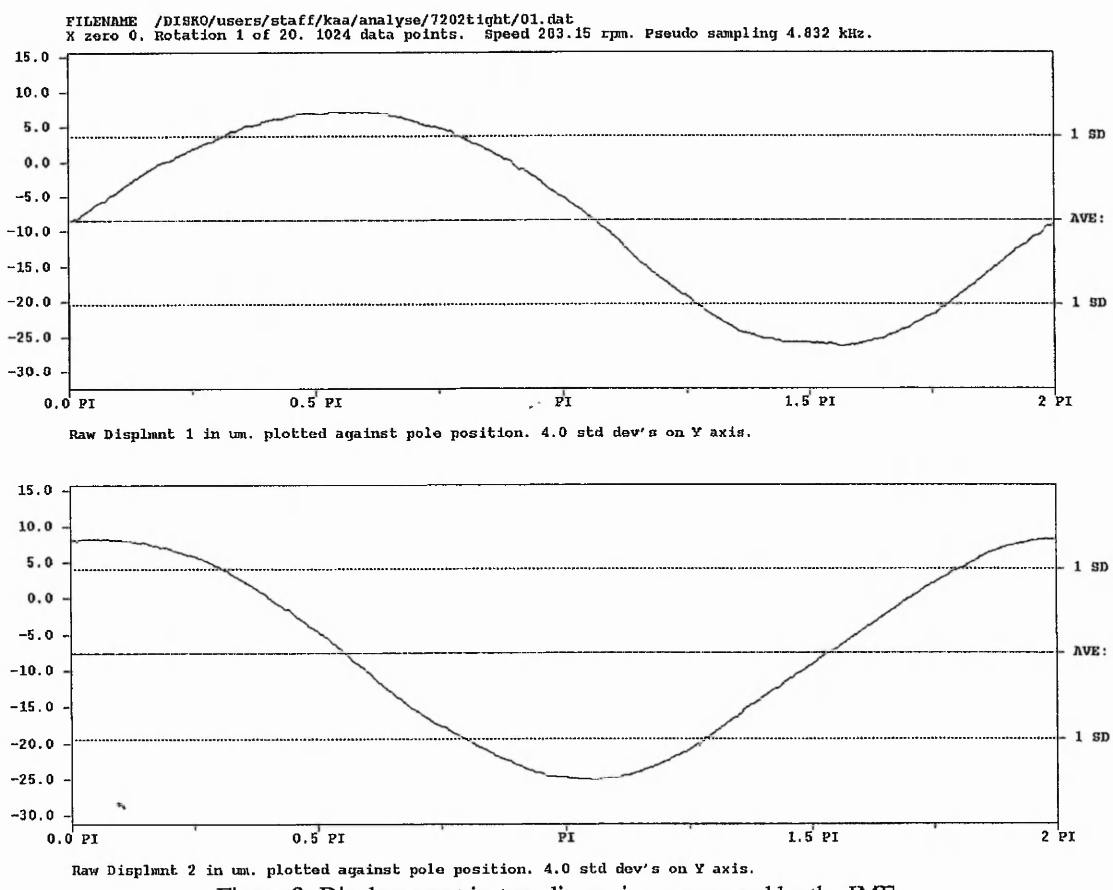


Figure 3. Displacement in two dimensions measured by the IME.

4.2 Motion of shaft centre supported by a normal bearing

The causes of shaft centre motion can be divided between two sources. Those which originate as a result of external systems to which the shaft is fixed such as dynamic imbalance of a flywheel or gear meshing, and those which are due to the relative motion of bearing elements. To date the design of this model has only considered motion due to the bearing elements as this is the primary indicator of bearing condition. For an undamaged bearing the motion of the shaft centre is due to the varying distribution of the rolling elements supporting the inner race as it rotates. The area of the bearing actually supporting the load of the shaft is termed the load zone. If the surfaces of the races and elements were perfect, shaft centre motion would result purely from the changes in loading due to the combination of positions that elements were able to take up in the load zone. The resulting differences in load distribution would produce small, regular and repeating changes in the shaft centre position as each rolling element passed through the maximum load position. In practice the surface irregularities due to manufacturing errors of the bearing components combine with the load distribution variation to produce more complex motion which repeats over a longer period. These irregularities are of two types and have different effects[17]. Surface finish errors, which take the form of a roughening of the contact surfaces of the bearing, can be expected to cause a random motion of the shaft centre. Geometrical errors, which take the form of varying diameter rolling elements and races with a waviness of a few cycles around the circumference, will result in motion of the shaft centre which will be periodic. This motion will repeat over the number of rotations taken for the elements to return to any given initial position relative to the load zone.

The effect of these surface irregularities on the results of conventional vibrational analysis has been investigated[17] and it has been found that the vibration spectrum of a normal bearing can have a similar pattern to that of a bearing with minor damage. This occurs because an irregularly shaped rolling element can excite the machine structure at the same frequency as one with a defect. As the IME senses the motion of the bearing via the shaft directly, these geometric irregularities present a lesser problem as the motion of the shaft due to rounded surface irregularities is different to that expected from bearing defects which are generally angular.

4.3 Effect of the bearing load zone on defect frequencies

Part of work done by McFadden and Smith to explain the frequency spectra from bearings monitored by accelerometers[18,19,20] can also be applied to the IME signal. The feature of their model relevant to the IME sensor is the use of the load zone to redefine the frequencies expected from a defective bearing. The geometrically based defect frequency equations[3] used in the interpretation of vibration spectra from conventional sensors are only directly applicable when the bearing is subject to a unit load distributed evenly about its whole circumference. This condition applies to axially loaded thrust bearings, but not to radially loaded deep groove ball bearings or roller bearings. For a radially loaded bearing the defect frequencies are modulated by the extent of the load zone. Impulses are only generated when the bearing elements strike each other under load. The greater the load, the greater the magnitude of the impulse. The distribution of the load around the circumference of a rolling element bearing is defined approximately by the Stribeck equation[19]:

$$q(\theta) = q_o[1 - (1/2\varepsilon)(1 - \cos\theta]^n \quad (2)$$

q_o is the maximum load intensity, $q(\theta)$ is the load at angle θ, ε is the load distribution factor and n = 3/2 for ball bearings and 10/9 for roller bearing. θ_{max} is the angle at which q becomes equal to zero defining the angular extent of the load zone. The terms q_o, ε and θ_{max} are all functions of the diametral clearance of the bearing and the applied load. For a bearing with positive clearance, $\varepsilon < 0.5$ and $\theta_{max} < \pi/2$, the load distribution is of the type shown in figure 4.

By applying this part of the McFadden and Smith model to the IME it can be predicted that the number of impulses the sensor would detect in each rotation will depend on the angular extent of the load zone. The magnitude of each impulse will depend on the position of the bearing elements between which the impulse was generated in the load zone. The inner race defect frequency only accurately predicts the number of impulses which would be generated in the case where ε = 1.0 and the load zone extends

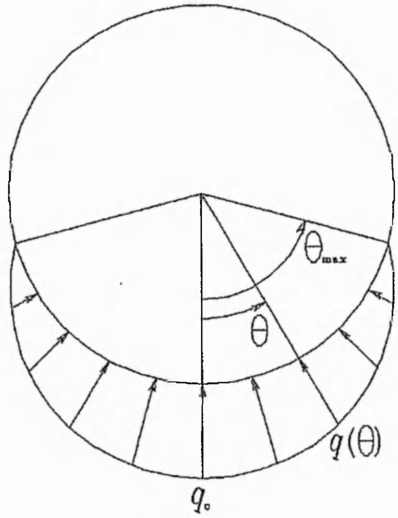


Figure 4. The load distribution in a bearing under radial load.

throughout the complete circumference of the bearing. The number of impulses which it would be expected that the IME would record in other cases from an inner race defect in each rotation is given by:

$$\textit{number of impulses} = \textit{inner race defect frequency} \; x \; \frac{2\,\theta_{max}}{2\pi} \tag{3}$$

To predict the signal the sensor would record for each impulse, a model for the response of the shaft centre motion to the presence of a defect is required.

4.4 Motion of shaft centre due to an inner race defect

In a loaded bearing stress is developed at the points of contact between the races and the rolling elements in the load zone. The motion of the rolling elements leads to fatigue which induces the nucleation and growth of sub surface cracks which, once developed, extend incrementally during successive load cycles. Eventually these cracks join and grow to reach the surface of the race resulting in a loss of material from the race surface known as a spall. Following the initial spalling, wear accumulates rapidly due to the damage caused by the freed particles and to rolling surfaces impacting the spalls.

Figure 5. is a diagram illustrating the model used to predict the response of the motion of the shaft centre to an inner race spall. Defects in other locations can be treated in a similar way. The size of the defect is exaggerated for clarity and the effect of other elements on the depth of penetration is ignored. D_i is the outside diameter of the inner race, D_{re} is the diameter of the rolling element, n_i is the rotational speed of the inner race, y is the depth of penetration of rolling element into spall, x is the width of the spall A - B, V_{re} is the velocity of the rolling element at point of impact, A is the point on the edge of spall about which the rolling element pivots, B is the point of impact between the rolling element and inner race, R is the centre of rotation, $q(\theta)$ is the load acting through point B.

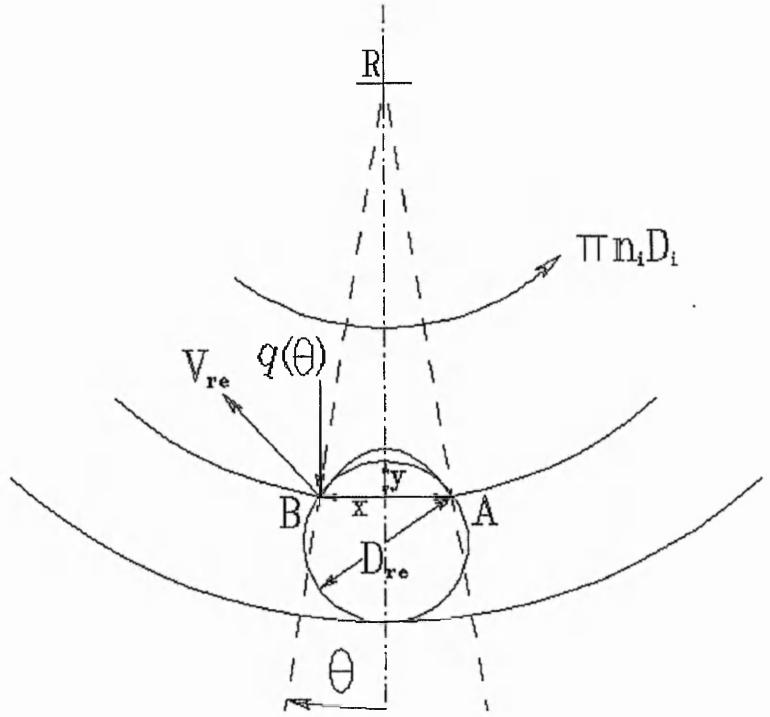


Figure 5. Model for the effect of a inner race spall on shaft centre motion.

The model assumes that from supporting the inner race's undamaged surface a rolling element will rotate around point A into the spall as the inner race pivots down to impact on the element at point B. The force of this impact is determined by the velocity of the rolling element at the point of impact V_{re} and the load acting through that point $q(\Theta)$. The magnitude of $q(\Theta)$ is dependent on the angular position of B in the load zone. It is this impact which generates the stress wave that propagates as an impulsive vibration through the machine structure. The depth of penetration y depends on the width of the spall x and the diameter of the rolling element D_{re} combined with the force of the impact. The motion of the bearing continues and the inner race pivots about point B to return to being properly supported on its undamaged surface by the rolling element. Slipping relative to the outer race may occur at this point.

The motion of the shaft centre which results from the defect is related primarily to the pivoting of the inner race at points A and B due to penetration of the element into the race rather than to the impulsive vibration. As the inner race pivots around point A the shaft centre R will move as the end of the radius R - A. This motion, comprising horizontal and vertical components, is directly related to the depth of penetration since this determines the amount of rotation that occurs before the impact at point B. Similar motion of the centre will occur when the inner race moves around point B. The motion around point A will be faster than that around point B since at point B the load acts against the motion. The resultant effect of this on the IME signal can be predicted as a rapid change away from, then slower change back to, the displacement values due to normal shaft motion.

As the shaft is part of the machine structure, the impulsive vibration due to the impact at point B will affect the motion of the shaft centre. A certain amount of resonant vibration of the shaft will occur which will be visible as fluctuations in displacement immediately following the effect directly due to the defect described above.

5. EXPERIMENTAL RESULTS

The experimental approach adopted to test the model defined for the IME output signal involved initially taking two new SKF NJ202 roller bearings and recording the output signal when each was mounted in the IME test rig. Different speeds were used between 1000 and 3000 rpm and load was changed by the addition of a flywheel for some tests. After a profile had been built up of the signal from each bearing under these conditions, a defect was induced on the inner race of one by electric discharge machining. The defect was 0.2 mm deep, 0.2 mm across and extended across the width of the inner race. Comparison was then made between the data collected previously for this bearing, the bearing without the defect and the behaviour predicted by the model.

Before the defect was induced the IME recorded smooth sinusoidal traces, as in figure 3., from the bearing that was subsequently marked. After the defect was induced the signals recorded became like the example shown in figure 6. in which displacement, on the y axis, is plotted against angular position for a single rotation from 0 to 2π. The most obvious features of this trace are the two small peaks occurring in its second quarter. Also clear is the waveform which appears superimposed on the sinusoid after the two peaks. The immediate supposition that these two peaks are due to the defect is confirmed by their angular separation. For any bearing the number of rolling elements which will pass any point on the inner race in a single rotation is equivalent to the inner race defect frequency. For the NJ202 this is 6.71. The angular distance such a point would therefore travel between elements is 0.937 radians. This is the angular distance between the two peaks in figure 6. This spacing has been consistently observed. The fact that two peaks are seen indicates that the load zone extends over the space of two elements, approximately 1.872 radians.

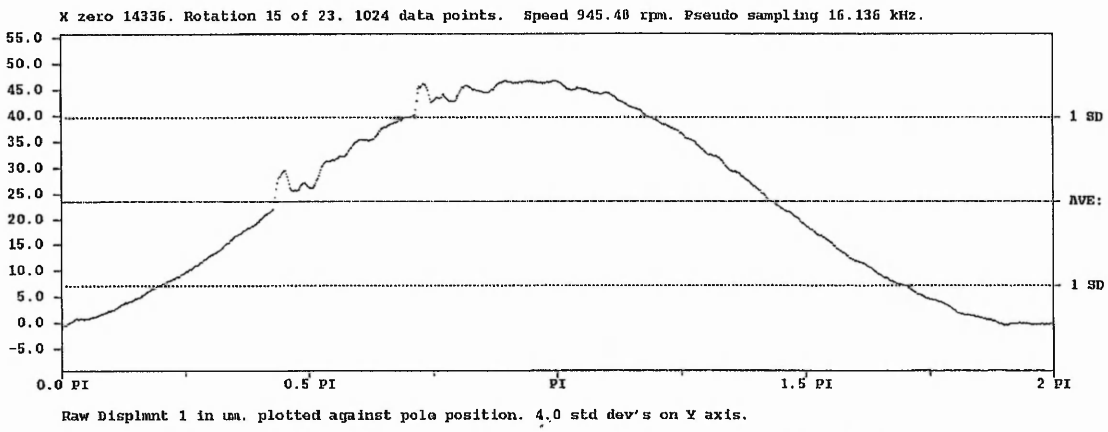


Figure 6. IME displacement signal from bearing with induced defect.

The shape of the peaks is as predicted by the model. The displacement data points from normal shaft motion are clustered so close together they form an almost continuous line. At the base of each peak, equivalent to point A, there is a sudden, sharp jump in the displacement values to a maximum peak, then a slower, but still sharp return to the previous trend. The motion suggested by the displacement values immediately after each peak is that of an oscillatory resonance due to the impulse generated when the defect and the rolling element meet.

The fact that each defect is seen to give a rapid change in displacement means that the presence of such a defect is most easily detected from examination of the acceleration of the shaft centre. This is calculated from the rate of change of the displacement between each angular position on the disk to give the centre's velocity then repeating the process with these velocity values. Figure 7 is a trace taken from the bearing before the defect was induced. It shows the acceleration of the shaft centre from the vertical axis of measurement for a single rotation. Angular position is on the x axis as before. The signal is random as predicted by the model since the main source of shaft movement in this case is the surface roughness of the bearing components. The acceleration in the motion due to the defect stands out against this extremely clearly. Figure 8 shows the shaft centre acceleration from the same rotation as the displacement values shown in figure 6. The vertical scale is 2.5 times that of figure 7 to allow for the peaks due to the defect. The impulsive nature of the signals initial response to the defect is marked, forming the two large positive and negative peaks followed by the exponential ebbing away of the vibration. The first peak is of a greater magnitude than the second, indicating an impact under greater load.

In order to be effective as a condition monitoring technique the use of the IME must allow the detection of incipient faults, as well as the large, fairly obvious one as presented in the preceding example. The change from a random signal to an impulsive one is the key to this.

6. KURTOSIS AND THE IME SIGNAL

Kurtosis is a dimensionless, statistical parameter which was first applied to bearing condition monitoring using accelerometers by Dyer and Stewart[21]. The kurtosis coefficient is calculated as the statistical fourth moment M_4, normalised to the root mean square value[11], equation 4.

$$kurtosis\ coefficient = \frac{M_4}{(RMS)^4} \tag{4}$$

When applied to a signal it gives a measure of the amplitude distribution. For a random signal, such as that from an undamaged bearing, the value of the kurtosis coefficient is 3.0. As bearing damage occurs, the signal becomes more impulsive and the kurtosis coefficient increases to values above 3.0. The parameter is sensitive to the size and distribution of the impulse peaks.

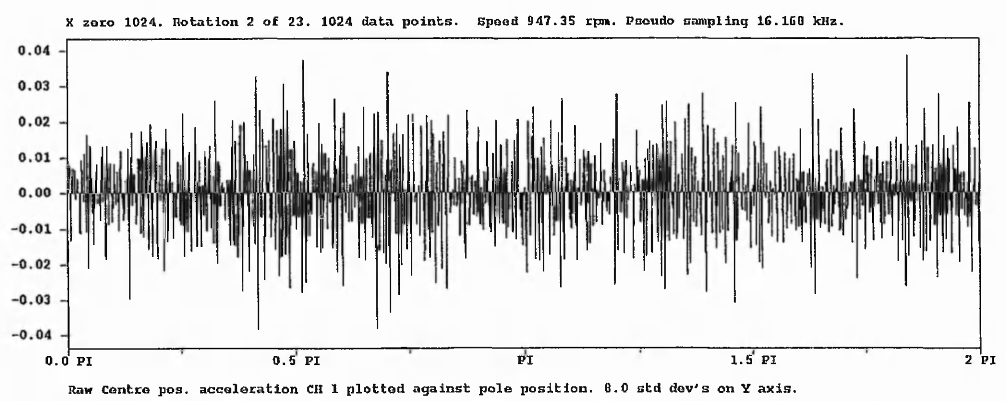


Figure 7. Shaft centre position acceleration recorded by the IME from an undamaged bearing showing random amplitude distribution. Kurtosis coefficient 3.37.

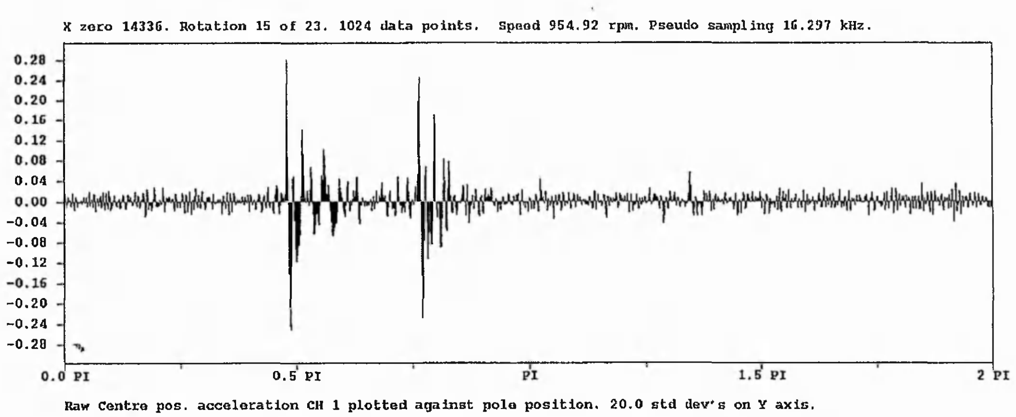


Figure 8. Shaft centre position acceleration from bearing with induced defect showing impulsive nature of the IME signal's response. Kurtosis coefficient 39.39.

Calculating kurtosis coefficients for the IME signal for centre position acceleration from the undamaged bearing gave an average value of 3.011. The average value recorded from the bearing after damage was 32.451. Although further work is needed to prove the ability of kurtosis to accurately determine the onset of bearing damage, from these results it is clear that it can unequivocally discriminate between a good and a damaged bearing from the IME signal. Once impulses are detected in this way, their source can be determined by looking at the angular distance between them. This could also be used to discriminate between impulse sources in the bearing and those from other parts of a machine.

7. CONCLUSIONS

A new technique for the condition monitoring of rolling element bearings has been described which uses a new type of optical sensor, the IME, to directly sense the mechanical motion of a shaft in response to a bearing defect. A simple model has been developed to explain the motion due to an inner race defect which has been shown experimentally to accurately predict the response seen in the sensor signal. The presence of impulses due to defects in the IME signal can be detected by the use of the statistical parameter kurtosis. The location of the defect can be determined from the angular interval between recorded impulses. Further work is required to extend the model for all possible defect locations and the effect of shaft motion related to other machine components.

A feasibility study to determine whether this method can be applied to aero engine bearings will shortly be undertaken. This will involve the analysis of data collected from an aero engine's existing electromagnetic angular position sensor. Further work to extend the IME principle to incorporate special factors relating to the use of electromagnetic sensors is also planned. It is believed that motion in up to six degrees of freedom may be able to be monitored.

8. ACKNOWLEDGEMENTS

The authors gratefully thank Mr. Sanjay Bhubuta of Hewlett-Packard Components Limited for providing the kit encoder parts for the encoder prototype.

This work is funded within the Nottingham Trent University by the Department of Computing.

9. REFERENCES

1. G. P. Miller and S. L. McClymonds, "Maintenance cost avoidance through comprehensive condition monitoring", Int. ASME/IEEE Power Generation Conference, pp. 21 - 25, ASME, Boston, MA, October 1990.

2. W. Scheithe. "Better bearing vibration analysis," *Hydrocarbon Processing*, vol 71 n 7 pp. 57 - 64 1992.

3. J. E. Berry, "How to track rolling element bearing health with vibration signature analysis", *Sound and Vibration*, vol 25, part 11, pp. 24 - 35, November, 1991.

4. J. Mathew and R. J. Alfredson, "The condition monitoring of rolling element bearings using vibration analysis," *Journal of Vibration, Acoustic, Stress and Reliability in Design*, vol 106, pp. 447 - 453, 1984.

5. B. T. Kuhnell and J. S. Stecki, "Correlation of vibration, wear debris analysis and oil analysis in rolling element bearing condition monitoring," Condition Monitoring '84, Pineridge Press Ltd, Swansea, United Kingdom, 1984.

6. T. M. Hunt, "What debris monitor for which application", Condition monitoring '94, Pineridge Press Ltd, Swansea, United Kingdom 1994.

7. M. Serridge, "What makes vibration condition monitoring reliable ?," *Noise and Vibration Worldwide*, pp. 17 - 24, Sept. 1991.

8. A. G. Ray, "Monitoring rolling contact bearings under adverse conditions," Proceedings of the 2nd International Conference on Vibrations in Rotating Machinery, Institution of Mechanical Engineers, pp. 187 - 194, 1 - 4 Sept., 1980.

9. D. B. Szrom, "Low Speed Bearing Analysis," Proceedings of the 10th annual meeting of the Vibration Institute , part 24 - 26, pp. 183 - 188, June 1988.

10. F. Schloss, "Accelerometer noise," *Sound and Vibration*, vol 27, n 3, pp. 22 - 23, 1993.

11. R. A. Collacott, , Vibration Monitoring and Diagnosis , George Godwin Limited, London, 1979.

12. P. A. Orton, Shaft Displacement Measuring System, International patent application number PCT/GB92/00811, World Intellectual Property Organization, International Publication number WO 92/19936, 12 November 1992.

13. A. Ernst, Digital linear and angular metrology , pp. 25-28, verlag moderne industrie, 1989.

14. P. A. Orton, K. Ayandokun, N. Sherkat and P. D. Thomas, "A fibre optic rotary motion sensor for the real-time condition monitoring of rotating machinery," in *Fibre Optic Physical Sensors in Manufacturing and Transportation*, John W. Berthold, Richard O. Claus, Michael A. Marcus, Robert S. Rogowski, Editors, Proc. SPIE 2072, pp. 195 - 207 (1994).

15. M. L. Courtwright, "Instrumented bearings get rolling with ABS," *Machine Design*, vol 64 n 1 pp. 84 - 88, 1992.

16. K. Ayandokun, P. A. Orton, N. Sherkat and P. D. Thomas, "An optical rotary motion sensor for the real-time condition monitoring of rotating machinery," Condition Monitoring '94, Mervin H. Jones, pp. 657 - 668, Pineridge, Swansea, 1994.

17. Y.-T. Su, M.-H. Lin, and M.-S. Lee, "The effects of surface irregularities on roller bearing vibrations," *Journal of Sound and Vibration*, vol 163 n 3 pp. 455 - 466 1993.

18. P. D. McFadden and J. D. Smith, "Vibration monitoring of rolling element bearings by the high frequency resonance technique," *Tribology International*, vol 17 n 1 pp. 3 - 10, 1984.

19. P. D. McFadden and J. D. Smith, "Model for the vibration produced by a single point defect in a rolling element bearing," *Journal of sound and vibration*, vol 96 n 1 pp. 69 - 82, 1984.

20. P. D. McFadden & J. D. Smith, "The vibration produced by multiple point defects in a rolling element bearing," *Journal of Sound and Vibration*, vol 98 n 2 pp. 263 - 273, 1985.

21. D. Dyer and R. M. Stewart, "Detection of rolling element bearing damage by statistical vibration analysis," *Journal of Mechanical Design*, vol 100 pp. 229 - 235, 1978.

Appendix G.

Tracking the development of rolling element bearing faults with the optical incremental motion encoder

K. Ayandokun, P. A. Orton, N. Sherkat and P. D. Thomas,
Proceedings fo the 8th International Congress on Condition Monitoring and Diagnostic Engineering Management (COMADEM '95), vol. 2, pp. 649-655, Kingston, Ontario, Canada, 1995.
ISBN 0-88911-720-9.

TRACKING THE DEVELOPMENT OF ROLLING ELEMENT BEARING FAULTS WITH THE OPTICAL INCREMENTAL MOTION ENCODER

K. Ayandokun, P. A. Orton, N. Sherkat & P. D. Thomas.

The Nottingham Trent University, Department of Computing, Burton Street, Nottingham NG1 4BU, United Kingdom, Tel: +44 (0)115 948 6538, Fax: +44 (0)115 948 6518, Email: kaa@doc.ntu.ac.uk

ABSTRACT

This paper describes the experimental application of a new technique for the condition monitoring of rolling element bearings. Conventional techniques rely predominantly on the measurement of transmitted vibration at the surface of the machine structure. The Incremental Motion Encoder (IME) is a new type of sensor which allows sub-micron measurement of the motion of a shaft relative to its supporting bearing. This has the advantage, as is shown, that this motion is directly related to the condition of the supporting bearing. In this paper we present results from recent experiments which show the IME's ability to detect changes in bearing condition related to the onset of bearing damage due to lubricant contamination.

1. INTRODUCTION

Manufacturing industry is increasing moving toward the development of production based upon computer integrated manufacturing (CIM) systems. An essential requirement for the successful operation of a CIM system is the high reliability of the sophisticated machine tools in the system. To avoid costly unplanned machine stoppages and maintain production standards, the condition of key machine components must be accurately monitored [1]. The aim is to diagnose any potential failure as early as possible allowing maintenance to be scheduled at the most convenient time [2,3]. Rolling element bearings are critical components in rotating machines such as machine tools. Various operating factors such as lubricant contamination, shaft misalignment or improper loading can cause bearing service life to vary significantly from manufacturers estimates, potentially leading to unexpected failures [4]. It has been estimated that only 10 to 20% of bearings reach their design life because of such factors [5]. To prevent unpredictable machine breakdown the monitoring of bearing condition is therefore extremely important.

A number of techniques have been developed to detect the start of failure in rolling element bearings based upon vibration monitoring. These include spike energy, shock pulse, kurtosis, crest factor and the high frequency resonance technique [5,6,7,8]. All of these methods basically rely on the determination of bearing condition from vibration measured at the surface of the machine structure. The vibration measured by a surface mounted sensor is determined by both the bearing condition itself and the mobility of the transmission path between the bearing and the sensor location [9]. This is illustrated by figure 1.

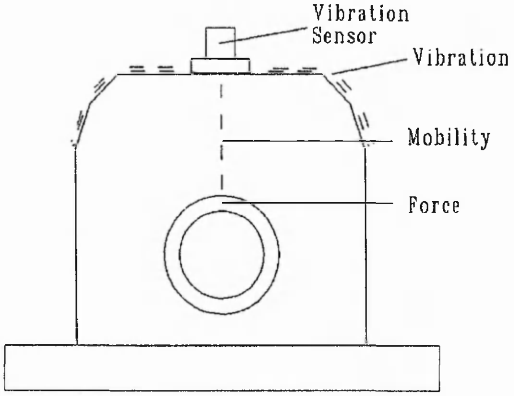


Figure 1. The mobility between sensor and bearing determines the level of vibration measured at the machine surface.

Ideal conditions for vibration monitoring are not found in most complex machines [10]. The vibration signal may suffer attenuation and distortion at component interfaces and the sensor may not be mounted in the location giving the best transmission path. The vibration measured will result from a number of sources and proper diagnosis of individual components may be difficult, especially in the early stages of bearing failure. Low speed machines present particular problems. Low rpm, commonly combined with massive structure, means that little energy is transmitted to the machine casing resulting in low levels of vibration [11]. Measurement of this low level, low frequency vibration is limited by the electromechanical efficiency of the piezoelectrical material used in the accelerometers used for vibration sensing [12]. The noise level produced by an accelerometer and its preamplifier overwhelm the low signal due to the actual vibration.

To overcome the inherent problems in monitoring bearing condition remotely from the machine casing, there are good arguments why relative shaft displacement measurement should be preferable [13]. Principly, relative displacement between a shaft and its mounting is directly related to the condition of the bearing supporting the shaft. Currently shaft displacement monitoring is almost exclusively limited to fluid film journal bearings which are designed to have large clearances. The extremely small clearances in rolling element bearings do not allow a significant amount of shaft motion relative to the bearing housing when measured conventionally. Techniques for monitoring rolling element bearings using eddy current and fibre optic displacement sensors have been developed [14,15,16]. However these rely on measuring deflections of the bearing outer race from within the bearing housing, which requires substantial modification.

The Incremental Motion Encoder (IME) extends the use of relative displacement measurement as a condition monitoring technique to rolling element bearings. It has the advantage of sensing the mechanical motion caused by the presence of a defect directly, allowing monitoring of low speed machines. It can also directly monitor the state of single bearings in complex machines.

2. THE INCREMENTAL MOTION ENCODER

The IME is based on the combination of a well understood principle used in high precision optical shaft encoders and a method for deriving a high degree of angular accuracy from a relatively low resolution encoder disk [17]. When using any optical shaft encoder, shaft centre movement is recognised as a possible source of error in the angular position measurement. High precision encoders are highly susceptible to this source of error and make a correction for the shaft axial eccentricity prior to their calculation of angular position [18]. This correction is achieved by determining the angular difference indicated between two, or four, diametrically opposed points on the grating disk. The measure of shaft centre position derived in this way is treated as an error and is discarded after correction has been made for it. The IME is constructed with four read heads mounted on the machine structure at 90° around a grating disk mounted on the shaft. This arrangement is shown schematically in figure 2.

A computer is used to determine the angular positions

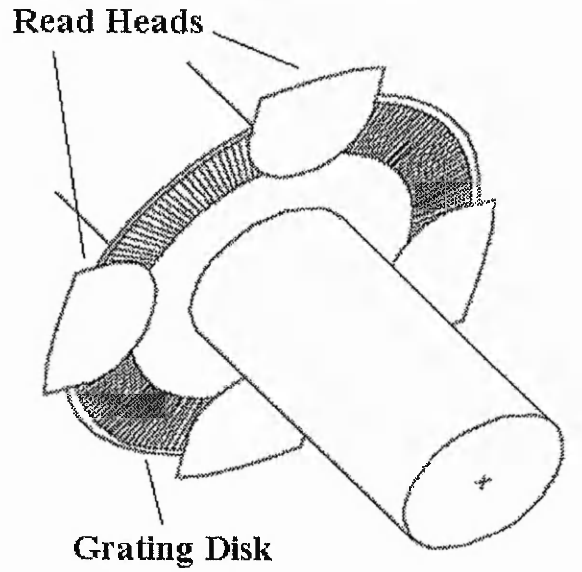


Figure 2. The arrangement of read heads around a grating disk forming the basis of an IME sensor.

measured at each read head from their square wave output signals [19]. The shaft centre position is derived trigonometrically from these four readings. Each diametrically opposed read head pair provides the shaft centre position in one axis. The computer therefore derives a two dimensional picture of the motion of the shaft centre in the plane of the sensor disk. Each read head pair makes one position reading for each line on the grating disk.

The IME is also able to measure angular position and velocity extremely accurately since it records to high resolution the time taken by the disk to rotate through the angular difference between grating lines. By precisely relating each angular velocity reading to an angular position on the disk's rotation, shaft torsional vibration may potentially be measured more accurately than with conventional encoders.

Sensors based on the IME principle can be implemented using a number of different technologies capable of giving a square wave output from a rotating marked surface. As well as opto-electrical read heads and a grating disk, optical sensors using a reflective drum on the shaft could also be used. For harsher environments, where contamination would limit the effectiveness of an optical sensor, the same principle may be applied using electromagnetic sensing elements. In many cases rotating machines have a sensor for measuring angular motion which produces the type of signal to which the IME principle can be applied with the addition of extra read heads. The development of bearings with integrated encoder sensors [20] offers the possibility, combined with the IME principle, for self monitoring bearings able to sense a number of machine parameters.

Within this project two experimental IME sensors have been constructed using optical devices. One has been built as an entirely optical device using fibre optic technology [19]. The other uses standard commercially available kit encoder parts from Hewlett Packard, four opto-electrical read heads and a 1024 line grating disk. This device has the higher resolution of the two and is being used for current experimental work.

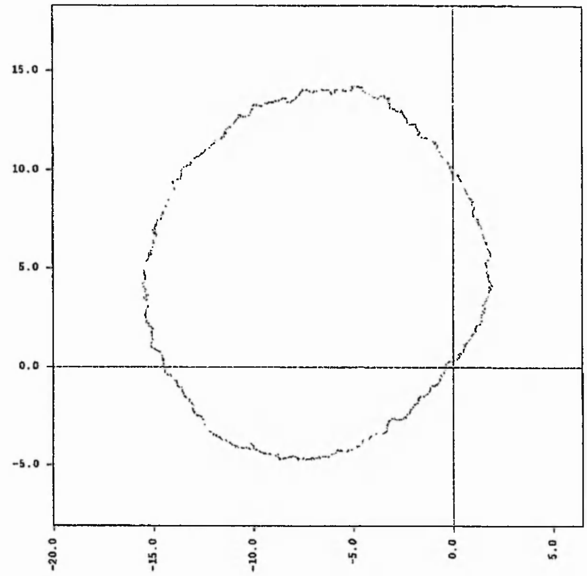


Figure 3. Combined two dimensional shaft centre displacement measured by the IME from an NJ202 roller bearing rotating at 790 rpm. X and Y values in microns.

3. THE IME SIGNAL AND BEARING CONDITION.

The signal from each read head pair of the IME consists of a displacement value, calculated in microns, taken at each line on the grating disk. Each value is therefore taken at an equal interval of angle. The start of data collection is synchronised to one point on the disk. This allows correlation of values recorded on each axis to form a trace that represents the motion of the shaft centre over successive rotations. Figure 3 shows an example of displacement measured by the IME from an NJ202 roller bearing. The two axes of measurement are shown plotted against each other. The intersecting zero lines represent the diameters joining the read head pairs. Each axis of measurement produces an approximately sinusoidal trace which, plotted together produce, the circular pattern shown. This represents the motion of the disk center as it moves around the actual axis of shaft rotation [21].

The irregular outline of the trace is due to the combined effect of irregularities on the bearing element surfaces and the variation in distribution of the elements in the load zone of the bearing. We have developed a model which describes the way the presence of a single inner race defect can be detected from the IME signal using the statistical parameter kurtosis [21]. It has been shown experimentally that the random motion resulting from surface irregularities gives way to impulsive motion when a large defect is created.

To be effective as a condition monitoring device the IME has to be capable of not only detecting the presence of defects after they have formed but also to accurately predict defect formation at the earliest point possible. To this end experimental work has been carried out to show that the IME is capable of detecting the breakdown of the lubricant contact between bearing elements caused by lubricant contamination.

4. THE IMPORTANCE OF LUBRICATION CONTACTS IN ROLLING ELEMENT BEARINGS.

The performance and life of rolling element bearings is strongly dependent on the performance of the individual lubrication contacts between balls and races [22]. The pressure developed in the contact zone between moving elements raises lubricant viscosity to the point where its behaviour approaches that of a solid [23]. A thin elastohydrodynamic lubricant (EHL) film is formed that physically separates the would be contacting surfaces inhibiting metal to metal contact. Under ideal conditions the surface rough spots or asperities of the elements would puncture into but not through this EHL film. In actuality contact between the peaks of the ball and race surfaces does occur, as shown schematically in figure 4. It is these contacts that cause the fatigue loading of the element surfaces which eventually result in the formation of spalls as the surfaces breakdown. The thickness and continued presence of the EHL film has a profound effect on bearing life [23]. Detection of the failure of the EHL film is therefore capable of acting as a valuable early warning of future bearing failure.

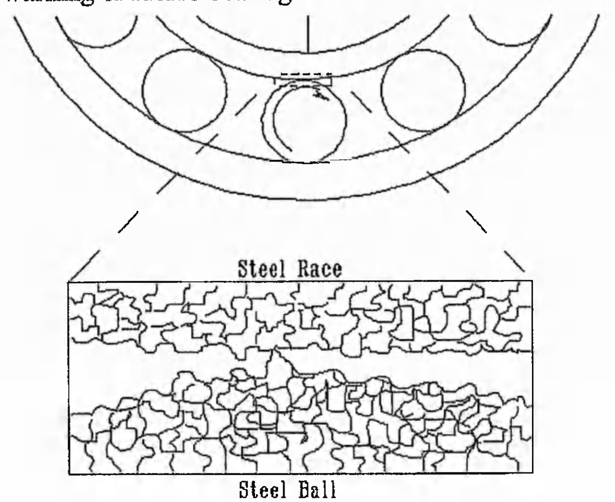


Figure 4. Asperity contact between ball and inner race otherwise separated by an EHL film.

Debris particle contamination of the lubricant is a factor that frequently leads to a reduction in bearing life [24]. As illustrated in figure 5, when lubricant born debris are entrained between the rolling elements and races the EHL film cannot be formed and stress contacts are generated. We have conducted a set of experiments which have shown the effectiveness of the IME at detecting lubricant contamination.

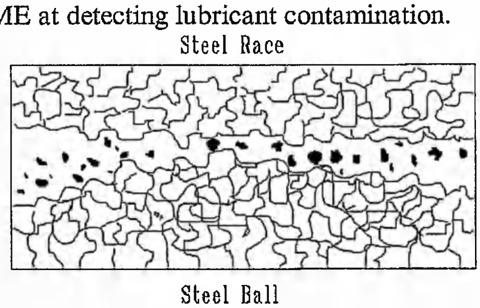


Figure 5. Debris contamination in the EHL film between ball and race causing stress contacts.

5. EXPERIMENTAL RESULTS

A rig has been constructed to allow testing of the experimental IME sensor providing the facility for bearings to be interchanged on the shaft supporting the encoder [19]. This rig was used for the contamination experiment with new NJ202 roller bearing. Shell Avalina R1 grease was used as the lubricant in the experiment.

The bearing was first run continuously under nominal load for 150 hours. At the end of this period the IME was used to collect a series of three displacement datasets. Each dataset consisted of the displacement values from 25 rotations of the shaft collected every 10 minutes over a 14 hour period. These were used to establish baseline values for average displacement, shaft centre position acceleration and kurtosis prior to the contamination of the lubricant.

The debris used to contaminate the lubricant came from a spark erosion machine. These were fragments produced during the normal operation of this machine which were filtered out of the oil bath in which the sparks are generated. The rating of the filter used gave a maximum particle size of 5μm, no estimation could be made for the minimum size of particle present. The particles were a random combination of steel, brass and various ceramics. Approximately $10mm^3$ of these particles were mixed into $1cm^3$ of grease. From this contaminated sample $4mm^3$ of grease was used to contaminate the bearing for the first experiment.

The displacement traces collected during the period prior to contamination of the lubricant were generally similar to that shown and described in figure 3. Immediately after the contamination the traces showed a marked change. Figure 6 shows an example of the displacement motion recorded by the IME as the bearing elements rolled over the particles in the lubricant.

Rather than the smooth rolling, sliding contact enabled by an EHL film, the shaft centre motion is dominated by the erratic, impulsive displacement which occurs as particles effectively act as fulcra between the rolling elements.

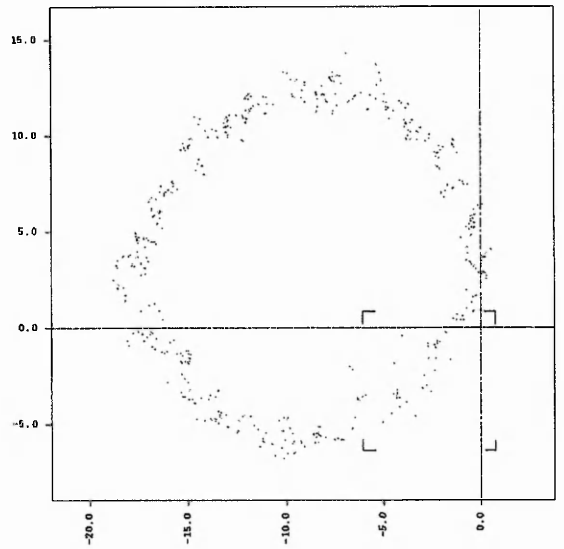


Figure 6. Shaft centre displacement recorded by the IME from an NJ202 roller bearing after contamination of its lubricant. *X* and *Y* values in microns. Speed 749 rpm.

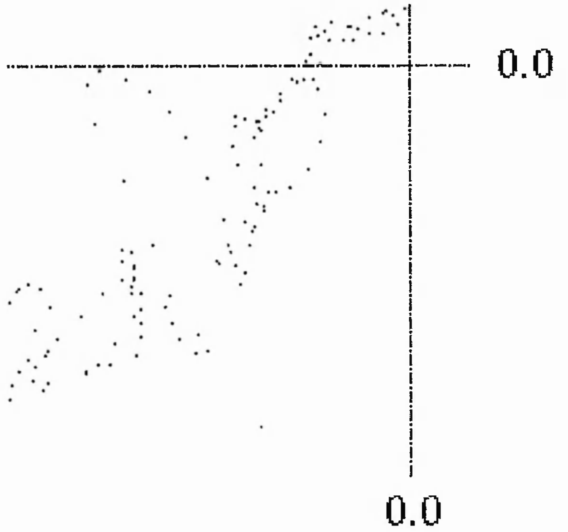


Figure 7. Enlargement of outlined area in previous figure. The motion of the shaft centre in response to particles of lubricant contaminant is a series of complex loops. Scale in microns.

Initially the pattern made by the points in the figure 6 might appear random. Figure 7 shows an enlargement of the small area of figure 6 outlined by the four corner marks. It can be seen that some successive points form loops with smooth curves. This motion of the shaft centre is consistent with that which would be predicted using part of the model explaining the IME's response to an inner race defect [21]. In both cases the shaft centre moves up and down in one axis and from side to side in the other as the ball-race contact makes a transition from smooth motion to impulsively riding over an obstruction.

The shaft centre acceleration, calculated as a simple double differentiation of the displacement values recorded at each angular position, showed an immediate and substantial increase after the lubricant had been contaminated. The baseline value for peak acceleration was 0.050 $\mu m/rad^2$ Immediately after lubricant contamination the peak acceleration value reached 0.350 $\mu m/rad$. This change is a clear and readily interpreted signal of a fault condition. Figure 8 compares the shaft centre acceleration from two separate rotations of the shaft. The one at the top of the figure is typical of the signal recorded in the period before the lubricant was contaminated. The second was recorded immediately after the contamination. The order of magnitude increase described is clearly visible, however it must also be noted that both signals display a random amplitude distribution with kurtosis coefficients of 3.0 in both cases. This indicates that no single repeating defect has been formed by the effect of the contaminating particles.

The bearing was monitored over a further 48 hour period after the lubricant had been contaminated. A rapid decrease in the impulsive motion due to the contaminating particles was observed over the first 3 hours. Over this period it appears that the larger particles were either crushed or moved clear of the element contacts. This would seem to indicate that the bulk of the contaminating particles were friable ceramics. Further work is required to quantify the constituents of the contaminant. At the end of 48 hours the kurtosis values recorded over the last hour averaged 3.7. This indicated that the centre position acceleration signal had become more slightly more impulsive than those recorded at the time of contamination. This would be consistent with the expected rise in surface roughness due to the abrasive nature of the contaminant. The still relatively low kurtosis value would seem to indicate that over the limited period of the study no large defect was formed. Again this was the expected result.

The contamination experiment was repeated three times and showed results consistent with the first experiment.

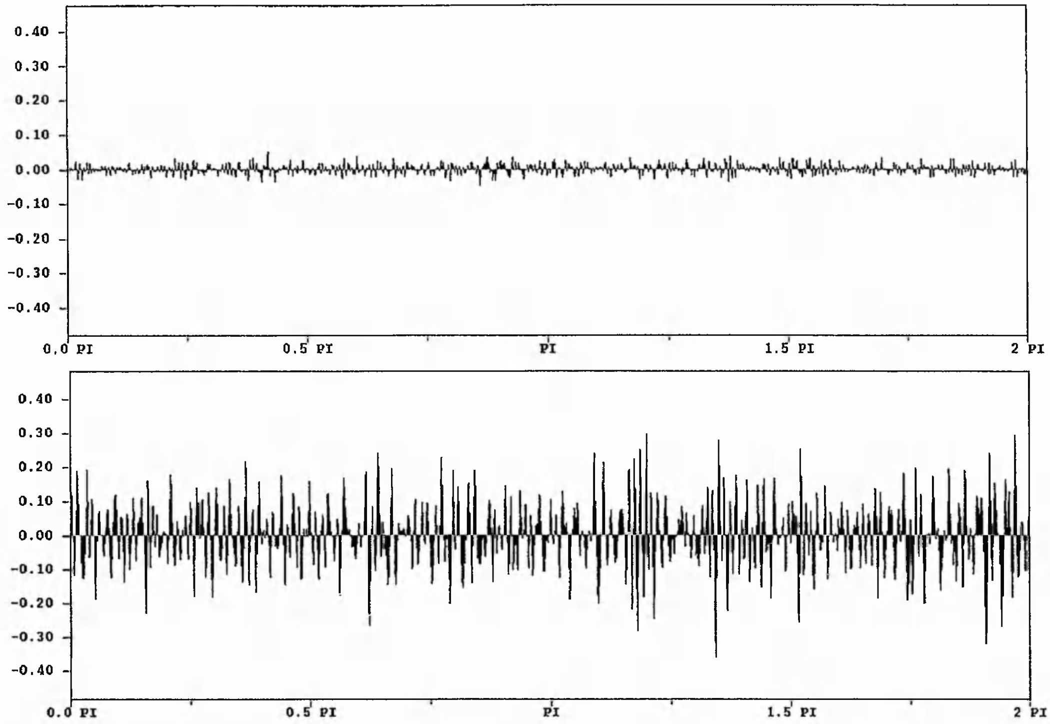


Figure 8. Shaft centre position acceleration before, top, and after, bottom, lubricant contamination. X axis is in radians, 0 - 2PI covering one rotation in each case. Y axis units are $\mu m/rad^2$.

6. CONCLUSIONS

The point at which the lubricant film in a rolling element bearing breaks down is a good early warning signal of potential bearing failure. Lubricant contamination is also a common cause of bearing failure. The incremental motion encoder has demonstrated experimentally that it has the ability to detect the point at which bearing lubricant contamination occurs. Changes in both the shaft centre motion patterns and acceleration levels have been clearly recorded which can be classified as indicative of a fault condition. Further work is currently in progress to determine how well the IME is able to follow the progress of the slight defects induced by the contamination.

The potential use of the IME method for monitoring the bearings of a high speed engine is currently being investigated.

7. ACKNOWLEDGEMENTS

The authors gratefully thank Hewlett-Packard Components Limited for providing the kit encoder parts for the experimental IME device.

This work is funded within the Nottingham Trent University by the Department of Computing and by a grant from the Japanese Electro-Mechanic Technology Advancing Foundation (EMTAF).

8. REFERENCES

1. A. Davies, "The intelligent machine", *Manufacturing Engineer*, Aug., 182-185, (1994).

2. K. F. Martin, "A review by discussion of condition monitoring and fault diagnosis in machine tools", *International Journal of Machine Tools and Manufacturing*, **34**(4), 527-551 (1994).

3. G. P. Miller and S. L. McClymonds, "Maintenance cost avoidance through comprehensive condition monitoring", *Int. ASME/IEEE Power*

Generation Conference, ASME, Boston, MA, 21-25th, October (1994).

4. W. Scheithe. "Better bearing vibration analysis," *Hydrocarbon Processing*, **71**(7), 57-64 (1992).

5. J. E. Berry, "How to track rolling element bearing health with vibration signature analysis", *Sound and Vibration*, **(25)**11, 24-35 (1991).

6. J. Mathew and R. J. Alfredson, "The condition monitoring of rolling element bearings using vibration analysis," *Journal of Vibration, Acoustic, Stress and Reliability in Design*, **106**, 447-453 (1984).

7. D. Dyer and R. M. Stewart, "Detection of rolling element bearing damage by statistical vibration analysis", *J. Mech. Design*, **100**, 229-235 (1978).

8. P. D. McFadden and J. D. Smith, "Vibration monitoring of rolling element bearings by the high frequency resonance technique", *Tribology Int.* **17**(1), 3-10 (1984).

9. M. Serridge, "What makes vibration condition monitoring reliable ?," *Noise and Vibration Worldwide*, Sept. (1991).

10. A. G. Ray, "Monitoring rolling contact bearings under adverse conditions," *Proceedings of the 2nd International Conference on Vibrations in Rotating Machinery*, Institution of Mechanical Engineers, 187-194, 1 - 4 Sept. (1980).

11. D. B. Szrom, "Low Speed Bearing Analysis," *Proceedings of the 10th annual meeting of the Vibration Institute*, part 24 - 26, 183-188, June (1988).

12. F. Schloss, "Accelerometer noise," *Sound and Vibration*, **27**(3), 22-23 (1993).

13. R. A. Collacott, , Vibration Monitoring and Diagnosis, George Godwin Limited, London (1979).

14. P. Y. Kim, "A review of rolling element health monitoring(III): preliminary test results on eddy current proximity transducer technique", *Institute of Mechanical Engineers Vibrations in rotating machinery conference*, C311/84, **5**, 119-125, (1984).

15. G. J. Phillips, "Rotating machinery rolling element bearing performance monitoring using the fibre optic method", *Reliability stress analysis and failure prevention methods in mechanical design*, published by ASME, 37-43, Aug. (1980).

16. "REBAM® an effective way to monitor and analyze rolling element bearing condition", Applications Note AN047-02, Bentley Nevada Co., Sept. (1992).

17. P. A. Orton, Shaft Displacement Measuring System, International patent application number PCT/GB92/00811, World Intellectual Property Organization, International Publication number WO 92/19936, 12 November (1992).

18. A. Ernst, Digital linear and angular metrology, 25-28, verlag moderne industrie, (1989).

19. P. A. Orton, K. Ayandokun, N. Sherkat and P. D. Thomas, "A fibre optic rotary motion sensor for the real-time condition monitoring of rotating machinery," in *Fibre Optic Physical Sensors in Manufacturing and Transportation*, John W. Berthold, Richard O. Claus, Michael A. Marcus, Robert S. Rogowski, Editors, Proc. SPIE 2072, pp. 195 - 207 (1994).

20. M. L. Courtwright, "Instrumented bearings get rolling with ABS," *Machine Design*, **64**(1) 84-88 (1992).

21. K. Ayandokun, P. A. Orton, N. Sherkat & P. D. Thomas, "Detecting rolling element bearing defects with the optical incremental motion encoder", in *Industrial Optical Sensors for Metrology*, Kevin G. Harding, H. Philip Stahl, Editors, Proc. SPIE 2349, 53-64 (1995).

22. A. Savan, H. J. Boving, F. Fluehmann & H. E. Hintermann, "Increased performance of bearings using TiC-coated balls", *Journal de physique IV*, **4**(7) supplement au journal de physique III, **3**, Nov. (1993).

23. W. A. Glaeser, R. C. Erickson, K. F. Dufrane & J. W. Kannel, "Tribology: The science of combatting wear", *Journal of the Society of Tribologists and Lubrication Engineers*, **50**(5), 381-384, (1994).

24. E. Ioannides, E. Beghini, B. Jacobson, G. Bergling & J. Wuttkowski, "Cleanliness and its importance to bearing performance", *Journal of the Society of Tribologists and Lubrication Engineers*, **49**(9), 657-663, (1993).

Appendix H.

The Incremental Motion Encoder: a sensor for the integrated condition monitoring of rolling element bearings in machine tools

K. Ayandokun, P. A. Orton, N. Sherkat, P. D. Thomas & J. F. Poliakoff, Machine tool. In-line. and Robot sensors and controls, SPIE proceeding vol. 2595, 1996.

The Incremental Motion Encoder: a sensor for the integrated condition monitoring of rolling element bearings in machine tools.

K. Ayandokun, P. A. Orton, N. Sherkat, P. D. Thomas & J. F. Poliakoff.

The Nottingham Trent University, Department of Computing, Burton Street, Nottingham NG1 4BU, United Kingdom,
Tel: +44 (0)115 948 6538, Fax: +44 (0)115 948 6518, Email: kaa@doc.ntu.ac.uk

ABSTRACT

Modern industry increasingly demands that machine tools operate continuously as parts of computer integrated manufacturing cells. Accurate monitoring of machine condition is the key to predicting failures which would result in quality defects or costly unplanned production stoppages. This paper presents research into the use of a novel rotary motion sensor for the condition monitoring of rolling element bearings. This sensor, the Incremental Motion Encoder (IME), is based upon a patented development of the optical encoder commonly used in machine tools for sensing angular position and rotational speed. The IME combines these measurements with that of shaft centre position in two dimensions. This motion of the shaft centre is directly related to the condition of the bearings supporting the shaft. Experimental results showing the ability of the sensor to distinguish between bearing defects and external sources of vibration and to measure shaft loading are presented to illustrate the IME principle. Currently machine tool condition is most often measured by external sensors, such as accelerometers or acoustic emission transducers, which are not part of the machine itself. The IME is ideally suited to being designed into a machine tool so as to integrate a condition monitoring facility into the computer control of the machine. The paper concludes by describing the current technology which will allow sensors based on the IME principle to be integrated directly into rolling element bearings for this purpose.

Keywords: incremental motion encoder, machine tools, predictive maintenance, condition monitoring, rolling element bearings, kurtosis, smart bearings

1. INTRODUCTION

Maintenance management has become increasingly important to industry with the growth of computer integrated manufacturing (CIM) within automated factories. The high capital cost of the sophisticated machine tools in CIM systems requires them to be operated at peak efficiency to generate maximum return. Effective maintenance is essential to guarantee the high reliability that successful operation of a CIM system requires. Using measurements of actual machine condition to plan maintenance schedules has become accepted as the best strategy to avoid expensive unplanned stoppages and maintain production standards.

Modern machine tools present a particularly challenging task for the maintenance engineer. Typically they are complex assemblies of electronic and mechanical components combined with a variety of drive, lubrication and coolant systems. Ideally specific condition monitoring sensors would be designed into the machine and the machine controller would incorporate a diagnostic system capable of detecting potential failure in any of the machine components. Combined with sensors for detecting unsafe or fault conditions in normal machine operation, such a system would provide the individual machine with the ability to monitor its own overall condition in an intelligent, semi-autonomous fashion. Such comprehensive condition monitoring facilities are not generally available commercially despite research in this area[1,2]. In the absence of such systems, sensors for separate components, such as accelerometers for spindle bearings, are used. The more sensors, and particularly sensor types, employed, the harder the monitoring system is to calibrate and maintain. As a number of different faults can be detected from a properly collected and analysed vibration signal this parameter is often used as the main basis of a rotating equipment condition monitoring program[3].

An alternative approach to specific condition monitoring sensors for individual machine tool components is the use of feedback from process control sensors to determine machine condition. Faults in machine operation can be detected by monitoring the feedback signals to, for example, a servo motor current controller, during machine operation and comparing the signal to that obtained when the machine was known to be in a healthy state[4]. For machine tools which contain rotating parts, a rotary motion sensor is an integral part of the control system and is used for speed regulation and position control. In machine tools the rotary motion sensor most often used is the optical shaft encoder (OSE). Shaft speed measurements made during drilling and milling operations for process control have been shown to be useful for condition monitoring[2,5].

The changes in speed and torque during cutting operations can be used to detect abnormal conditions such as excessively worn or broken tools.

2. THE INCREMENTAL MOTION ENCODER

Recent research into the automatic monitoring and control of precision grinding machines made extensive use of velocity and torque measurements made by optical shaft encoders[6]. Highly accurate angular velocity profiles were obtained for the grinding wheel during machine operation using a reciprocal timing method. Drive models could then be obtained online using least squares parameter estimation. Analysis of these drive models enabled the estimation of important drive system characteristics such as motor speed control performance and drive belt tension. The use the OSE in this work led to the conclusion, later patented, that a new type of sensor could be developed based upon the OSE principle combined with a novel method of angular position estimation[7]. This new sensor, the incremental motion encoder (IME), measures shaft centre position in the plane of the encoder disk as well as the usual angular motion parameters. Our current research aims at developing the potential of this device as a condition monitoring sensor with an emphasis on rolling element bearings.

When using any optical shaft encoder, shaft centre movement is recognised as a possible source of error in the angular position measurement. High resolution encoders of the type used for accurate positioning are particularly susceptible to this source of error and make a correction for shaft axial eccentricity prior to their calculation of angular position[8]. This correction is achieved by determining the angular difference indicated between two, or four, diametrically opposed points on the grating disk. The measure of shaft centre position derived in this way is treated as an error and is discarded after the correction is made. An IME device is constructed with four encoder read heads mounted on the machine structure or bearing housing at 90° around a grating disk mounted on the shaft. This configuration is illustrated in figure 1 which shows the experimental IME device constructed to allow testing and development of the basic IME theory. The sensor is mounted on a supporting pillar which allows the bearing at that end of the shaft to be easily interchanged. This allows the IME's response to various bearing conditions could be experimentally assessed.

This experimental IME device uses photoelectric read heads and a 1024 line encoder disk available as standard kit encoder parts of the type commonly used for industrial motion control. Treated conventionally the square wave signals from read heads of this type would not provide significant angular resolution for the shaft centre position to be determined. However, by using a high frequency clock oscillator, the computer system to which the IME is interfaced can subdivide the interval

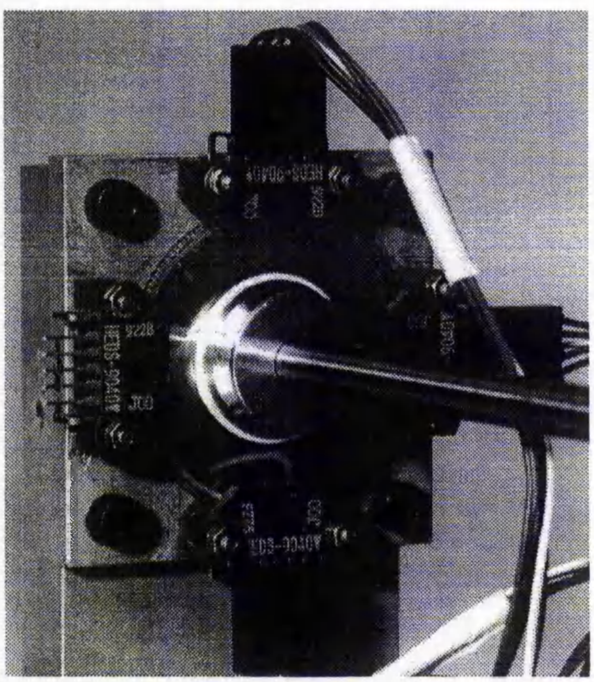


Figure 1. Experimental incremental motion encoder showing the configuration of encoder disk and read heads used to sense the motion of the shaft centre. Shaft dia. 8mm.

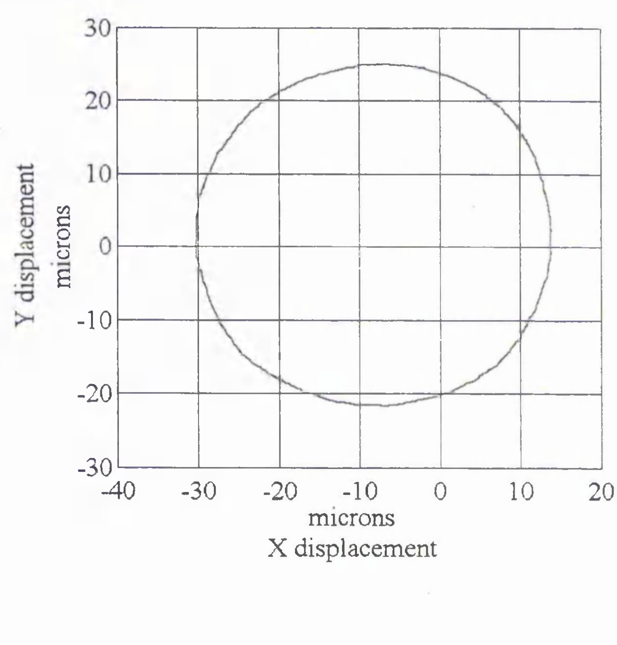


Figure 2. Shaft centre motion measured by the IME from an NJ202 bearing rotating at 806 rpm. Displacement values averaged over 24 rotations.

between grating lines to a degree limited only by the frequency of the clock oscillator[9]. The difference in timings recorded between grating lines by each pair of diametrically opposed read heads is a measure of their angular separation. This separation changes with the motion of the shaft centre along a line perpendicular to the diameter joining the read head pair. The two read head pairs allow the motion of the shaft centre in two dimensions to be measured. An x and y value is recorded for each line on the grating disk. The axes of measurement produce approximately sinusoidal traces that when plotted together produce the circular pattern shown in figure 2. The data plotted in this way resembles the shaft orbital motion plots obtained using eddy current proximity sensors, but which actually represent the motion of the shaft centre rather than the gap between the shaft surface and the machine structure. A single synchronising point on the grating disk allows each displacement measurement to be made at a specific angular position, corresponding to a grating line, on each rotation. This allows the displacement at the same point on successive rotations to be precisely compared over any period of time. This is a unique feature of the data produced by IME sensor.

3. MONITORING BEARING CONDITION WITH THE IME

The rolling element bearings (REB) that support shafts and spindles in machine tools are critical components whose failure can occur unexpectedly without effective condition monitoring. The majority of commercial systems for the on-line monitoring of machine tool bearings are based on analysis of vibration measured at the machine surface by a sensor such as an accelerometer. Several signal processing methods have been developed to determine bearing condition from the vibration signal. These include the use of such parameters as spike energy, shock pulse, kurtosis and crest factor which are measures the magnitude of the signal and the high frequency resonance technique which examines the frequency content of the signal[10,11,12,13]. The problem inherent in all these methods is that the vibration signal at the machine surface is dependent not only on the condition of an individual bearing, but also on other sources of vibration in the machine and the transmission path of the signal through the machine structure to the sensor. The signal will be attenuated as it is transmitted particularly at the component interfaces within the machine. As a result proper diagnosis of an individual bearing fault from the vibration signal can be difficult, particularly in the early stages of bearing failure when vibrational energy levels are low[14].

The measurement of the motion of a shaft relative to its mounting has been suggested as an alternative method for REB monitoring as this motion is directly related to the condition of the bearing supporting the shaft[15]. Although widely applied to fluid film journal bearings that have large clearances, this method has been impractical to apply in the same way to REB which have extremely small clearances by comparison. The advantages of directly monitoring motion related to shaft displacement has been shown by the work of Harker and Hansen[16] who successful used eddy current displacement sensors within the bearing housing to determine bearing condition from deflections of the bearing outer race. Despite becoming a commercially available system[17], this technique cannot be applied without significant modification of a machine's structure which prevents its widespread application.

The development of the IME has created the potential for the advantages of shaft relative displacement measurement to be extended to the monitoring of REB in machine tools and other industrial plant machinery. Experimental results have shown that the sub-micron accuracy of the displacement measurements made by the IME allow the motion of the shaft to be related to the positions of individual elements within the supporting bearing. A model has been developed, based on this work and established bearing condition monitoring theory, that relates the motion of the shaft centre recorded by the IME to the surface texture of the bearing elements[18]. The random roughness of normal bearing surfaces produces shaft centre acceleration with a random amplitude distribution. This distribution will change, becoming more impulsive, with the development of a spall-type defect on one of the bearing surfaces. Changes in the amplitude distribution of the acceleration signal can be measured using the statistical parameter kurtosis[12]. A signal with a totally random amplitude distribution has a kurtosis coefficient of 3.0 which increases as the signal becomes more impulsive. The kurtosis coefficient derived from the shaft centre acceleration signal of a normal NF202 roller bearing is typically in the range 2.8 to 3.5. Figure 3 shows the shaft centre acceleration signal from a single rotation of a bearing of the same type with a defect 0.2mm wide and 0.2mm deep etched across the width of its outer race. The kurtosis coefficient of this signal is 10.15.

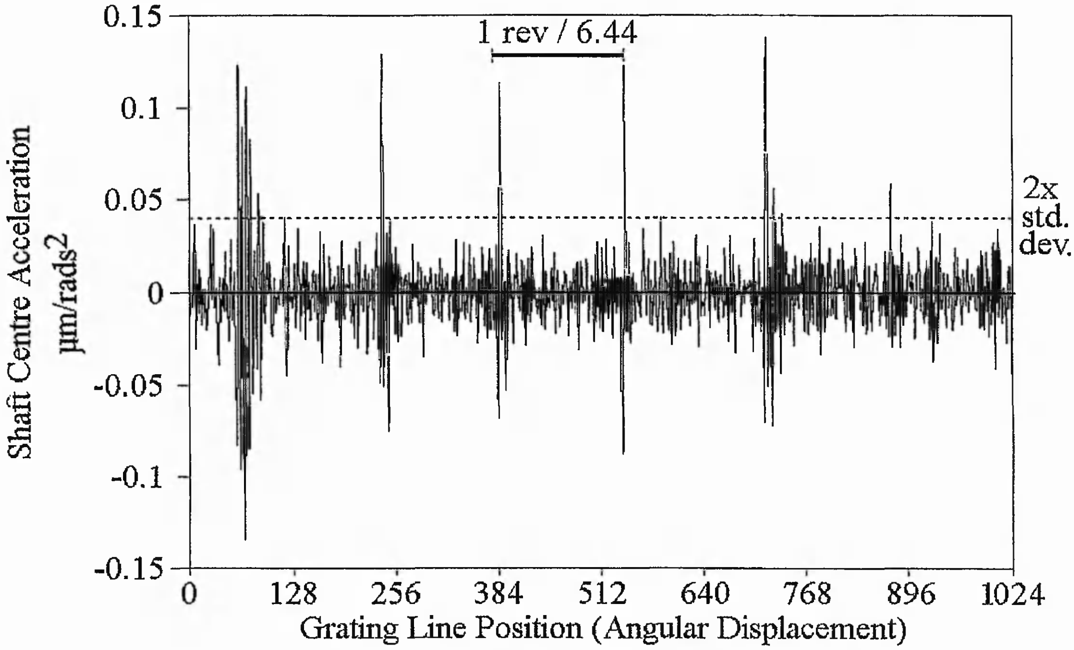


Figure 3. The impulsive shaft centre acceleration signal from a bearing with an outer race defect. The angular interval measured between impulses indicates their source.

By monitoring the kurtosis level of the shaft centre acceleration the presence of impulses in the signal can be detected, but the source from which they originate cannot be determined. Discriminating between defects on the rollers and the inner or outer races is not generally useful since in the event of damage on one part the whole bearing will be replaced. However it is important to be able to differentiate between impulsive signals that originate in the bearing supporting the shaft and those which might arise from other sources in a machine, such as gear meshing. The nature of the IME signal, collected at equal angular intervals, allows the direct measurement of the angular distance between impulses in the signal. This is an advantage over conventional vibration sensors which need a separate angular motion sensor to relate their signals to the rotation of the shaft. The geometry of a bearing's races and rolling elements allows the frequency of the impulsive signal due to any spall-type defect to be related to shaft speed using a set of commonly applied bearing defect equations[10]. Conventionally these equations express the defect frequency in Hertz. By omitting the term for shaft speed from these equations, the frequencies for defects on any of a bearing's components can be calculated in terms of impulses per rotation. Equation 1 gives the frequency, f, per revolution which results from an outer race defect. D = bearing pitch diameter, d = rolling element diameter, n = number of balls or rollers, θ = rolling element contact angle.

$$f = \frac{n}{2}\left(1 - \frac{d}{D}\cos\theta\right) \tag{1}$$

From this frequency the angular interval to be expected in the IME signal between each impulse from an outer race defect, expressed as a number of grating lines, can be calculated. Figure 3 shows how the value of f calculated for an NF202 bearing, 6.44, compares to the angular interval between the impulses actually recorded by the IME. The calculated interval was 159 lines, the average value in this example of the IME signal was 166 lines. The slight difference can be attributed to the slippage which occurs between the rolling elements and the races as the bearing rotates. The NF202 does not have a cage so slippage can cause the interval between impulses to vary. In this example the spacings ranged between 150 and 184 lines. The impulse amplitude is irregular due to the changes in rolling element position within the load zone of the bearing. This changes the relative amount of load at the location of the defect which determines the amount of movement recorded by the IME[18].

Once impulses in the shaft centre acceleration signal are detected from a rise in the kurtosis level, the angular interval between impulses can be compared to the values calculated for bearing defects. On an amplitude distribution curve the impulsive components of a signal, which produce a high kurtosis coefficient, will lie above the level of twice the standard deviation. This is the level used by the experimental IME system to set the amplitude level that counts as an impulse. For each impulse in twenty rotations of data the system measures the angular distance to impulses in the three succeeding and preceding rotations of data. By dividing the distances measured between impulses in the signal by the intervals calculated for bearing defects the system can determine whether or not the source of the impulses is part of the bearing. At the same time, by dividing the length of a rotation by the interval measured between impulses, the system is able to determine if the interval is an integer fraction of a rotation. The impulses due to bearing defects are always spaced at intervals that are non-integer fractions of a rotation as a result of bearing geometry. Shaft vibration sources such as gear meshing or fan blade passing always result in an integer number of impulses per rotation corresponding to the number of gear teeth or fan blades present. This difference provides the method by which the two types of vibration source can differentiated.

Figure 4 shows the shaft centre acceleration signal from the experimental IME when a small gear wheel was attached to the shaft and used to drive second gear wheel attached to a flywheel. The bearing supporting the shaft in this example was a normal NJ202 roller bearing. When the gear wheel on the shaft was out of contact with the one on the flywheel, the acceleration signal had an average kurtosis value of 3.15 indicating a random signal from normal, defect free, bearing surfaces. The kurtosis value of the signal in figure 4, with the gear wheels in contact, is 18.82. The small gear wheel had 16 teeth. An impulse interval equivalent to 1/16 of the rotation can be detected in this signal which would clearly allow its source to be determined. Our investigation into the use of the IME for gear monitoring is in its initial stages.

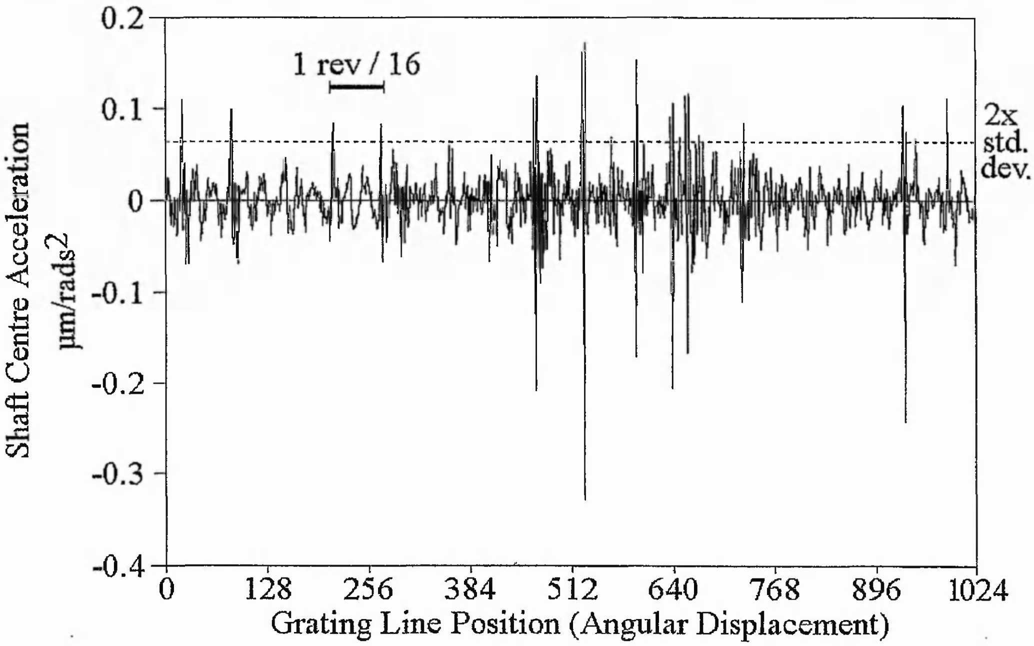


Figure 4. Shaft centre acceleration signal from a shaft driving a gear wheel with 16 teeth. Impulses occur at an approximate spacing of 1/16 of a rotation.

5. IME SHAFT LOADING MEASUREMENTS

There are numerous industrial settings where the loading on a rotating machine's shaft is an important process control parameter. Many industrial processes involve the winding, rolling or drawing of material using rollers in a manner similar to that shown in figure 5. Optical encoders are commonly used to control the rotation of such rollers while separate strain gauges or other sensors would measure the load on the roller and the tension in the material. The bearings supporting the roller in such an situation have been recognised as a suitable site for mounting sensors to measure shaft loading and a number of companies market bearings with integral strain gauges for this purpose[19]. The forces on the spindle bearings of machine tools such as lathes and grinding machines can be used to monitor tool wear. As cutting tools wear the forces on them during machining operations increases. Trending these forces can allow tool breakage to predicted and avoided.

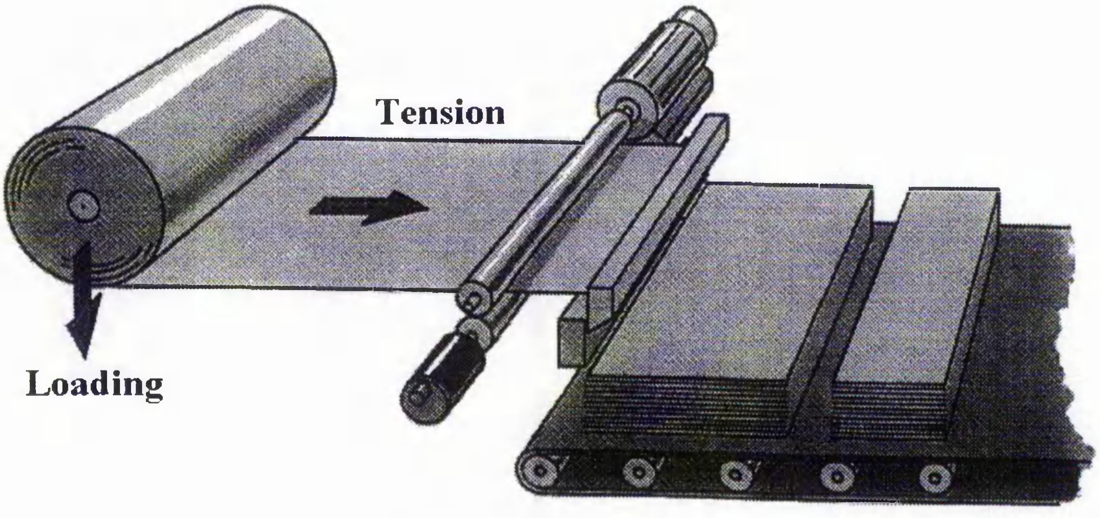


Figure 5. A typical industrial rolling application.

The IME offers the capability to measure the load on a shaft from the same sensor as might be used by a control system for speed and position monitoring. To simulate the progressive increase in shaft loading that would occur in a situation similar to that shown in figure 5 a set of weights was used to load the shaft of the experimental IME. Figure 6 shows a graph of the shaft center displacement in the sensor's y axis as the load on the shaft was increased from 0 to 2000g. From 0g to 1000g the increment was 50g and from 1000g up to 2000g the increments were 100g. Readings were also made at 1250g and 1750g. Each point is an average of data collected with a particular load over 24 shaft rotations. In each rotation the 1024 points of the sinusoidal signal are averaged to give a single value for displacement which is then averaged with the values for all the rotations. In total 24,576 values are averaged for each point on the graph. Figure 7 shows the displacement signals from the x and y axes of the sensor for the same data. The two sets of sinusoids together produce the orbit plots whose centres are the points in the graph shown in figure 6.

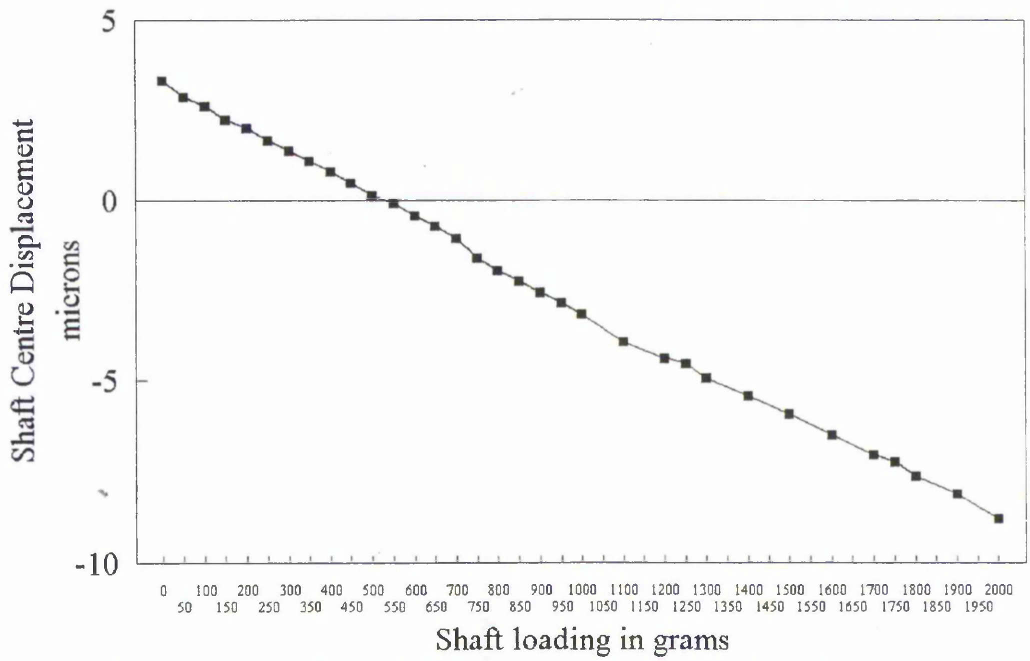


Figure 6. Shaft centre displacement in response to progressive shaft loading.

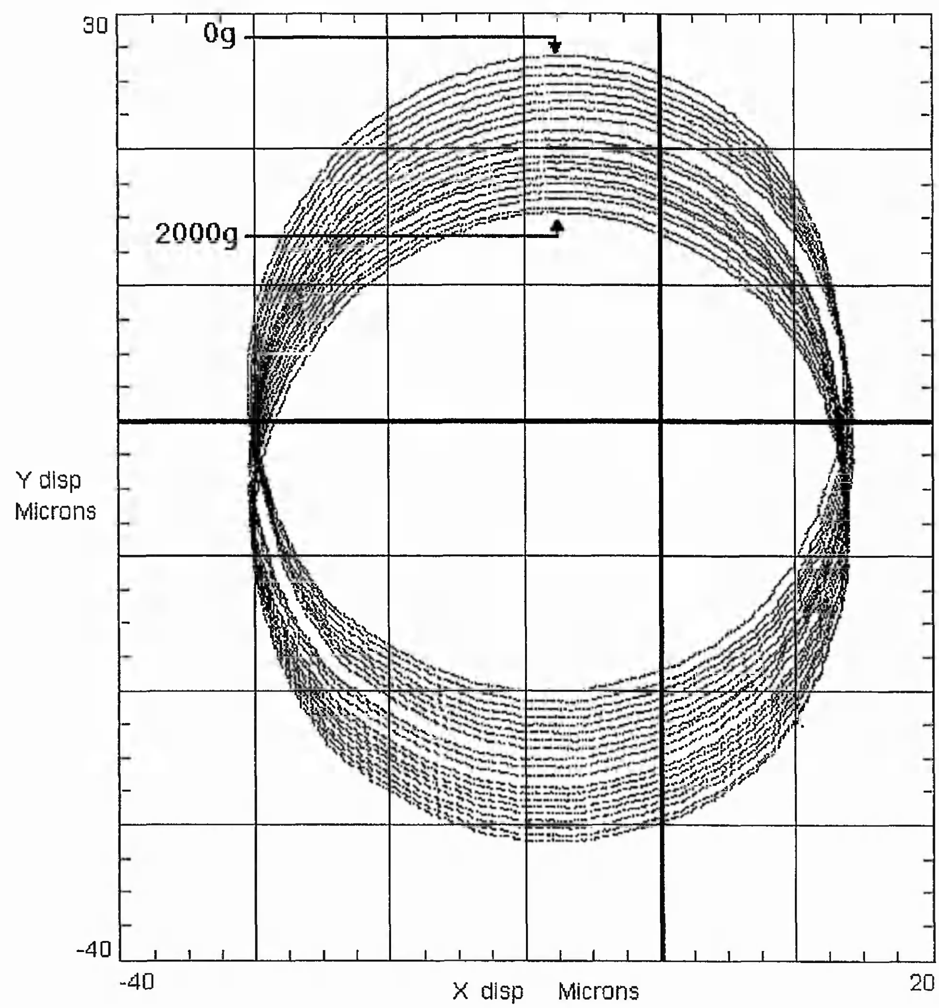


Figure 7. Orbital motion plots of shaft displacement due to loading from 0 to 2000g in 100g Increments. Each orbit averaged from 24 rotations.

The total shaft displacement shown on the graph for the load applied was 12.13 microns. This displacement was approximately linear with a gradient of 0.006 microns/g. The average displacement measured for each 50g increase in load was 0.323 microns, with a maximum value of 0.553 microns and minimum value 0.240 microns. This displacement increment and the variation recorded were similar in the three experiments of this type conducted. Changes in the experimental setup during the test were thought the most likely explanation for the variation. The weights were suspended from the shaft between two low friction tapes, used to distribute the load across a 5cm length of shaft. Although every attempt was made to ensure that each weight increment was mounted at the same point on the shaft, some small differences in the distance of the load from the disk were inevitable. The NJ202 bearing supporting the shaft had a load rating of 9800Kg which meant that most of the appreciable displacement from a 2Kg load was expected to come from the bending of the shaft. The exact amount of displacement measured therefore depended on the load and the distance of the load from the disk.

6. SMART BEARINGS

We believe that the IME has great potential as an industrial control and monitoring sensor. Although our experimental work to date has used a sensor based on optical components, sensors based on the IME principle can be implemented using any technology capable of giving a sufficiently high resolution square wave signal from a rotating, marked surface. In particular the recent development of REB with integrated speed sensors opens up the possibility of a completely new type of bearing combining an IME sensor. The widespread introduction of anti-lock braking systems (ABS) by the automotive industry has produced a number of bearings designed to accurately sense road wheel speed[20]. These use encoder-like electromagnetic sensors with read head mounted on the outer ring of the bearing and a grating rotating with the inner ring. This type of sensorised bearing has proven successful and in recent years some bearing manufacturers have begun to develop and market sensorised bearings for general applications, particularly speed sensing in machine tools.

Using the technology of these speed sensing bearings, with the addition of extra read heads and higher resolution encoder rings, an IME sensor could be developed capable of being integrated into existing machine designs with only minimal modifications for cabling. A bearing of this type could be used for process control and monitoring during normal machine operation as well as for detecting fault conditions, both in the machine and the bearing. Such bearings could monitor machine parameters such as angular velocity, angular vibration, torque and shaft loading together with their own condition. A number of conventional sensors and their respective data analysis systems could be replaced with a single unit. As a system component, a smart bearing like this would a be significant step forward in the direction of the type of intelligent machine required for use in CIM systems[1].

Figure 8 is an artist's impression of a smart bearing based upon the design of a speed sensing bearing currently available commercially. This sensorised bearing uses a single hall effect sensor in the seal attached to the bearing's outer ring and a magnetic multipolar ring fixed to the rotating inner ring. To achieve the basic functionality of an IME device only the addition of the second sensor shown in figure 8 would be required. This would allow the monitoring of shaft centre motion in one dimension. The relative simplicity of the calculations required to process the data from the IME would allow existing micro computer based machine controllers to interface directly to a smart bearing with the addition of minimal data acquisition hardware.

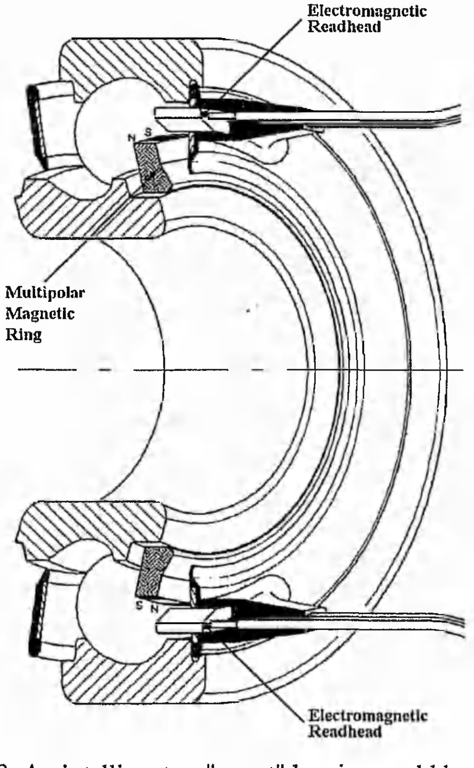


Figure 8. An intelligent or "smart" bearing could be created by the addition of a second read head to a speed sensing bearing.

7. CONCLUSIONS

Machine tool condition monitoring is a important factor in the smooth and economic running of modern industry. A new type of sensor for machine process control and condition monitoring has been described with examples of its experimental use. This sensor, the incremental motion encoder, is uniquely able to relate the position of impulses in its signal to the angular rotation of the supporting shaft. It has the ability to detect bearing defects and can differentiate these from other sources of machine vibration by determining the angular separation of impulses in its measurement of shaft centre acceleration. From its basic measurement of shaft centre displacement it can also measure incremental loading applied to a shaft.

This work forms part of an ongoing project to develop a comprehensive interpretation framework for all the IME's possible data output types. The work so far conducted suggests this device has considerable potential. Using the IME principle, a new class of sensorised bearings can be developed with current technology. These would be able to sense a number of machine parameters and their own condition. Within an intelligent machine tool, a smart bearing of this type would allow a higher degree of control and monitoring information integration than is currently possible for machine tools and other rotating machinery.

8. ACKNOWLEDGEMENTS

The authors gratefully thank Hewlett-Packard Components Limited for providing the kit encoder parts for the experimental incremental motion encoder.

This work is funded within the Nottingham Trent University by the Department of Computing and by a grant from the Japanese Electro-Mechanic Technology Advancing Foundation.

9. REFERENCES

1. A. Davies, "The intelligent machine", *Manufacturing Engineer*, pp. 182-185, Aug. 1994.

2. K. F. Martin, "A review by discussion of condition monitoring and fault diagnosis in machine tools", *Int. J. Mach. Manufact.*, vol. 34, n. 4, pp. 527-551, 1994.

3. M. Serridge, "What makes vibration condition monitoring reliable ?", *Noise and Vibration Worldwide*, pp. 17-24, Sept. 1991.

4. K. F. Martin, S. M. Hoh and J. H. Williams, "Condition monitoring machine tool drives via health indices", IFAC Symp. SAFEPROCESS '91, pp. 311-316, Baden Baden, Sept. 1991.

5. C. N. Ramirez and R. J. Thornhill, "On estimating rotational speed and torque from the period of an encoder", *ASME Sensors, controls and quality issues in manufacturing*, 1991.

6. P. A. Orton, Instrumentation and control for precision grinding machines , Ph.D. Thesis, Department of Computing, Nottingham Polytechnic, 1990.

7. P. A. Orton, Shaft Displacement Measuring System, International patent application number PCT/GB92/00811, World Intellectual Property Organization, International Publication number WO 92/19936, 12 November 1992.

8. A. Ernst, Digital linear and angular metrology , pp. 25-28, verlag moderne industrie, 1989.

9. P. A. Orton, K. Ayandokun, N. Sherkat and P. D. Thomas, "A fibre optic rotary motion sensor for the real-time condition monitoring of rotating machinery," in *Fibre Optic Physical Sensors in Manufacturing and Transportation*, John W. Berthold, Richard O. Claus, Michael A. Marcus, Robert S. Rogowski, Editors, Proc. SPIE 2072, pp. 195-207, 1994.

10. J. E. Berry, "How to track rolling element bearing health with vibration signature analysis", *Sound and Vibration*, vol. 25, n. 11, pp. 24-35, Nov. 1991.

11. J. Mathew and R. J. Alfredson, "The condition monitoring of rolling element bearings using vibration analysis", *Journal of Vibration, Acoustic, Stress and Reliability in Design* , vol. 106, pp. 447-453, 1984.

12. D. Dyer and R. M. Stewart, "Detection of rolling element bearing damage by statistical vibration analysis", *J. Mech. Design*, vol. 100, pp. 229-235, 1978.
13. P. D. McFadden and J. D. Smith, "Vibration monitoring of rolling element bearings by the high frequency resonance technique", *Tribology Int.* vol. 17, n. 1, pp. 3-10, 1984.

14. A. G. Ray, "Monitoring rolling contact bearings under adverse conditions," Proceedings of the 2nd International Conference on Vibrations in Rotating Machinery, Institution of Mechanical Engineers, pp. 187-194, 1-4 Sept. 1980.

15. R. A. Collacott, , Vibration Monitoring and Diagnosis , pp. 133-134, George Godwin Limited, London, 1979.

16. R. G. Harker and J. S. Hansen, "Rolling element bearing monitoring using high gain eddy current transducers", ASME Paper 84-GT-25, October 1984.

17. "REBAM® an effective way to monitor and analyze rolling element bearing condition", Applications Note AN047-02, Bentley Nevada Co., Sept. 1992.

18. K. Ayandokun, P. A. Orton, N. Sherkat & P. D. Thomas, "Detecting rolling element bearing defects with the optical incremental motion encoder", in *Industrial Optical Sensors for Metrology*, Kevin G. Harding, H. Philip Stahl, Editors, Proc. SPIE 2349, pp. 53-64, 1995.

19. M. L. Courtwright, "Instrumented bearings get rolling with ABS", *Machine Design*, vol. 64, n. 1, pp. 84-88, 1992.

20. O. Bayer, "A new generation: ABS capable wheel bearings", SAE International Congress and Exposition , Detroit, Michigan, Feb. 25 - Mar. 1, 1991.

Appendix I.

Smart bearings: Developing a new technique for the condition monitoring of rotating machinery

K. Ayandokun, P. A. Orton, N. Sherkat & P. D. Thomas, Proceedings of INES'97 The IEEE International conference on intelligent engineering systems, pp. 505 - 511, Budapest, Hungary, Sept 15 - 17, 1997

SMART BEARINGS: DEVELOPING A NEW TECHNIQUE FOR THE CONDITION MONITORING OF ROTATING MACHINERY

K. Ayandokun, P. A. Orton, N. Sherkat & P. D. Thomas.

The Nottingham Trent University, Department of Computing, Burton Street, Nottingham NG1 4BU, United Kingdom, Tel: +44 (0)115 948 6538, Fax: +44 (0)115 948 6518, Email: pao@doc.ntu.ac.uk

Abstract - The useful life of rolling element bearings can be significantly shorter than their design life due to factors in their operating environment. To prevent potentially unexpected bearing failure effective condition monitoring is required. This paper describes the experimental use of a new technique for the condition monitoring of rolling element bearings. The Incremental Motion Encoder (IME) is a new type of rotary motion sensor that can measure the relative motion between a shaft and its supporting bearing. This motion is directly related to the condition of the bearing. The advantage of this method, over the measurement of transmitted vibration, is that the signal is not attenuated in transmission or distorted by other vibration sources. An experiment demonstrating the ability of an experimental IME device to detect distinct changes in the shaft centre displacement related to bearing corrosion is described to illustrate the IME principle. Based on our work we propose that a bearing constructed with an integrated IME sensor could sense its own condition as well as external factors such as loading and imbalance. A sensorised or smart bearing of this new type could be an effective condition monitoring and process control sensor for machine tools and other rotary equipment.

I. Introduction

Condition based maintenance (CBM) has become accepted by industry as the key to avoiding expensive unplanned machine stoppages and maintaining high production standards. The growing importance of computer integrated manufacturing (CIM) is increasing the need for better machine monitoring and fault detection. The high capital cost of the sophisticated machine tools in CIM systems requires that they be operated at peak efficiency to generate maximum return. High reliability, achieved through effective maintenance, is an essential prerequisite for the successful operation of a CIM system [1].

Rolling element bearings are common subjects for condition monitoring as they are generally critical components and are prone to unexpected failure. Actual bearing service life can be significantly shorter than the design estimate derived from fatigue life studies [2]. Bearing failure can be hastened by common operating factors such as lubricant contamination, shaft misalignment or improper loading. It has been estimated that only 10 to 20% of bearings reach their design life [3].

The basis of an effective CBM strategy is the accurate assessment of machine component condition using suitable sensors. Several methods of determining bearing condition by analysing machine vibration have been developed [3,4,5,6]. All of these techniques depend on the fidelity of the vibration signal transmitted through the machine structure to a sensor mounted on the machine casing. This signal will suffer attenuation as it is transmitted and may be distorted at component interfaces within the machine. The sensor location may also not give the best transmission path. As each bearing may be only one of several vibration sources, proper diagnosis of an individual fault from the vibration signal can be difficult, particularly in the early stages of bearing failure [7]. The most common vibration sensor, the accelerometer, has a limited ability to measure the low level, low frequency vibrations from bearings in machines operating at low rpm [8].

To overcome the problems inherent in monitoring bearing condition via transmitted vibration, proximity measurement of shaft relative displacement has been suggested as an alternative method [9]. The displacement between a shaft and its mounting is directly related to the condition of the bearing supporting the shaft. At present, however, shaft displacement monitoring is limited almost exclusively to fluid film journal bearings which are designed to have large clearances. The extremely small clearances of rolling element bearings do not allow enough relative motion between shaft and bearing housing for conventional condition monitoring measurements. The advantages of proximity monitoring have been shown by the successful use of eddy current displacement sensors to determine bearing condition. These are used to measure deflections of the bearing outer race from within the bearing housing [10,11]. The difficulties involved in mounting the sensors for this technique have prevented its widespread application.

Our research aims to extend the use of relative displacement measurement to rolling element bearings using a new type of sensor, the Incremental Motion Encoder (IME). This new device has the advantage of sensing the mechanical motion caused by the presence of a defect directly, allowing monitoring of bearings in low speed machines. It also has the potential to directly monitor the state of single bearings in complex machines.

II. THE INCREMENTAL MOTION ENCODER

The IME is based on a patented evolution of the optical shaft encoder that allows a high degree of angular accuracy to be derived from a relatively low resolution encoder disk [12]. Four encoder readheads mounted on the bearing housing are spaced at 90° around a grating disk mounted on the shaft. Measurement is made between each pair of diametrically opposed readheads as illustrated at positions A and B in fig. 1. One measurement of the displacement in microns, d, of the disk centre from the readhead centreline is calculated for each line of the grating as it rotates with the shaft. A computer is used to calculate the angular difference, θ, measured between each readhead from the time difference, Δt, of their square wave output signals [13]. Equation (1) is used for the calculation of the disk centre displacement.

$$d = \frac{(\pi - \theta)L}{4} \qquad (1)$$

Plotting the displacement values against angular position produces a sinusoidal trace similar to that shown in fig. 1. The sinusoid results from the motion in one dimension of the centre of the grating lines on the disk around the actual centre of shaft rotation. The sinusoid is modified by defects on the bearing surfaces, the varying positions of the bearing elements and the motion of the shaft due to external forces such as loading, imbalance and gear meshing. Combining the sinusoidal signal from each of the two axes of measurement allows the construction of a shaft centre motion plot. Shaft centre velocity and acceleration are calculated by simple differentiation of the displacement values.

A single synchronising point on the grating disk allows each displacement measurement to be made at a specific angular position, corresponding to a grating line, on each rotation. This allows the displacement at the same point on successive rotations to be precisely compared over any period of time. This is a unique feature of the data produced by IME sensor. Comparison of the maximum and minimum displacement values for the same point over a number of rotations is one of the ways that an IME sensor can give an indication of bearing condition.

III. DETERMINING BEARING CONDITION WITH THE IME

In a properly lubricated, defect free, rolling element bearing the motion recorded by an IME sensor will be related to the normal contacts between the bearing's races and its elements. The shaft centre acceleration will show a random amplitude distribution due to the random roughness of normal bearing surfaces. The presence of a spall-type defect on one of the bearing surfaces

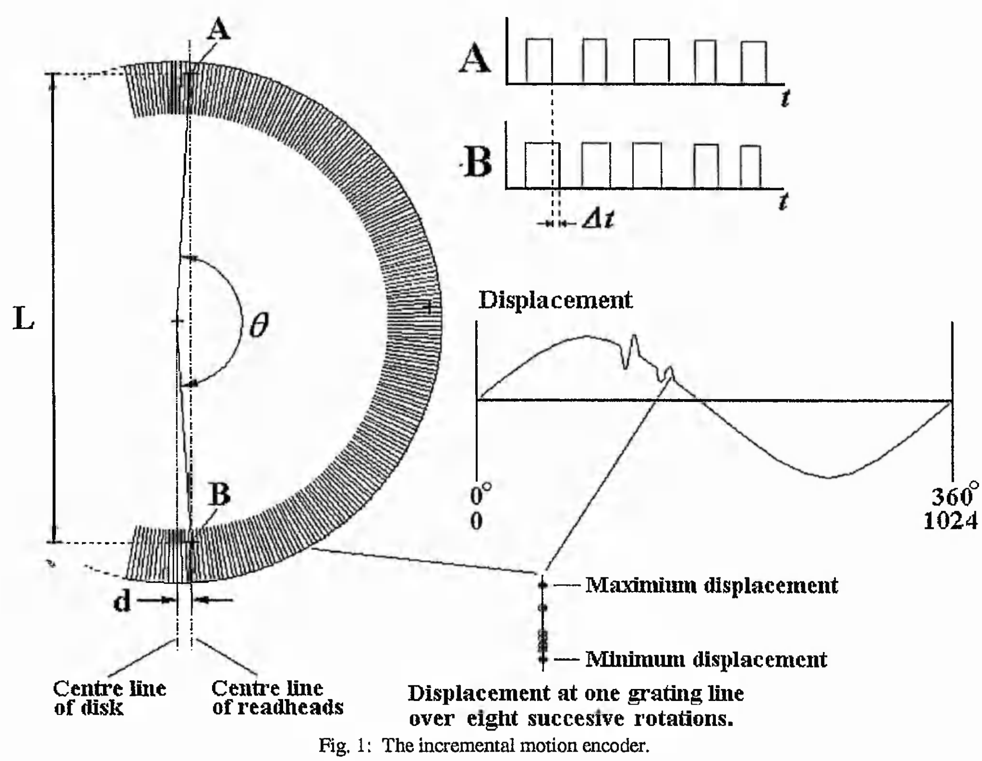


Fig. 1: The incremental motion encoder.

introduces an impulsive component to the acceleration signal. This can be detected from the resulting change in the amplitude distribution measured by the statistical parameter kurtosis [14]. The angular separation of the impulses can be used to determine whether they are from the bearing or another source affecting shaft motion.

Some bearing conditions that may precede actual defect formation do not produce an impulsive signal when monitored by the IME. Detection of such conditions as lubricant contamination, lubricant starvation and bearing corrosion can provide important early warning of bearing failure. Monitoring of the kurtosis value has not been found appropriate for trending all bearing conditions in the experiments so far conducted. In the case of lubricant contamination experiments have shown the IME registers an increase in overall acceleration level while the amplitude distribution remains random [15].

Fig. 2 shows the acceleration amplitude frequency distribution from two NF202 roller bearings. Using electric discharge machining a groove 0.2mm wide and 0.2mm deep was made across the outer race of one to simulate a spall-type defect. This produced the amplitude distribution with the sharp peak in fig. 2 which has a kurtosis value of 10.15.

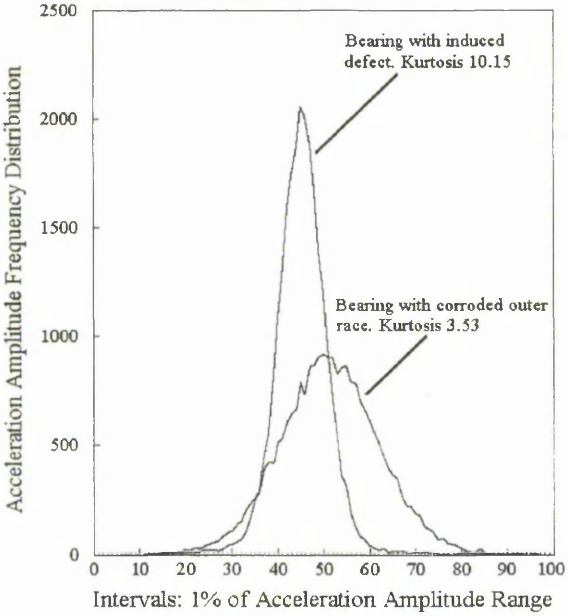


Fig. 2: Acceleration amplitude frequency distribution for two bearings with different types of defects.

The second bearing's acceleration amplitude shows a close to random frequency distribution with a kurtosis value of 3.53. From this measurement the bearing would appear normal, but inspection of its outer race, shown in fig. 3, reveals the presence of greyish black streaks, coinciding with the roller spacing, around approximately

Fig. 3: Corroded outer race of an NF202 bearing showing the effect of water etching due to moisture exposure during storage

one third of the race circumference. These are a result of galvanic corrosion between the race and the rollers due to the presence of moisture in the bearing during storage [16]. In the corroded areas the race surface has been etched increasing its roughness and therefore the stress generated between it and the rollers. This stress would accelerate the breakdown of the race surface, due to fatigue loading in normal operation, leading to premature bearing failure.

The surface roughness produced by the corrosion has a random nature similar to that of the original race surface. The acceleration signal therefore remains random in the presence of the corrosion. However the loss of material from the race surface in the corroded areas leads to an increase in the internal clearances of the bearing that can be seen from displacement measurements.

Fig. 4 shows the displacement in two dimensions measured by the IME from an undamaged NF202 roller bearing over four rotations. The successive rotations are superimposed and show that the displacement measurements for the same angular position have only a small degree of variation as the circular traces are largely coincident. For a normal bearing the changes in displacement recorded at a particular angular point are due to the variations in position of the rolling elements in the bearing load zone. The motion of the elements is cyclical. They return to the same positions after a number of rotations determined by bearing geometry and the amount of slippage between elements and races. By calculating the difference in displacement recorded at the same angular position over a number of rotations a measure of the clearances in the bearing can be obtained. In particular the cumulative sum of these values taken over several rotations has been found to provide a useful parameter for the trending of bearing condition.

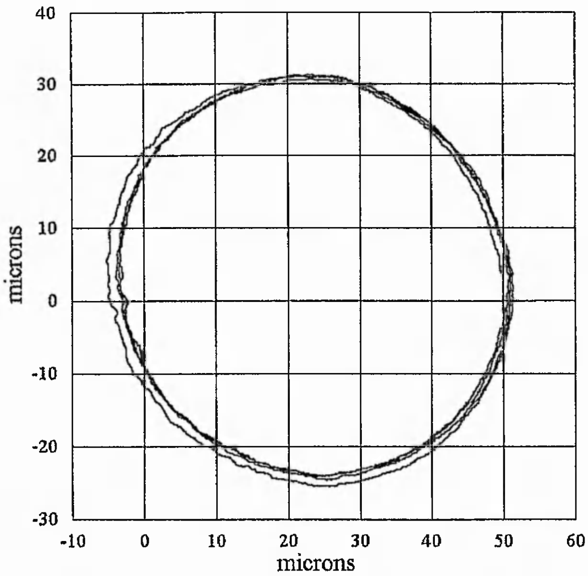


Fig. 4: Shaft displacement in two dimensions measured by the IME from an undamaged NF202 roller bearing rotating at 678 rpm.

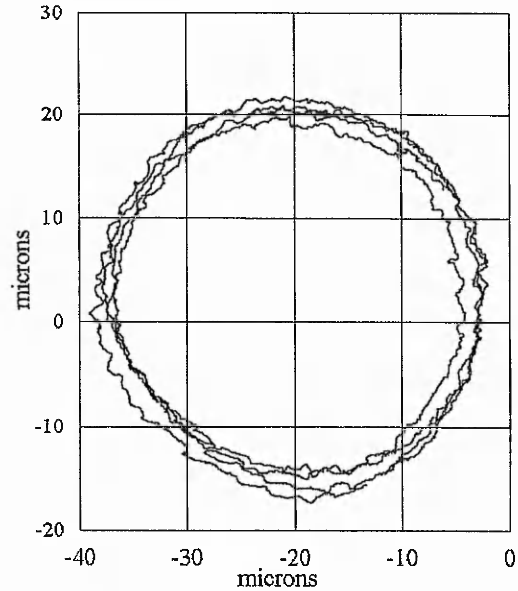


Fig. 5: Shaft centre displacement from an NF202 roller bearing with corroded outer race. Speed 588 rpm.

Fig. 5 shows the shaft centre displacement from four successive rotations of the corroded bearing described above. It was mounted with the corroded area in the load zone with a nominal load on the shaft. The irregular outline of the traces and the amount by which they diverge from one another clearly shows the sensitivity of the IME to the state of the bearing in this test. The corroded portions of the outer race do not present sufficient relief relative to the normal areas to create the sharp changes in displacement which produce impulsive acceleration. Instead the motion seen is that which would be expected from the passage of a rolling element across a wavy, irregular surface.

Fig. 6 compares the cumulative displacement differences measured in the horizontal axis of movement from the normal and the corroded bearing. Each point on the traces represents the sum of the differences in displacement measured between eight successive rotations at one of the 1024 angular positions of the encoder disk. The corroded bearing's race allows a larger amount of shaft centre movement, both in absolute terms with a maximum value of 8.24μm compared to 5.09μm, and in relative terms with a range of 5.02μm against 3.17μm for the normal bearing. These values

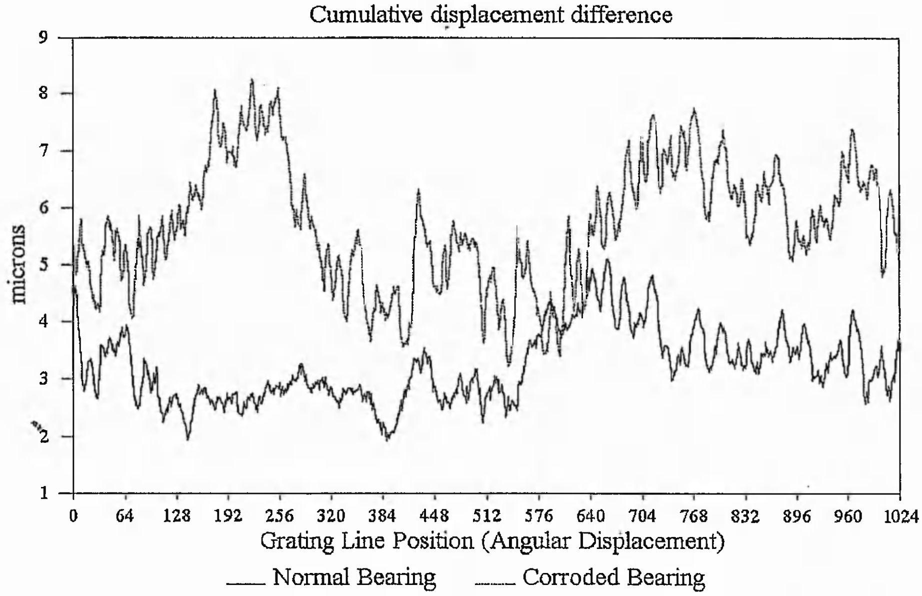


Fig. 6: Cumulative displacement difference values calculated over eight rotations for a normal and a corroded bearing.

were found to be consistent from a series of eighteen tests. The corroded bearing data was compared against two undamaged bearings whose maximum cumulative displacement differences differed by less than 0.5μm.

IV. SMART BEARINGS

Until recently rolling element bearings were monitored, but had no sensory ability of their own. Their rotational speed might be measured for a machine control system by a shaft encoder and their condition monitored by a separate diagnostic system. The introduction of anti-lock braking systems (ABS) by the automotive industry created the first application for bearings with integrated sensors. Car road wheel speed could be most accurately measured by incorporating an encoder-like electromagnetic sensor into the wheel bearing. Several technologies have been developed to be used by ABS and the sensorised bearing has proven very successful in this area [17]. In the past five years some bearing manufacturers have begun to develop and market sensorised bearings for general applications, particularly in machine tools. The emphasis has primarily been on rotary speed sensing although bearings with internal strain gauges to measure shaft loading have been developed [18]. Shaft loading can be used to monitor tool wear in machine tools and roller tension in a variety of industrial processes.

Based on our work developing the incremental motion encoder, we believe that there is a great potential for a new generation of bearings incorporating the IME principle. Such bearings could monitor machine parameters such as angular velocity, angular vibration, torque and shaft loading together with their own condition. A number of conventional sensors and their respective data analysis systems could be replaced with a single unit. This could be integrated into existing machine designs with only minimal modifications for cabling. A bearing of this type could be used for process control and monitoring during normal machine operation as well as for detecting fault conditions. As a system component, a smart bearing would a significant step forward in the direction of the type of intelligent machine required for use in CIM systems [19].

Fig. 7 is an artist's impression of a smart bearing based upon the design of a speed sensing bearing currently available commercially. This sensorised bearing uses a single hall effect sensor in the seal attached to the bearing's outer ring and a magnetic multipolar ring fixed to the rotating inner ring. To achieve the basic functionality of an IME device only the addition of the second sensor shown in fig. 7 would be required. The relative simplicity of the calculations required to process the data from the IME would allow existing micro computer based machine controllers to interface directly to a smart bearing with the addition of minimal data acquisition hardware.

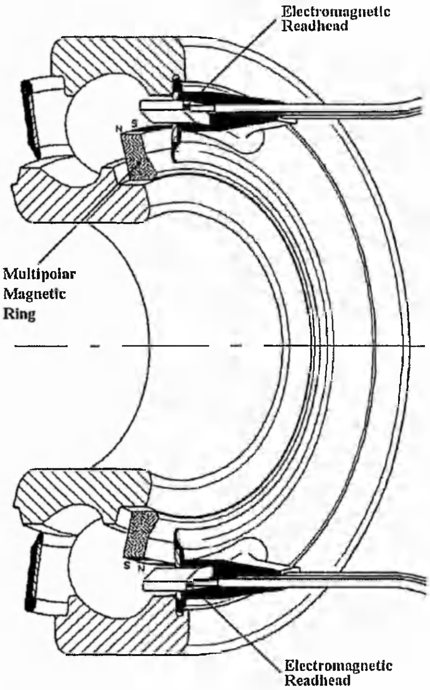


Fig. 7: An intelligent or "smart" bearing could be created by the addition of a second readhead to a speed sensing bearing.

V. CONCLUSIONS

Bearing corrosion, an example of pre-spalling race damage, is one of several conditions that can provide a warning of potential bearing failure. The sensitivity of the IME to changes in the internal clearances of a rolling element bearing related to corrosion has been demonstrated. Distinct patterns that can be related to the effect of the corrosion damage on the race surface were observed in the shaft centre displacement measured by the IME. While the shaft centre acceleration was similar to that from a normal bearing, the corrosion gave irregular displacement values that could be easily distinguished. The use of the cumulative displacement difference measured between the same point on successive rotations, a unique feature of IME data, was found to be an effective parameter for evaluating displacement data.

A comprehensive interpretation framework for all the IME's possible data output types is still in the development stage, however from the work so far conducted this device has considerable potential. Using the IME principle, a new class of sensorised bearings could be developed with current technology. These would be able to sense a number of machine parameters and their own condition. Within an intelligent machine tool, a smart bearing of this type would allow a higher degree of control and monitoring information integration than is currently possible for rotating machines.

VI. ACKNOWLEDGEMENTS

The authors gratefully thank Hewlett-Packard Components Limited for providing the kit encoder parts for the experimental IME device.

This work is funded within the Nottingham Trent University by the Department of Computing and by a grant from the Japanese Electro-Mechanic Technology Advancing Foundation (EMTAF).

VII. REFERENCES

1. J. Lee and B. Kramer, "On-line fault monitoring and detection using an integrated learning and reasoning approach", Japan/USA symposium on flexible automation, vol. 1, ASME, 1992.

2. T. E. Tallian, "Simplified contact fatigue life prediction model-Part 1: Review of published models", *ASME Journal of Tribology*, v 114, p 207-213, 1992.

3. J. E. Berry, "How to track rolling element bearing health with vibration signature analysis", *Sound and Vibration*, vol 25, part 11, pp. 24 - 35, November, 1991.

4. J. Mathew and R. J. Alfredson, "The condition monitoring of rolling element bearings using vibration analysis," *Journal of Vibration, Acoustic, Stress and Reliability in Design*, vol. 106, pp. 447-453, 1984.

5. D. Dyer and R. M. Stewart, "Detection of rolling element bearing damage by statistical vibration analysis", *J. Mech. Design*, vol. 100, pp. 229-235, 1978.

6. P. D. McFadden and J. D. Smith, "Vibration monitoring of rolling element bearings by the high frequency resonance technique", *Tribology Int.* vol. 17, n. 1, pp. 3-10, 1984.

7. A. G. Ray, "Monitoring rolling contact bearings under adverse conditions," Proceedings of the 2nd International Conference on Vibrations in Rotating Machinery, Institution of Mechanical Engineers, pp. 187 - 194, 1 - 4 Sept., 1980.

8. F. Schloss, "Accelerometer noise," *Sound and Vibration*, vol 27, n 3, pp. 22 - 23, 1993.

9. R. A. Collacott, *Vibration Monitoring and Diagnosis*, ch. 7, pp. 131-135, George Godwin Limited, London, 1979.

10. R. G. Harker and J. S. Hansen, "Rolling element bearing monitoring using high gain eddy current transducers", ASME Paper 84-GT-25, October 1984.

11. REBAM® an effective way to monitor and analyze rolling element bearing condition, Applications Note AN047-02, Bentley Nevada Co., Sept., 1992.

12. P. A. Orton, Shaft Displacement Measuring System, International patent application number PCT/GB92/00811, World Intellectual Property Organization, International Publication number WO 92/19936, 12 November 1992.

13. P. A. Orton, K. Ayandokun, N. Sherkat and P. D. Thomas, "A fibre optic rotary motion sensor for the real-time condition monitoring of rotating machinery," in *Fibre Optic Physical Sensors in Manufacturing and Transportation*, John W. Berthold, Richard O. Claus, Michael A. Marcus, Robert S. Rogowski, Editors, Proc. SPIE 2072, pp. 195 - 207, 1994.

14. K. Ayandokun, P. A. Orton, N. Sherkat & P. D. Thomas, "Detecting rolling element bearing defects with the optical incremental motion encoder", in *Industrial Optical Sensors for Metrology*, Kevin G. Harding, H. Philip Stahl, Editors, Proc. SPIE 2349, pp. 53-64, 1995.

15. K. Ayandokun, P. A. Orton, N. Sherkat and P. D. Thomas, "Tracking the development of rolling element bearing faults with the optical incremental motion encoder", Proceedings of the 8th International Congress on Condition Monitoring and Diagnostic Engineering Management (COMADEM '95) vol.2, pp. 649-655, Kingston, Ontario, Canada, 1995.

16. Bearing failures and their causes, SKF publication PI 401 E, 1994.

17. O. Bayer, "A new generation: ABS capable wheel bearings", SAE International Congress and Exposition, Detroit, Michigan, Feb. 25 - Mar. 1, 1991.

18. M. L. Courtwright, "Instrumented bearings get rolling with ABS," *Machine Design*, vol 64 n 1 pp. 84 - 88, 1992.

19. A. Davies, "The intelligent machine", *Manufacturing Engineer*, Aug., pp. 182-185, 1994.