# Automatic Classification of Tubing Defects by Analysis of their Eddy Current Signals

A Thesis submitted to the Faculty of Engineering
of the University of Glasgow
for the degree of Doctor of Philosophy

by

Kenneth T. Macfarlane, B.Sc.

November 1987.

**Summary**

This thesis contains the results of a study conducted upon the application of pattern recognition techniques to the signals from a rotating eddy current probe in essentially two circumstances. These were (firstly) the inspection of 316 finned austenitic stainless steel heat exchanger tubes destined for an Advanced Gas Cooled Reactor using single frequency eddy currents with a differential probe, and secondly, the inspection of Inconel 600 tubing for Pressurised Water Reactor steam generators with a multiplexed 4 frequency pancake probe.

In the first case, an automatic defect detection algorithm has been developed which will isolate any discernable axially oriented crack-type defect. Various classifiers (multiclass linear, k,l nearest neighbour, nearest mean, scan cross correlation, and adaptive learning network) have been coded and tested on a 3 class sample set and 96.9% correct classification achieved with the two least easily separable classes, whilst the harmless third class was removed with the segmentation algorithm.

In the second case, a new logging system was developed, and 38 tubes containing 45 synthesised defects were logged to disk and scans from each were plotted in the impedance plane, and against time. There were six types of defect, mostly in 4 different classes of depth. Seven different segmentation methods were devised and tested on the scans and a fitted linear threshold in the impedance plane was found to be best for sensitivity and speed of operation given the background (Pilger) noise present in the tube set,though another method which uses a convex hull algorithm is able to discriminate against any definable undesirable background signal. An

automatically annulable method of generating a subtractive mix of two channels was also developed for use with simple magnitude– based thresholding segmentation. Ad hoc geometrical and spectral featuresets were tested on the 7 class defect set with a multiclass linear and a nearest neighbour classifier. The correct classification rate achieved was 61.5%, and in view of this, and in an attempt to produce a site-trainable system, some development was done on parametrically-based feature extractors (FFT and Fourier Descriptors), and on a class distribution–independant feature set selector.

**Acknowledgements**

I would like to express my thanks to Prof. Lamb, Head of the Department of Electrical Engineering, University of Glasgow, for the provision of facilities for my research. I would also like to thank Dr. J.E.S. Macleod for starting the project off in the first place, and then supervising my research and the production of a thesis. My gratitude is also extended to my industrial supervisors Dr. J.M. Farley, Jim Hamilton, and particularly Dr. Patricia Roberts.

This research would not have been possible without the provision of eddy current instrumentation and sets of tubing by Babcock Power Research Ltd.. The inspection of the first set of tubes was done by the late Robert Sinnamon, a technician at Babcock's, who died tragically in 1984.

The financial support of the Science and Engineering Research Council and Babcock Power Ltd. is gratefully acknowledged.

I would also like to thank Miss Vi Romanes for her help in typing parts of this thesis, the staff of the Babcock Power Research Centre Library for the help they gave me in finding the more obscure reports to which I needed access, and the staff of the Computer Services Group for tolerating my monopolisation of their PDP11/23 whilst I compressed my suite of programs to run on their machine.

Most of all, I would like to thank those friends with whom I have lived whilst writing up, for their constant interest in my work and encouragement to finish. They were:my friends Andy Hunter and Sandy Day, my brother David, and my parents, Tom and Davina.

TABLE OF CONTENTS

# 1 Chapter 1: Introduction

This chapter contains a general introduction to nondestructive testing of steam generator tubing in pressurised water nuclear reactors, followed by a literature review of automatic eddy current inspection, and finally a guide to the organisation of this thesis.

## 1.1 Cracks in Steam Generator Tubing

A nuclear power station can be schematically represented by the diagram in fig.1.1. The reactor cooling fluid, the "primary coolant" is circulated around the primary circuit by the pump as shown, from the reactor to a heat exchanger or steam generator where heat is tranferred to the "secondary coolant" (water), which is then allowed to boil and thus expand in volume. The expansion of the steam is harnessed by turbines which convert the heat to electricity because they are connected to large generators. For the purposes of this introduction, the main point is that there are two separate cooling circuits, and the coolants are brought into close proximity with each other only within the heat exchanger, where they flow on the opposite sides of the thin metal walls of the tubes within that vessel.

Different types of reactor are usually named either according to the primary coolant or according to the method of cooling. For instance in an advanced gas cooled reactor (AGR), the coolant is a gas under pressure (carbon dioxide) and in a pressurised water reactor (PWR), it is pressurised water. The work in this thesis is relevant to eddy current inspection of tubing generally, but particularly to tubing used in these two types of reactor because the tubes used

in the experimental work were destined for AGR heat exchanger tubing (Chapter 4) or PWR steam generator (SG) tubing (Chapter 5).

The primary circuit in most reactor types is run considerably above atmospheric pressure, in order to increase the capacity of the coolant to carry heat away from the core. The heat is transferred from the primary to the secondary coolant across an interface within a heat exchanger. The secondary coolant is water, which is allowed to boil as it exits from the heat exchanger, having absorbed considerable thermal energy. The steam produced is used to drive turbines, which in turn drive generators to produce electricity.

A practical layout for a PWR power station is shown in fig. 1.2. This particular arrangement of 4 steam generators around one core is the Westinghouse "F" design. Secondary coolant is injected into the feedwater inlet halfway up each steam generator, and falls to the bottom, rises to the surface and then boils and leaves via the steam outlet. The pressurizer visible in the diagram is a device which limits the primary circuit pressure in normal operation, by spraycooling the hot primary water with cold water.

Each steam generator consists of a pressure vessel (fig.1.3), typically 20 metres tall, containing several thousand inverted U-tubes, each typically of internal diameter 1.5cm, and wall thickness of about 1mm. The inverted U-tubes are about 12 metres high, and braced against any lateral movement periodically up their length by 10 cm. thick tube support plates, particularly to prevent vibration of the tubes due to the turbulent passage of large amounts of secondary cooling water.

Both ends of the U-tubes are sealed into a matrix of holes in a thick sheet at the bottom called the tubesheet, arranged so that all of the inlets are on one side and the outlets on the other, which makes connection to the rest of primary circuit simpler than with any other arrangement of tube ends. Both the vessel wall and the tube sheet are 20 cm. thick. Access to the tube ends for inspection is via a manway on either side of the large bore primary fluid connections at the bottom of the steam generator (fig. 1.3). There are also period access holes up the side of the vessel, to allow visual inspection of the tube bundle exteriors in the vicinity of the tube support plates.

Under service conditions, the most troublesome component for leakages has been the steam generator (**Mundis and DeYoung**,1983), which contains the only interface between the high pressure primary circuit and the lower pressure secondary water/steam circuit. Due to the amount of tubing in it, the steam generator accounts for over 50% of the entire surface area of the high pressure primary circuit (see figs.1.2 & 1.3(a) and (b)), the remainder being accounted for by the thick walls of the primary pressure vessel and large bore pipes, where although the external pressure is atmospheric and the pressure differential greater, growth of millimetric cracks is less troublesome because of the difference in relative scale between the cracks and the wall thickness. Thus it is in the thin walled steam generator tubing that the defects that are of greatest concern occur.

## 1.2 The Necessity for In Service Inspection

In a steam generator in a PWR, the primary coolant flows through U-tubes surrounded by the secondary coolant (see figs. 1.1 & 1.3(a)&(b)). The steam

generators in Sizewell "B" (**Layfield**,1987), when it is built, will run with internal and external water pressures of 2250 and 1100 psi respectively (**Glasstone and Sesonske**,1981), figures that are typical for PWR steam generators manufactured by Westinghouse (the most prolific designer and manufacturer). In such a design, each of the steam generators has a mass flow rate of 5 tonnes of water per second which represents $1.25ms^{-1}$ along each of the 5600 tubes in each steam generator. The primary water enters the hot leg of the U-tubes at 325°C and leaves the cold leg at 275°C.

The occurrence of any crack through which the pressure in either circuit might be dissipated must be prevented, but especially so in the case of the primary circuit. The PWR design in particular is such that loss of pressure by failure of any one of the primary components under service conditions must be impossible. The reason for this is partly that since the core of a PWR is very compact (much more so than an AGR core), that the heat removal would be so much impaired by a loss of coolant pressure that core overheating would occur. The other reason is that the primary water, although at a pressure which is not particularly high or dangerous in itself (at 2250 psi, not much higher than that of the hydraulics in a JCB roadside digger), is also superheated and so is still liquid at 325°C solely because of the pressure under which it is held. If a crack were to propagate so that it finally breached a tube wall, the coolant would expand explosively through the crack as the water turned to steam, possibly causing some of the surrounding tubes to buckle or burst as well.

There is no alternative to using thin walled tubing: a large amount of tubing along with sufficient pressure in order to prevent the coolant from boiling is

required in order to achieve a high enough heat flux. Leaks tend to occur in the tubing, rather than in the thick section parts of the circuit, because of the thinness of the tube walls.

Thus it is that the detection of defects is essential because it is impossible to guarantee that defects will not occur as a result of deficiences in the tube manufacture or in the design itself.

Nondestructive testing is required because the plant is too expensive to be discarded after a safe life, and even if it was, it would still need to be inspected before going into service for the first time. **Birkle** (1976) gives the standards to which PWR steam generators have to be tested.

### 1.3 Possible Types of Defect

There are basically 6 types of defect:

**(i) axial and circumferential cracks.**

Cracks occur almost entirely aligned along these two orientations because of the dominant modes of deformation in tubes. Cracks tend to be almost closed, that is the two sides mesh, because when the tubes are inspected they are not subject to the forces present while the reactor is running. Cracks are typically a few mm. long and need to be of the order of 0.2mm deep to be detected by any method. They may occur due to residual stress from the manufacturing process, especially at the top of the inverted-U in the tube, where it has been bent through 180°; or where the tube diameter has been enlarged by explosive sealing of the tube

ends into the tube sheet at the bottom of the steam generator.

**(ii) small volume pits or holes, or clusters of pits.**

These typically originate from acidic attack of inclusions in tubing that occurred during the manufacturing process.

**(iii) Thinning.**

Uniform tube wall thinning is due wear caused by vibration of tubes within section of tubing constrained by the tube support plates. The vibration can be set up either by the turbulent flow of the primary coolant through the steam generator tubing or by the secondary coolant swirling around the tubing bundles as it boils.

**(iv) intergranular stress corrosion cracking (IGSCC).**

Stress due to thermal effects and the turbulent passage of coolant can cause a patch of tube wall to display myriad cracks because that area was contaminated with grease during the manufacturing process, and the composition of the grain boundaries impaired to the extent that when subsequent acid washing takes place, cracks propagate along a large number of the grain boundaries within a small area.

**(v) intergranular attack (IGA).**

Grain boundary dissolution due to acid attack by chemicals present in the coolant water, especially where they precipate out at tube fixing points and may lie in a sediment of corrosion by products for months or even years, until the next routine cleaning and inspection of the steam generator from the secondary side.

**(vi) Denting**

Tube wall distortion due to denting also occurs as a result of corrosion by products, but here to the extent that the tube is compressed and deformed.

Some defects are shown schematically in fig. 1.3(b). The tube support plates, which exist to brace the bundles of tubes against the high,turbulent flow rates within the steam generator, are commonly the site of defects such as fretting (also known as elliptical thinning) due to flow induced tube vibration, and denting due to the buildup and swelling of deposits of corrosion by-products (fig 1.3) from other parts of the secondary circuit, in the gaps between the tubes and the plates. There are also some rarer defects such as laps, candleflames and scallops which have not be included in this study.

The method of sealing the tubes to the tubesheet used in a lot of existing designs (by explosive expansion of the tubing to fit the hole) tends to encourage local corrosion due to sludge build-up and, from which cracks emerge because of the high residual stresses, and where leaks will eventually occur if the defects go undetected. The "tube roll" shown in fig. 1.4 refers to the sudden expansion of diameter at this point

It was not until many PWR's were in service that manufacturers realised that failure to expand the tubing over the entire thickness of the tubesheet left a crevice (see fig. 1.4) which would fill with chemically aggressive sludge, leading to the formation of secondary side corrosion phenomena such as intergranular attack, and intergranular stress corrosion cracking. Further attention to types of defect is paid in Chapters 4 and 5.

### 1.4 Detection of Defects

In the case of PWR steam generators, safety from losses of pressure can only be assured by pre-service and 2 yearly in- service inspections, in order to plug badly corroded or cracked tubes, thus pre-empting the loss of primary coolant through them. It is therefore a necessity to detect any degradation of the tube walls, and to be able to do this with certainty. For instance, in order to be granted a site operating licence for Sizewell "B", the CEGB has undertaken to ensure that all defects of any consequence, including those in the tube sheet region (where they are particularly difficult to detect) shall be demonstrably detectable.

**Tomlinson**(1985) explains the requirements made of the CEGB by the Nuclear Installations Inspectorate more fully in a paper demonstrating the capability to satisfy them with eddy current testing, in which the experimental work was performed by the author's industrial supervisor (**Dr. P.T.E. Roberts**). The tube set and equipment used to do this experimental work were also used for the work of Chapter 5.

In a steam generator, as fig. 1.3 shows, access to the inner surface of long tubes for inspection is very limited and largely dictates the choice of nondestructive test method. The candidate methods are reviewed below.

#### 1.4.1 Ultrasonics

Ultrasonic testing, the most obvious alternative to eddy current testing, is best suited for finding the range and bearing of reflective discontinuities buried deep in thick sections of metal or in welds, e.g. in the main pressure vessel

wall of a steam generator (**Thomson and Thomson**,1985). Defect detection with ultrasonics is usually based either upon pulse time of flight measurements or upon spectral analysis of the reflected sound.

The beams from ultrasonic probes are too wide and so do not have the necessary resolution to determine the shape (and hence the type) of defects, especially when of approximately the same size as the required minumum detectable size (0.2mm i.e. 20% through wall). It is commonly believed impossible to distiguish between the reflection from a defect and the far side wall reflection in thin walled tubing, because the reflected pulse front returns to a pulse echo transducer (which functions as the transmitter and then the receiver) before the emmited pulse has finished, and in addition, it is not possible to separate in time the reflections from a surface breaking defect and the surface itself, due to the short time difference between the echoes (~120 nanoseconds for a 0.2 mm. crack) and the length of the pulse (microseconds)(**Spinks**,1987). Recently, **Broere et al**(1987) announced a system capable of operating with sufficiently short pulses, and of challenging the superiority of rotating eddy current probes in the tube sheet region, but the system has still to be made available more generally. In addition, the defects are often virtually coincidental with the water/metal interface at the outside of the tubes, which will reflect the sound much more strongly and so tend to mask the defect signal. Furthermore, using ultrasonics requires a liquid couplant to acoustically match the piezoelectric resonator to the metal tube wall. This means either transporting the couplant with the probe, or pushing the probe through a tube full of couplant. Propulsion of a probe along small bore tubing whilst containing the couplant between the probe tip and the tube wall would be difficult, particularly if the probe tip has to be rotated; and

pushing a reasonably tight fitting probe through a tube full of couplant (which is at least as viscous as light oil) is not consistent with rapid inspection.

Nevertheless, **Furlan et al**(1982) have demonstrated an internal probe with a rotating ultrasonic reflector to sweep the beam, and **Tomlinson**(1985) has experimented with the dual effects of IGA on ultrasound: velocity reduction (measurable by time of flight) and attenuation increase. **Temple** has analysed the effect of the degree of grain boundary attack, the grain shape and the presence of water between the grains, on the decrease in ultrasonic velocity through IGA.

### 1.4.2 Older Methods

The most widely used methods in NDT (Magnetic Particle Illumination, Magnetic Dye Penetration and X-ray Radiography) require direct visual inspection of the surface of the test piece and are thus not suitable for use in this application due to the innaccessiblity of the tube surface both internally and externally (because of the small diameter and long length of the tubes, and their close packing in bundles).

### 1.4.3 Methods Under Development

Although steam generator defects are of a predominantly non- ID-surface breaking nature, an internal video inspection probe has been developed (**Hall, Meckstroth & Orphan**,1981). It consists of a circle of optical fibres ends, and 3-D graphical display software.

Secondary side inspection with fibre optic endoscopes has been performed by **Hinton and Chagnon** (1981) but is difficult in practice due to tip steering problems and the very localised illumination associated with fibre optic bundles. The close packing of the tubes into tubing bundles makes 100% OD wall inspection impossible: in fact all that has been demonstrated is a bundle exterior inspection, rather than individual tube inspection. Thus, although this is obviously useful for determining the extent of sludge build up, it is not suitable for defect detection due to the small size of the defects and the poor visibility.

### 1.4.4 Eddy Current Testing

Eddy current testing is the only method certified for operational use which can operate along such long narrow tubes, and has been developed to a sufficient level for use outside a laboratory. It has been in routine use in less critical rod, wire, tubing and non-contact hardness inspection tasks since the 1960's, and thus its capabilitities and limitations are well defined (**Van Drunen and Cecco**,1984). It guarantees 100% volumetric inspection of the tube wall (not just the inspection side surface), and it can do this quickly, typically with a probe speed of $0.5ms^{-1}$. This ability to completely inspect tubing, from the inside through to the outside, is important since most of the defects in steam generator tubes propagate from the outside towards the inside, and thus do not break the surface on the inspection side until the moment of rupture.

### 1.4.5 Method of Operation

Eddy currents are currents induced in a conductor by an adjacent coil, carrying an alternating current, due to the magnetic field associated with the coil current. The induced circulating currents mimic those in the coil in geometry (see fig. 1.5 and also section 2.3 on bobbin and pancake probes), but act so as to induce in the coil a subsidiary voltage which alters the net current flowing by vectorial phasor addition. The phase and magnitude of the net current, and therefore of the coil's impedance, vary in response to disturbances of the flow of the eddy currents, for example when a defect in a tube wall is traversed.

Most instruments at present available display the coil instantaneous impedance as a dot upon an oscilloscope with a storage facility. Defects and other phenomena appear as closed figures representing the locus of values of the test coil impedance, with the origin representing the coil impedance whilst in a nulled condition. These loci can be used by an experienced operator to characterise particular types of defect. In older equipment, the signal was presented as two somewhat more conventional oscilloscope traces, one each for the resistance and reactance, representing their variation with time.

## 1.5 Present Practice

At present when a steam generator in a nuclear power plant has a routine inspection (which is required every 2 years), its contribution to the grid must either be replaced by a fossil fuel-burning power station, or lost altogether. For a typical 600Mw PWR station with Westinghouse model F steam

generators (similar to the type intended for the Sizewell "B" reactor), having one steam generator of the four not in use costs about £600,000 per day.

That part of the primary circuit within the steam generator i.e. the inverted U-tubes and the inlet and outlet channels must be drained. Water washing under pressure then removes any precipitate or obstructions adhering to the tube walls, in order that the probe may be driven along as quickly as possible. The probe had to be inserted by hand into the end of each tube, a job resulting in high radiation exposure. Radiation levels in the steam generator are too high for the lengthy probe pulling process to be done by hand, so a remotely-controlled probe manipulator robot is installed in the chamber into which the primary coolant nozzle is connected (the "water box") the through the lower manway (see fig. 1.3(a)). This task even though done by a technician in a protective suit, heavily shielded against radiation still results in a considerable dose of radiation (prior to the development of the "finger-walker" probe insertion robot, the entire job was done manually, resulting in even higher levels of radiation exposure (**Layfield**,1987)).

The robot allows the probe to be remotely inserted in sequence along the tube ends in the tube plate, and blasted through each of the tubes in turn with compressed air, retracted through each in a more controlled manner, whilst the signal is taped. The cable is gripped between two pinch rollers in order to meter its position with respect to tube support plates within the steam generator, and which cause large signals. This gives a regular check on the progress of the probe during the pulling operation. Switching to and from a mixed signal (see section 2.5) allows improved defect detection in the vicinity of the masking effects due to these same support plates and tube sheets.

Before any steam generator goes into service, an eddy current inspection of all of the tubes is performed. These tubes will have already have been tested by the tubing manufacturer. Subsequently, when a PWR steam generator first begins to be inspected "in service" after it has been running for some 2 years, only 35% of the tubes are tested (this is only true in Europe **De La Pintiere**(1982): in the USA only 33% are tested (**Combustion Engineering**, 1983)). Those tubes that are inspected are distributed evenly across the vessel cross section, and the subset of tubes inspected is a different one on each occasion. This of course required a tube end numbering system and careful remote probe manipulation using closed circuit television, in order to make sure that the same tubes were monitored for degradation on subsequent inspections. When tube degradation was found to be occuring faster than expected, and tube rupturing caused the steam generators in several PWRs in the USA to have to be taken out of service for repair (e.g. Surry Unit 2 and San Onofre Unit 1: see section 2, **Brown**,1983), some utilities began to practice inspection of all of the tubes. Even with inspection of only 35% of the tubing, the expense was considerable. With 100% inspection, any way of reducing the time spent probing tubes, either by simultaneous use of several probes or by on-line defect detection, is highly desirable. It would be especially desirable to be able to perform the signal analysis so that it did not slow down the probing process, and to plug badly defective tubes during one and the same out-of-service period.

In the past, two separate sessions were necessary for an ISI: one to do the inspection and another, several months later, after all of the tapes recorded during the first "outage" have been analysed manually off-line (because the probe travels so fast and also because an operator will usually look at a the

recorded signal from a defect several times before deciding to have the tube plugged). Having two separate outages involved introduces the possibility of errors due to either tube labelling errors or errors in filing the tapes. On occasions the wrong tube has been plugged (**Russell & Richardson**,1983).

**1.6 Why Automation is Desirable** The problems of variations in operator concentration and judgement when watching a display on an oscilloscope in order to analyse several thousand tubes are obvious. Thus the reasons for interest in automating the inspection of steam generator tubing are:

(i) to prevent missed defects

(ii) to be able to analyse tubing at the same speed as the the probe can travel through it, (preferably concurrently with the probe drawing operation)

(iii) to fully exploit the information available especially when using a high definition rotational scan probe and/or multifrequency excitation.

(iv) to attain objective, repeatable and consistent detection and classification of defects, both by type and by depth, even when those defects coincide with other things which produce much larger signals, such as tubesheets or dents.

(v) to have a record of the signals from defects recorded which where the timewise segments of signal not of interest have been removed in order to compress the storage required to manageable proportions.

The difficulty experienced by operators in discriminating manually between iga and tube wall thinning due to fretting has been a major motivating factor in the development of automatic detect classification by type. This is because the two phenomena yield signals which appear to be virtually

identical, and it has been thought for some time that if these defects can be told apart from their eddy current signal loci at all, then it will only be by the use of a computer. The importance of resolving this problem lies in the fact that the iga is much more of a problem than thinning due to the greater likelihood of the origination of cracks from dissolved grain boundaries, and also because it is more difficult to assess the depth of penetration through wall.

### 1.7 Meaning of Automatic Signal Analysis

From the human point of view, the first signal analysis task required is the segmentation of the significant portions of the signal from the insignificant parts. This is then followed by the second task, classification: a review of each the "significant" sections of signal, in order to determine the type and depth of the defect indicated by the signal, based upon the shape, size and slope of the associated locus in the impedance plane.

From the point of view of system categorisation, there are in fact three types, corresponding to the three distinct phases in signal analysis as perceived by most authors. These are:

(i) automatic logging systems

(ii) automatic defect detection systems

(iii) automatic classification systems.

### 1.8 Automatic Tube Logging Systems

More papers have been published about systems which fit into this first category

than on either of the others. These systems do not really perform signal analysis, though the title of the papers frequently imply the opposite. They are included here because their titles and keywords cause them to appear in literature searches for pattern recognition/signal analysis within eddy current testing , and because signal logging seems to be perceived as equivalent to signal processing within NDT literature.

The Electricitie de France's **Estelle** system described by **Baumaire and Meugrey**(1987) is typical of an expanding class of system which uses a PC AT or similar specification desktop computer to log the eddy current signals to disk, perform rudimetary processing and then upload the tube data files to a much more powerful computer where more powerful signal processing techniques may be perfromed off-line (e.g. FFT spectral analysis based filtering). In the case of the Estelle system, the host is an IBM 3083, and the system is particularly for recording the very large amounts of data (Megabytes per tube) generated by the inspection of tube rolls with a rotating probe. Other systems which have been developed recently include one by Combustion Engineering (**McMillan**, 1987) using an HP9836 and 55 Mbytes of disk space, a portacabin based one by Laborelec (**Dobbeni**,1987) and the CF2000 by Framatome (**Poulet and Grozellier**,1987) using an HP2000 to log a rotating probe to an optical disk. Phyically smaller (luggable) systems have also been developed with less data storage capacity but not necessarily less processing speed or sophistication, such as the Failure Analysis Associates Smarteddy (Texas Instruments PC running MS-DOS with 1 floppy disk), Zetec's MIZ-18 (**Houserman et al**,1983) using a 68020-based HP200 series computer and a cartridge data logger, and Lufthansa'a PC-based Archimedes (**Schur**,1987) for bolthole inspection for the Airbus 300.

## 1.9 Automatic Segmentation

Automatic segmentation is desirable because of the increasing use of rotating probes, and the increase in the amount of signal to be analysed in comparison to bobbin probes. Until recently, almost all published work on segmentation has been directed towards bobbin probes, with the exception of work done by **Macleod** (1982).

In most existing work upon defect detection or "segmentation" is done by reference to the amplitude or magnitude of a signal or signals. An amplitude threshold is applied to the x or y axes (or both) to detect defects, treating each axis in isolation from the other, or a ring is superimposed around the null point. In the most advanced systems, sloping lines boundaries are used. Systems using amplitude alone are described first, then those which make some use of the phase information available. Work upon this is described in chapters 4 and 5.

At the Chalk River Nuclear Laboratories in Canada, **Wells**(1981), **Sharp**(1982), and **Jarvis and Cranston**(1982) have developed a similar system for inspection of the 10mm dia. tube in CANDU-type reactors. The system completely controls the insertion and withdrawal of the probe, into one tube after another, in an orderly sequence. It was the first to be tested on-line in a real in-service inspection(1981). It worked as follows. During an anomaly (or defect), ie a y-signal excursion unaccounted for by the preprogrammed positions of the tubesheet and tube support plates, the x and y samples are stored to tape, along with two quantities derived from analogue filters, ybar (the recent signal mean) and $y_{pp}$, the signal peak to peak. The use of ybar allows discrimination against gradual changes in signal level (such as

that due to probe heating) which do not constitute defects. The group have published complete specifications of their library of subroutines which together accomplish the above task (i.e the input and output parameters for each routine, although they do not actually give the Fortran listings).

In the UK, **Russell and Richardson**(1983) developed an offline system with the acronym **DART** (Data Reducing Eddy Current Tester)(see also **Whittle**,1985), which categorised any signal by amplitude, in terms of the number of coincident 1.5mm dia. holes (between 1 and 8) required to produce a similar size of signal (a widely-applied plugging criterion is that any tube must be plugged if it contains a defect producing a locus bigger than that of 4 holes).

The size of the locus due to a defect is actually more strongly related to defect volume than depth. It is the phase of the locus which varies with depth. Since it is the through wall depth of a defect that is most important in deciding whether or not to plug a tube, the use of an amplitude criterion, though simpler to implement than a phase sensitive criterion is not the best method. These matters are discussed in more detail in chapter 3.

**Hocking**(1985) developed an on-line version which allowed failed tubes to be plugged on the spot. The system can plot the frequency of defect occurrence along the length of the tubes, enabling diagnosis of the cause of batches of defects. Such symptoms might be that all of the defects lie in a plane which coincides with the top of a tube support plate, implying a problem local to that plate. Also, it can sense the approach of the expanded tube/tube sheet section, thus speeding the inspection by preventing the probe overshooting the tube-end and simutaneously preventing wear and tear on the probe and cable. Sensing the tube roll (see fig. 1.4) also allows the probe position

measuring equipment to be re-registered if there has been any cable slippage due to obstruction. Samples from the probe are only stored to tape for sections of tubing where the amplitude threshold is exceeded.

Commercially available, portable systems all tend to fall into this class of systems which use horizontal or vertical lines, or rectangles in the impedance plane. Examples of such instruments are the Hocking AV100, the Nortec NDT-19 and NDT-25L and the Failure Analysis Associates SmartEddy. This reflects the differing priorities in the design of these instruments, i.e. that they must segment on line, and be as portable as possible. The segmentation is frequently carried out by analogue comparators, even although the instruments contain microprocessors (which are there primarily to imitate a storage CRT without the normal weight and power consumption penalties). Certainly, they do not have the processing time available between samples and screen updates to perform anything more complicated than a simple comparison.

**Levy, Caussin and Dombret**(1982) at Vincotte in Belgium were the first to describe a phase and magnitude based segmentation method. They have superimposed concentric warning and rejection threshold rings in the impedance plane centred upon the null point. By subdividing these rings into sectors, the authors use phase to categorise loci as from internally and externally originated defects, and their through wall depths (and as such, the system might be regarded as not more than a segmentation system). The unit that they used had an eddy current unit with 3 differential channels and an absolute one in order to mix out undesirable signals. When a tube sheet or support is encountered, the signal being tested for amplitude is automatically

switched to a mixed version. The entire signal segment with a defect is stored once the threshold is exceeded, along with the magnitude and phase at the extremity of the locus.

As is explained in Chapter 3, the relationship between phase and depth is compressed in a way not easily describable by the mixing of channels of different excitation frequency (**Selig & Neumaier**,1983). Furthermore, the relationship between phase and depth is also dependent upon the type of defect, and thus for accurate depth assessment the defect type must be determined before deciding upon its depth. It is surprising that **Levy et al** mention nothing of this, or of how their categorisation system copes with this, or to what extent the categorisation borders are changed by using a mixed signal. Some attention is paid to this matter in Sections 2.5 and 5.3.

**Stepinski**(1987) has described a segmentation system designed to "bolt on" to an existing system, which uses a microprocessor to compare each sample's position against the boundaries of a user adjustable rectangle which is contrained to have with one corner aligned with the null point, and also with two radii which emanate from the null point, defining a sector. The system's main innovative feature is that it operates in real time. It can process (i.e. digitally filter and test against the threshold), 500 samples per second. The comparison involves a number of multiplication operations to test for $y \geq mx$ which have only become feasible in stand-alone systems at these sampling rates in real time since the advent of microprocessors with built-in multiply instructions.

The Zetec MIZ-18 (**Houseman**,1983,1987)has the most advanced segmentation methods of any commercially available system, being capable of

superimposing a parallelogram with a major axis lying at any angle in the impedance plane, and of qualifying this with the maximum rate of change of signal (the furthest distance between consequitive samples). It can also use a quadratic fitted to the signal from good tubing at the tube roll section as the threshold, because of its ability to closely track the probe position by re-registering the position of the probe each time it passes the upper or lower edge of a tube support plate.

A rather specialised scheme is that of **Cleveland, O'Brien and Wereley** (1986), who have developed an anomaly detector for bobbin probes passing through support plates which detects any lack of symmetry about the signal. They did this by testing, firstly on a gross level, that there were exactly two peaks in the signal magnitude, one on the way in and one on the way out, and secondly on a fine level, by testing for a minimum level of correlation between the two lobes in opposite directions (due to the differential probe).

### 1.10 Automatic Classification Systems

There exist six other published sets of work on automatic classification applied to eddy current testing, apart from the work by **Macleod**(1982,1982,1983). Most of the systems are intended for use with a differential bobbin probe, with the exception of the system devised by **Neumaier**(1983).

**Hsiung and Cox**(1972) were the first to address the subject. They used simple peak size and width features applied to the y signal to discriminate between circumferential interior and exterior (termed internal diameter("ID") and outer

diameter("OD") slots.

**Doctor and Harrington(1980)** have discriminated between spark notches, elliptical wastage and thinning (tube fretting), and have been able to correlate the features with the depth of thinning using a set of adhoc features, namely geometrical features of the impedance plane loci; autocorrelations with various lags of the time series, and moments; parameters of the spectral peak and spectral moments; and some measures of relative power in the x and y signals. Classification was variously by Bayesian, linear and k-nearest neighbour classifiers, and by k- means clustering (these are all explained in Chapter 2). A very similar paper has also been published by the same authors upon the use of pattern recognition to discriminate between acoustic emissions from defects and spurious sounds (**Doctor and Harrington**,1980).

**Neumaier** (1983), at the Institut Dr. Forster in Germany, has used a differential rotating surface probe, and has differentiated between ID and OD notches, and between elliptical wastage and small diameter (1mm) holes. For the first pair, he used the y signal interpeak time (which is a measure of the phase providing that the probe wobble signal has been rotated to lie along the x-axis and that there is sufficient phase spread; see section 2.5). For holes and patches of wastage, he used the number of peaks in the entire signal segment relating to the defect. It is not surprising that this worked, because the number of peaks is just a measure of the axial length along the the probe's helical path, and due to the nature of wastage (or fretting) it will always be of greater extent than a small hole.

**Satish** (1983), at Colorado State University (see also **Udpa and Lord**(1984), working with **Lord's** 2-dimensional finite element bobbin probe simulation(1981), used Fourier Descriptors to discriminate between eight classes of defects and other phenomena. Ida and Udpa also attempted to incorporate Fourier Descriptor-based pattern recognition in Zetec's DDA-4 system, the prototype MIZ-18, which was at that time based upon a Hewlett Packard 9836 computer (**Brown**,1984). A full discussion is given in chapter 3.

One particular type of nonlinear classifier (termed an "Adaptive Learning Network") with a feature product term pruning mechanism has been repeatedly advocated in automated NDT by the staff of Adaptronics (see for example **Mucciardi and Dau**(1979), a firm set up to manufacture computer systems dedicated to ALN deduction. The method is discussed in detail in section 3.7.4.

**Cleveland, O'Brien and Wereley** (1986) have used 65 ad-hoc time domain correlation, geometric shape and energy features to discriminate between 16 holes of 20, 40, 60 or 80% through wall depths in 4 different diameters, each recorded in one of 17 positions with respect to a tube sheet. The requirements of such a large number of features in terms of samples imply that their results are questionable. The requirements are discussed in Chapter 3.

## 1.11 Scope of Thesis

A collaborative project between the Department of Electrical Engineering and Electronics at the University of Glasgow and the NDE group at Babcock Power Research Centre in Renfrew began in 1980 with a feasibility study on the application of pattern recognition to inspection of 316 stainless steel tubing

with a rotating probe (**Macleod**,1983) following suggestions in several reports to the effect that this might prove fruitful (**Davis**,1979 and 1980). The aim of this research at its inception was to classify various types of defects using the impedance plane loci from existing single frequency eddy current equipment. At the start of the present author's research, the advent of multifrequency equipment, and the increased difficulty it brought with it in signal comprehension, stimulated a shift towards a multifrequency re-logging of the defect set and redesign of the feature set. In addition, previously untested types of defects were to be manufactured and a means of discriminating defects from clean tubing devised. Unfortunately, defect detection using the particular multifrequency eddy current (MFEC) equipment available was found to be impossible on stainless steel, and a change to a different tubing material (Inconel® 600 ) took place at a fairly advanced stage. This complied with a change in the sponsoring bodies' favoured material for steam generator tubing, due to reported problems with stainless steel tubes in Westinghouse PWR steam generators.

This thesis therefor describes the application of pattern recognition to single frequency helical scan eddy current signals from thin walled tubing, initially to inner wall manufacturing defects in finned 316 AGR heat exchanger tubing, then latterly, and of great topical interest, to defects on the outside of Inconel 600 tubing, logged using a multiplexed 4 frequency excitation eddy current system.

**Chapter 1** has explained how and why eddy currents are used in NDT, why eddy current techniques are particularly necessary for steam generator tubing, and why automatic inspection is desirable on economic and safety grounds.

The ways in which the eddy current signals from steam generator tubing can be automatically processed to detect and classify defects have been reviewed. **Upda & Singh** (1986) have also published a review paper upon digital signal analysis applications though it was concerned with in NDT in general and thus was not so thorough with respect to this particular application. Systems capable of automatic logging, segmentation and classification have been catalogued separately, since frequently systems do not perform all of these functions. Commercially available systems have also been included.

The remainder of this thesis is arranged as follows. The parts of the pattern recognition system are introduced in the order in which the signal proceeds through them: **Chapter 2** deals with the sensor, and **Chapter 3** with the preprocessor, feature extractor and classifier.

**Chapter 2** reviews eddy current inspection in detail, with particular emphasis upon thin walled tubing, and discusses the choice of probe type and frequency. The deficiencies in analytical and numerical simulations of the interaction between defects and eddy current distributions are shown to be the main reason for tackling automatic classification with statistical pattern recognition rather than by comparison with synthesized prototype signals.

**Chapter 3** contains a review of the pattern recognition methods used in the experimental chapters, 4 and 5, along with a description of original experimental work on several topics. This relates particularly to work upon Fourier Descriptors (a type information preserving parametric featureset), upon a non- parametrically based class separability measure which does not impose the usual restrictions upon the form of the pattern point class distributions, and upon adaptive learning networks which have been applied

so frequently and erroneously by certain authors.

**Chapters 4** and **5** contain the bulk of the experimental results. The former is a continuation of the work described by **Macleod**(1983) centring upon the development of a segmentation technique and upon the application of different features and classifiers from those originally applied. The upper limit on the correct classification rates is also assessed.

In the course of the work for **Chapter 5**, all of the software had to be upgraded in order to handle the signals generated by multifrequency inspection. A new set of defects have been logged, and a new set of loci now correspond to the defect types, which are also now categorised by depth. Plots of the signals by type, depth and frequency of probe excitation have been made and are discussed. A segmentation method capable of dealing with the more difficult background conditions has been developed, and some feature extraction and classification has been attempted.

**Chapter 6** contains conclusions drawn from the results as a whole, and suggests the manner in which work should be continued. The feasibility and method of construction of a stand-alone automatic inspection systems is considered, along with specific suggestions for theoretical work that has either been outside the scope of this thesis, or impossible thus far for practical reasons.

## 1.12 References

1. **Mundis, J.A. and De Young, G.W.** "Solutions to NDE Problems in PWR Steam Generators - An Overview", pp 93-98, Proc. 2nd ASME Conf. on Quality Assurance in the Nuclear Industry, San Diego, 1982.

2. **Layfield, F.**,"THe Sizewell B Public Enquiry",,Vol II, Section 4,Chapter 22,HMSO, 1987.

3. **Glasstone,S. and Sesonske,A.**,"Nuclear Reactor Engineering",pp30-31,740-748, and pp449-452,Van Nostrand Reinhold, 1981.

4. **Birkle, A.J.**,"PWR Steam Generator Inservice Inspection", IMechE C202/76, p. 61-65.

5. **Tomlinson, J.R.**, "Detectability of Intergranular Attack in PWR Steam Generator Tubing",pp95-114,20th Ann. Brit. Conf. on NDT, Aug. 85.

6. **Thomson, R.B. and Thomson, D.O.**, "Ultrasonics in Nondestructive Evaluation",p.1716-1755, Proc. IEEE, Vol.73, No.12, December 1985.

7. **Spinks, A.H.**, Central Laboratories (NDT), ministry of Works, Auckland, Lower Hutt, New Zealand. Private Communication, Sept.,1987.

8. **Broere, C., Ketelaar, K. and Lodder,M.**, "Practical Experience on Crack Detection of Tube Bundles in the Tube Sheet Area by Means of Ultrasonics", Paper SS09/05,Proc. 4th Euro Conf. NDT, London, 13th-17th Sept., 1987.

9. **Furlan, J., Soleille G., Mogavero R. and Chalaye, H.**,"Ultrasonic Testing of Steam Generator Tubes Using An Internal Probe", pp110-112, 20th Ann. Brit. Conf. on NDT, Aug. 85.

10. **Temple, J.A.G.**,"Calculations on the Feasability of Ultrasonic Inspection of Steam Generator Tubing in Pressurised Water Reactors",AERE-R-10980,HMSO,April,1985.

11. **Hall,G.D.,Meckstroth,E.A. and Orphan V.J.**,"Optical Scanner System for

Internal Inspection of Steam Generator Tubes",EPRI-NP- 1944,July,1981.

12. **Hinton, M.S. and Chagnon, C.W.**,"Visual Inspection Equipment for the Secondary Side of Steam Generators", EPRI-NP-1859, May,1981.

13. **Van Drunen, G. and Cecco, V.S.**, "Recognising Limitations in Eddy Current Testing", Fig. 6, page 13, NDT International, Vol. 17, No. 1, pp. 9-17, Feb., 1984, Butterworth.

14. **De La Pintiere, L.** "Experience and State of Art of Steam Generator Tubing Examination in France", pp 113-116, Proc. 2nd ASME Conf. on Quality Assurance in the Nuclear Industry, San Diego, 1982.

15. **Combustion Engineering Power Systems**, "Remote PWR Steam Generator Inspection Service", Promotional Brochure, 1983.

16. **Baumaire, A. and Meurgey**,"Automatic Signal Analysis of Rotating Probe ECT Inspection of SG Tube",Proc. 4th European Conference on NDT, London, Sept 1987.

17. **McMillan, John D.**,"Imaging and Beyond - What to do with the Data?",Proc. 4th European Conference on NDT, London, Sept 1987.

18. **Dobbeni,D.**,"Steam Generator Inspection Engineering",Proc. 4th European Conference on NDT, London, Sept 1987.

19. **Poulet, J.P. and Grozellier, M.**,"Computerized Eddy Current Inspection and Data Analysis", Proc. 4th European Conference on NDT, London, Sept 1987.

20. **Houserman,H., Lamb.,L., and Kitson, W**,"Digital Analysis of Steam Generator

Tubing",pp465-480, Proc. 3rd Conf. on Q.R.N.D.E.,Zurich, 1983.

21. **Houserman,H.E.**,"The Application of Improved Eddy current Inspection Techniques For Nuclear Steam Generator Tubing",Proc. 4th European Conference on NDT, London, Sept 1987.

22. **Schur, F.**,"Computerised Evaluation of Eddy Current Signals", Proc. 4th European Conference on NDT, London, Sept 1987.

23. **Wells, N.S.**,"An Automated Eddy Current In Service Inspection System for Nuclear Generator Tubing", Proc. 13th Symp. on NDE, pp419-434, San Antonio, 1981.

24. **Sharp, F.L.**, "Semi- automatic Eddy Current System for heat exchanger Inspection", AECL Report 7258, Chalk River Nuclear Labs., Feb., 1982.

25. **Jarvis,R.G. and Cranston, R.J.**, "EDDY - A Fortran Program to extract significant Features from Eddy Current Test Data - the Basis of the CANSCANsystem." AECL Report 7779, Chalk River Nuclear Labs., Sep. 1982.

26. **Russell, M. and Richardson, K.P.**,"Automatic Inpection and Evaluation of Power Plant Condenser Tubing",pp 25/1-14, CEGB internal Report.

27. **Whittle,D.W.**,"Evaluation of the CEGB SWR(South West Region) SSD (Scientific Services Department) designed "DART" System for Eddy Current Inspection of Condenser Tubing", Babcock NDT Report No. 85/4/5, April '85.

28. **Hocking Associates** Private communication concerning D6 Specification, Hocking Associates, St. Albans, 1985.

29. **Levy, R.**,"Eddy current Inspection of Heat Exchanger Tubes: Partl:

Multifrequency Eddy current Testing",translated from Fils,Tubes,Bandes, Profiles, 1982,15(92),13-15.

30. **Caussin,P. and Dombret,Ph.**,"Eddy current Inspection of Heat Exchanger Tubes: Part 2: Automatic Real time Analysis",translated from Fils,Tubes,Bandes, Profiles, 1982,15(92),13-15.

31. **Selig and Neumaier, P.**,"Testing of Steam Generator Tubing by the Eddy Current Method - Improvement of Defect Sizing",pp1-6, Proc. 3rd ASME Conf. on Quality Assurance in the Nuclear Industry, 1983.

32. **Stepinski, T.**,"Real time Signal Analysis in Eddy Current NDT Equipment", Proc. 4th European Conference on NDT, London, Sept 1987.

33. **Cleveland D., O'Brien L.J. and Wereley, N.M.**,"An Automated Aid for Isolating Defects in the Heat Exchanger Tubing of Nuclear Steam Generators",ISA Transactions, Vol. 25, No.3, p.61-71, 1986.

34. **Hsiung, C. and Cox,C.W.**,"Pattern Classification in Scan-type Non destructive Tests", Int. J. NDT, 4,231-247,1972.

35. **Doctor, P.G. and Harrington, T.P.**, "Analysis of Eddy Current Data Using Pattern Recognition Methods", pp137-139,IEEE 5th Conf, on Patt. Rec., Miami, 1980.

36. **Doctor, P.G. and Harrington, T.P.**, "Acoustic Emission Analysis Using Pattern Recognition Methods", pp1204-1207, IEEE 5th Conf. on Patt. Rec., Miami, 1980.

37. **Neumaier,P.**,"Testing Heat Exchanger Tubes using Eddy Current Techniques

with Computerised Signal Analysis",Br.J.NDT, 25,(3),223-237,Sept,1983.

38. **Satish, S.R.**,"Parametric signal Processing for Eddy Current NDT", Ph.D. thesis, Colorado State Univ.,1983.

39. **Udpa, S.S. and Lord, W.**,"A Fourier Descriptor Classification Scheme for Differential Probe Signals",Materials Evaluation/42, pp1136-1141, Aug. 1984.

40. **Lord, W.**,"Development of a Finite Element Model for Eddy- Current NDT Phenomena", EPRI-NP-2026, Sept.,1981.

41. **Brown, S.**, EPRI 1984 Annual Research Report, Section 19, EPRI NDE Research Centre, 1984.

42. **Mucciardi, A.N. and Dau, G.J.**, "Development of Adaptive Learning Network for Eddy Current Signal Analysis". "Nondestructive Evaluation Program: Progress in 1979", EPRI NP- 1234-SR, Special Report, December, 1979.

43. **Macleod, J.E.S.**,"Automatic Inspection of Tubes Scanned By A Rotating Eddy Current Probe", Signal Processing, Vol. 1., No.5,pp445-450,1983.

44. **Macleod, J.E.S.**, "Feasability Study on Computer Sorting of Flaws in Steam Generator Tubes Scanned by a Rotating Eddy Current Scanner", Babcock Power Research Report No. 07/82/46, July 1982.

45. **Macleod, J.E.S.**, "Pattern Classification in the Automatic Inspection of Tubes Scanned by a Rotating Eddy Current Probe", 6$^{th}$ IEEE Conf. on Patt. Rec., Munich, Oct. 1982.

46. **Davis, T.J.**,"Advances in Multifrequency Eddy Current Instrumentation", I. Mech. E. C37/7, pp257-265, 1979.

47. **Davis, T.J.,**"Multifrequency Eddy Current System for inspection of Steam Generator Tubing", EPRI-NP-1621, Nov., 1980.

48. **Singh, G.P. and Udpa,** S., "The Role of Digital Signal Processing in NDT",pp177-189, NDT International 19 (3), 1986.

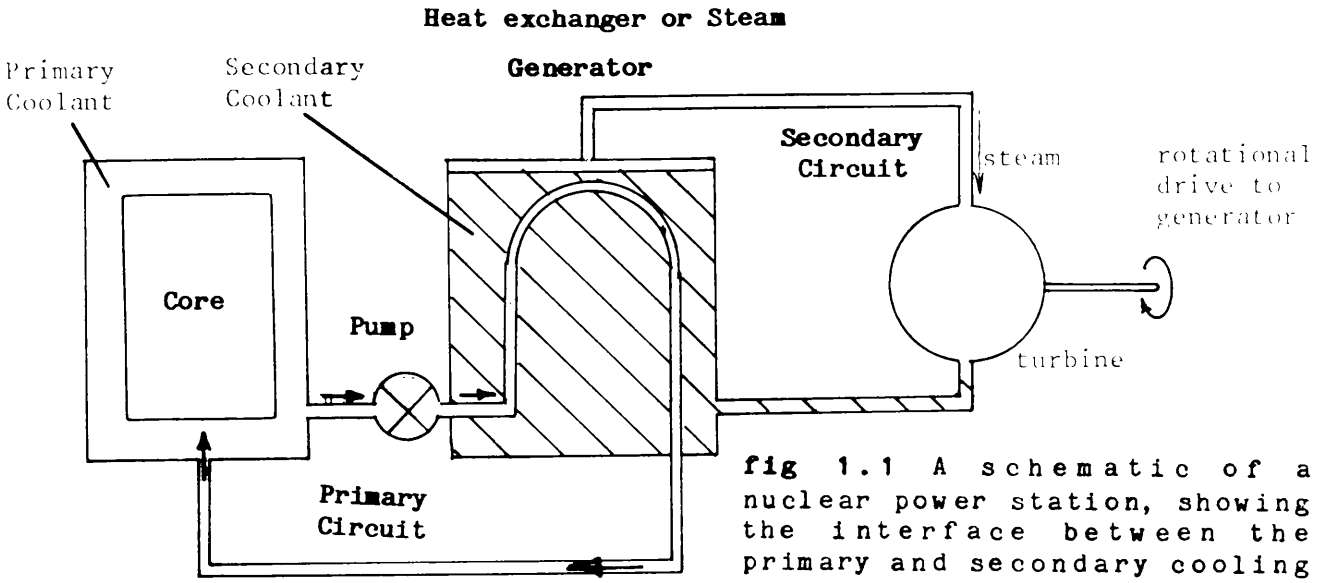



**fig 1.1** A schematic of a nuclear power station, showing the interface between the primary and secondary cooling circuits.

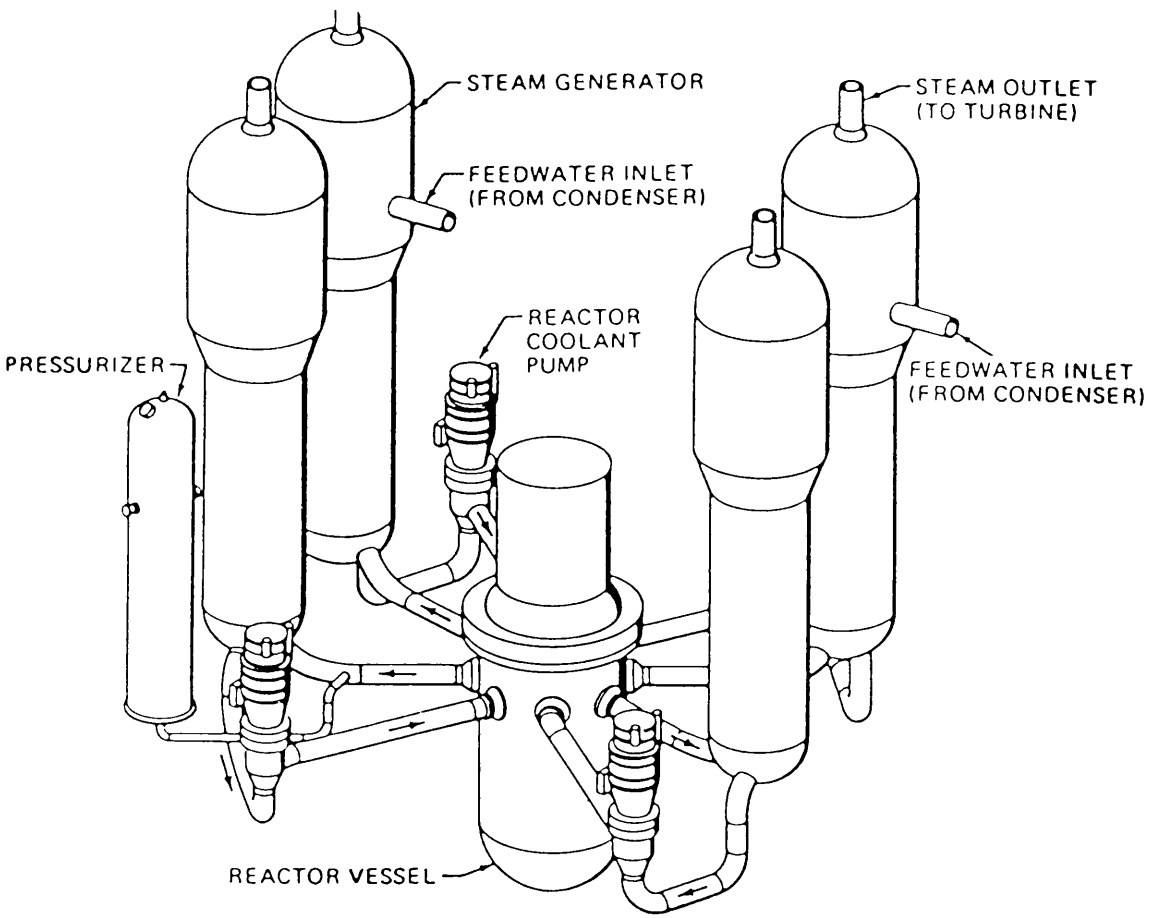



**fig. 1.2** A practical PWR primary circuit, showing the 4 components: the reactor, the pressurizer, the 4 pumps, and the 4 steam generators. This design will be used in the Sizewell "B" reactor, if it is built.

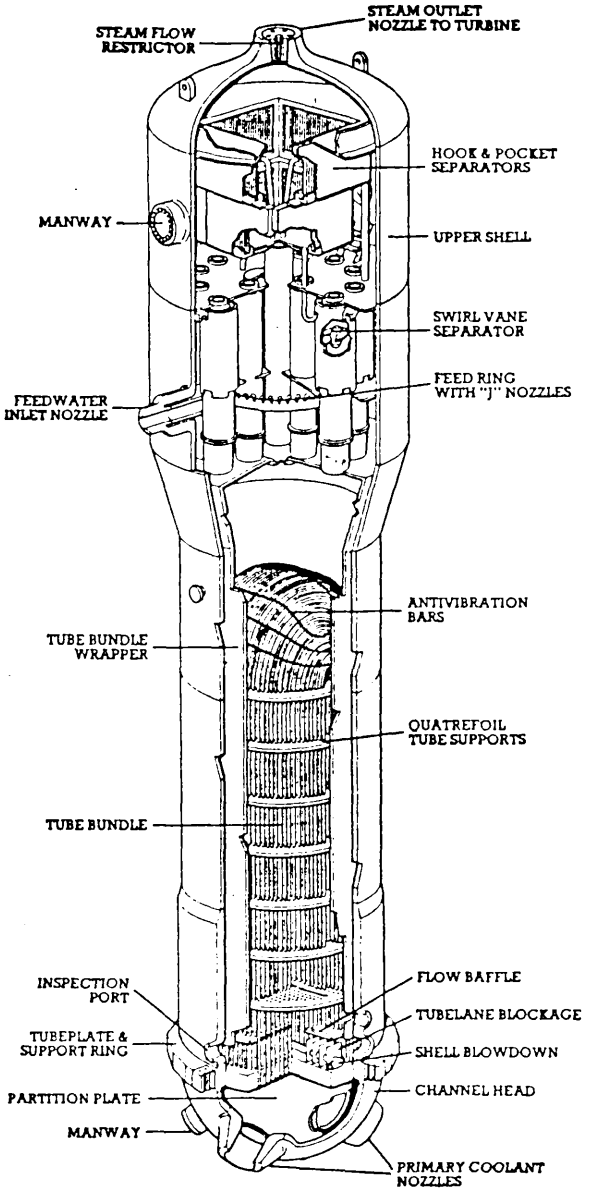



**fig. 1.3(a)** A cutaway diagram of a Westinghouse Model "F" steam generator (from **Tomlinson**, 1985)

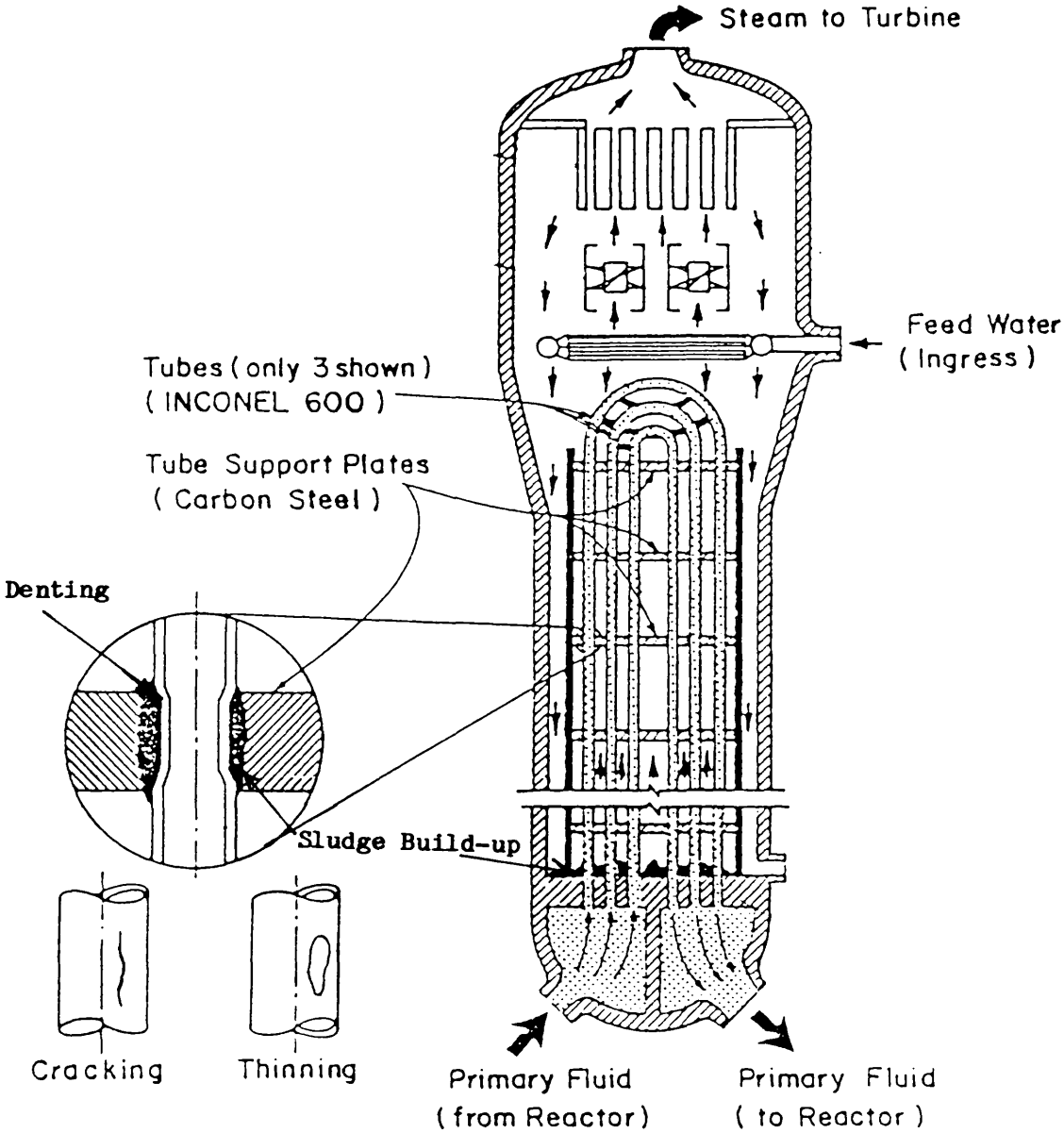



**fig. 1.3(b)** A cross sectional view through a steam generator, showing water flow and the likely sites of defects (from **Satish**, 1983).

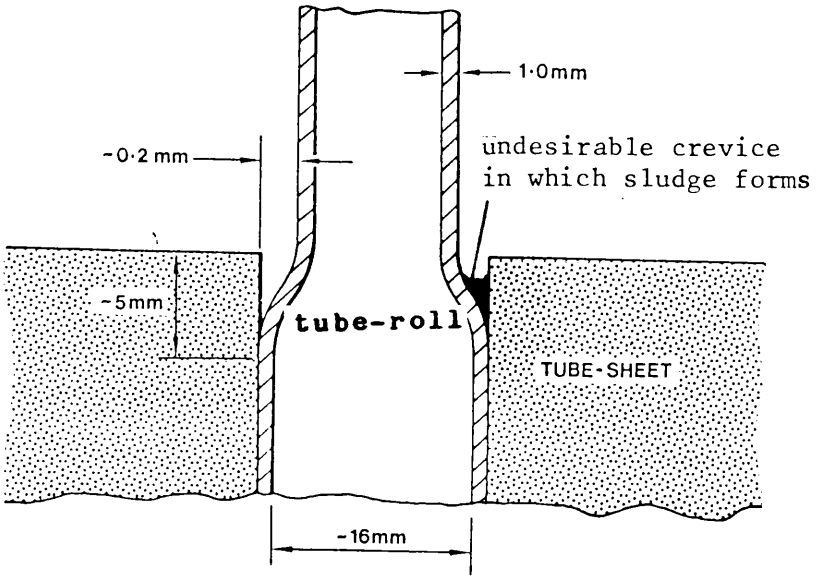



**fig. 1.4** The expanded diamter tube area at the tubesheet. A sludge forms in the crevice and sometimes in as much as a 200mm. layer, encouraging local corrosion of the outer tube wall (from **Tomlinson**, 1985).

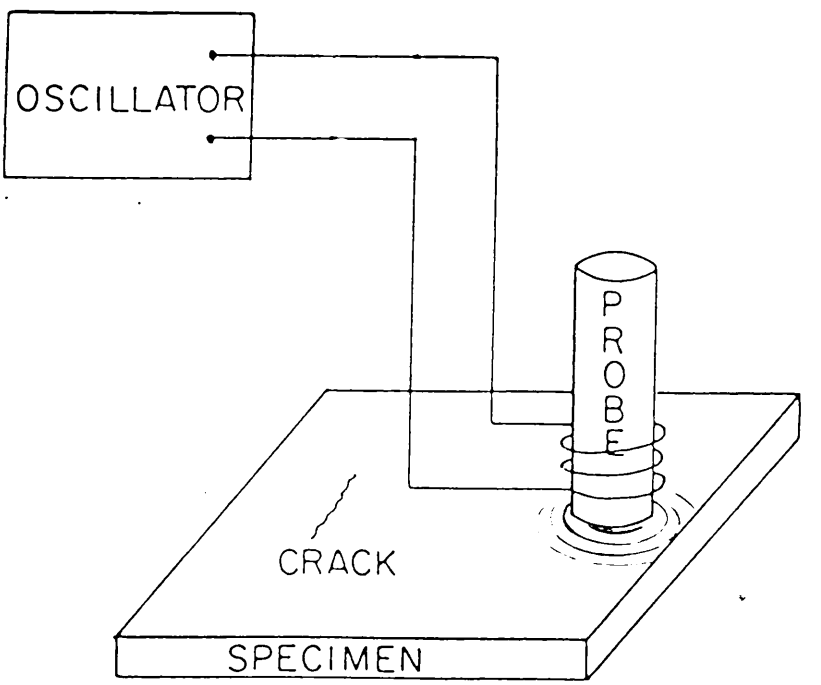



**fig. 1.5.** A schematic of a ferrite cored coil, and the mimicking of the coil current flow by the eddy currents induced by the flux linking the coil and the metal surface.

# 2 Chapter 2: Eddy Current Tube Testing

## 2.1 Introduction

This chapter summarises the phenomena involved in eddy current testing and then explains the factors which influence the choice of probe and operating frequency (or frequencies) especially where complete volumetric inspection of thin walled tubing is the goal.

Theoretical work upon analytical and numerical (finite element) simulation of the behaviour of eddy currents in the vicinity of defects is reviewed, in order to show that a statistical approach to determining the shape for particular defects is necessary because the quantitative effects of defect shape and field shape upon the shape of the associated impedance plane loci cannot yet be calculated.

Work on mixing, the subtractive combination of loci derived from simultaneous (multiplexed) inspection at several frequencies, is also reviewed, and some original work on the development of computationally optimised mixes is introduced.

## 2.2 Basic Principles

Eddy current testing is primarily a comparative method of testing: an eddy current instrument is first of all calibrated using defects of known depth and geometry in a section of material identical to that about to be tested. Then, any indications appearing during testing are characterised by comparison

with the calibration signals. Alternating current in a coil causes an alternating magnetic field, with the flux linking the coil. If the coil is adjacent to a conductor, eddy currents are induced near to the surface. The current flows in a circular manner, mimicking the circular cross section of the coil. The eddy currents produce a secondary magnetic field which acts so as to oppose and partially cancel the original field, thus modifying the original coil current by vectorial addition of an induced current. The probe coil impedance, the ratio of driving voltage to current flowing, is therefore dependent upon the intensity and phase of the eddy currents in the inspected metal wall. The magnitude and phase of the eddy currents vary with depth into the tube wall, and depend upon the material thickness, conductivity and permeability of the material, the probe excitation frequency, the extent of probe "lift-off" from the surface, and the localised current-disrupting effects of defects. A crack is in effect a discontinuity in conductivity which causes the currents to flow around and underneath it.

## 2.3 The Skin Depth and Phase Lag Effects

At any given frequency, current decays exponentially beneath the surface, according to the relationship:

$$I_x = I_o . e^{-x/d} . e^{-jx/d}, \qquad (2.1)$$

where $I_x$ is the current density at depth x mm., $I_o$ is the surface current density (inspection side), and d is the skin depth as defined below. The term "skin depth", actually refers to the depth by which the current density has fallen to 1/e of its surface value, and can be shown to be

$$d=50(\rho/f.\mu_r)^{1/2} \text{ mm.} \qquad (2.2)$$

where $\rho$ is the material resistivity in micro-ohm.cm, f the frequency in Hz, and $\mu_r$ the (dimensionless) relative incremental permeability . Thus with rising frequency, current increasingly concentrates near to the surface of a conductor.

From equation (2.1), the nature of the skin effect is such that the subsurface current phasors lag the surface ones. The phase lag $\beta$ is linear with increasing depth, varying according to

$$\beta(\text{rads.})=x(\text{mm.})/d(\text{mm.}) \qquad (2.3)$$

The coil impedance is proportional to the vectorial sum of the eddy currents flowing at all depths in the metal. It is entirely possible that a small inspection-side surface defect and a large subsurface or outer surface defect will cause an impedance change of the same magnitude. Under such circumstances, discrimination between them is only possible by reference to the phase of their impedance loci.

Although eqns. 2.1-3 are widely quoted (**Hagemaier**,1983), they are only strictly applicable when the incident magnetic field is planar i.e. uniform and parallel to the metal surface. For a flat plate of infinite extent, production of such a field would require an infinite sheet of current. For the inspection of tubing, an infinitely long solenoid would induce a magnetic field satisfying the same requirements at the tube surface. In practice, the field from a bobbin coil (which is after all a short solenoid) is sufficiently close to this ideal for the standard skin depth equations to be applicable. When these equations are quoted in EC NDT literature, it is seldom stated that they apply only for bobbin coils in tubing. The next section (2.4) is concerned with the analogous equations for pancake probes, where the induced field at the

metal surface is not planar.

### 2.4 Probe Geometries: Bobbin vs. Surface Probes

For tubing inspection there are basically only two practical probe geometries:

(a) with tube and probe concentric (bobbin probe);

(b) with the probe axis perpendicular to the surface and the probe tip in contact with the metal (surface or pancake probe).

It has previously been mentioned that the skin depth and phase lag expressions assume that the magnetic field is planar, and that the conductor is infinitely thick. For bobbin coils where the coil is usually not by any means "an infinitely long solenoid", and even in the case of thin walled tubing where the wall thickness is less than the 3 standard skin depths that it should take for the field to decay to a negligible intensity, eqns. 2.1-3 hold approximately. For surface probes, the field is not at all planar (see figs. 2.1 and 2.2). However, a simulation resulting in a contour map of complex valued impedance sensitivity (**Dodd et al** 1969: see fig 2.2) has shown that the current density still decays in a broadly exponential form, but more rapidly than for the bobbin coil. The depth of penetration is no longer purely a function of frequency. The maximum achievable depth of penetration

$$d_{max}=D_o/4 \qquad (2.4)$$

where $D_o$ is the mean coil diameter(**Cecco, Van Drunen and Sharp**,1983). This is the depth below which the current flow represents less than 1% of the total current flows (see fig. 2.2), and holds regardless of how low a frequency is used. It is analogous to approximately 3 skin depths below the inspection

side for a bobbin probe.

The effective diameter of the interrogating field is given by another rule of thumb(**Cecco et al**,1983):

$$D_{eff}=D_o+4d \qquad (2.5)$$

The same authors have stated that the frequency at which the loci from defects at the inner or outer surfaces of a tube (usually referred to as ID and OD defects respectively in the literature) are separated in phase by 90° in the impedance plane is given by:

$$f_{90}=k\rho/t^2 \text{ kHz} \qquad (2.6)$$

where t is the tube wall thickness in mm, and k is a coefficient dependant upon the ratio t/d: the ratio of the tube wall thickness to the standard skin depth as defined in eqn. 2.2. Typically, at $t=1.1d$, k is 3, and for $t=0.8d$, k is 1.6. This formula is empirically derived (though based upon eqns.2.1-3) but makes allowance for the disturbance of the field due to the finite depth of the conductor (a thin sheet in effect) and the non-planarity of the magnetic field associated with a surface probe.

For Inconel 600 tubing, with 1.1mm thick walls, eqn. 2.6 predicts $f_{90}$ to be 130kHz. At 100kHz, the phase spread between ID Pilger noise and a 56% through wall OD circmuferential notch is about 70° (see fig 5.41). It is not until the frequency reaches 220kHz that the part of the defect loci is perpendicular to the Pilger locus, and it really not until 600kHz that the defect locus is perpendicular along all of its length. This can be explained to some extent by the fact that all real defects have some depth, so the phase separation expected tends to be somewhat less than would be achieved with

purely OD and ID phenomena.

### 2.5 Phase Spread

At $f_{90}$, the phase spread between purely ID phenomena (ie those having no penetration into the wall) and similarly defined purely OD effects is 90°. If, as is often done, the plane is rotated so that ID effects such as probe wobble and Pilger noise (due to tube bore rifling, see section 5.6.1) lie along the positive x-axis, and the OD wall thinning locus lies along the negative going y-axis (see fig. 2.3), then the remaining 2 directions, -ve going x and +ve going y are inward denting and wall thickening respectively. Denting increases probe to wall coupling because of the force applied by the spring carrying the probe, and is essentially an ID effect. Wall thickening is the opposite: an OD effect.

An ID defect locus lies not exactly along the +ve x-axis (increasing lift off), but slightly into the fourth quadrant towards wall thinning. This is best explained by an example. At the bottom of an ID crack, the defect might be thought of as either a localised thinning of the wall or as lift off of the probe from the metal surface within the crack. In fact both effects take place and appear in the resultant phase of the locus, but lift-off predominates because the more intense surface currents affect the resultant more severely than the weaker currents flowing at the outer surface. An OD notch causes no such movement along the x-axis, because only the wall thickness is affected. Any ID defect with a significant depth will have a locus rotated towards the wall thinning axis, so consequently at the $f_{90}$ frequency, the separation of ID and OD defect loci can be expected to be less than 90° in practice.

Probe wobble can be difficult to distinguish from shallow ID defects, since both phenomena are largely a matter of lift-off. Frequently they have similar magnitudes, as well as identical phase angles. Under these circumstances, only multifrequency inspection can tell them apart.

### 2.6 Differential vs. Absolute Probes

There are two basic types of probe, from an electrical point of view: absolute and differential ones (fig. 2.4). An "absolute probe" in fact consists of two identical probes heads (each with a single coil winding) connected to opposite sides of the bridge. The heads are physically separate. Each head is positioned inside a tube, one on the region of wall to be inspected and the other on clean metal for reference. The reason for keeping the both coils in contact with separate samples of the same type of metal is that it aids balancing of the bridge when a range of test frequencies are in use. For both types, a Wheatstone bridge is used to measure changes in probe impedance, so the "absolute" type is actually a misnomer.

In the differential probe, two coils are wound close together, both in contact with the specimen, and connected to form the lower two sides of the bridge i.e. the part which causes the difference voltage to arise. The essential difference between the two types of probe is thus that in a differential probe both sensing coils are coupled to the area of material currently under test, whereas with an absolute probe one coil is always coupled to clean metal. The differential coil is so-called because it effectively differentiates the magnetic field. It has a driven primary, wound on top of the adjacent test coils which form the secondary. The secondary coils are wound in opposite

directions, one on each leg of a ferrite rod which has been cut to resemble a tuning fork. The coupling of both coils is normally equal. Differentiation of the field in the direction of separation (and motion) of the two test coils occurs.

In both electrical inspection configurations of the coils in the bridge shown in figure 2.4, the excitation sinusoid is applied across the opposite corners of the bridge to which the measurement circuitry is connected. This is the way in which the bridge is connected in the Zetec equipment used in Chapter 5. The Forster circograph, and indeed all other more recent Forster equipment such as the Defectomat series use a an electrically separate primary (wound on top of the secondary coil or coils which from part of the bridge) (**Davis**,1980) in order to produce the eddy currents. $V_{out}$ in figure 2.4 is normally connected to a set of four quadrant mixers which multiply the output sinusoid by the input one, and then by lower intermediate frequencies in order to produce the x and y signals which drive the impedance plane display.

The differentially wound secondary/separately wound primary configuration was favoured in the first EC instruments developed by Forster in the 1950's, because the low impedance primary allowed more power to coupled into the coil, whilst a high impedance secondary could be made very sensitive with a lot of turns, and benefit from the advantages of matching a high impedance input stage.

The absolute probe produces a characteristic single lobed figure (somewhat like the perimeter of a machine printed comma: see for example fig 5.12) as it passes over a defect. The differential probe produces a figure-of-eight (see fig 2.6), with lobes lying at the same angle as for the absolute probe, but as if the point describing the loci had been rotated about the origin half-way through

describing the figure. These characteristic signals are independant of coil geometry, whether the coil be surface or bobbin type. Due to the easier-to-interpret defect loci from an absolute probe and its non-directionality, and the absence of any problems in detecting very gradual defects, the absolute probe is preferred to the differential in all but a few circumstances. Its use is not however entirely without problems: Pilger noise (Chapter 5) may occur, and in prolonged use, the probe impedance may drift due to thermal effects because of tip friction.

For the tubes in chapter 5, two probes, a test and a reference, were used. For chapter 4, a differential probe was used.

## 2.7 Synthesis of a Differential Signal

It was found during the course of some low pass digital filtering experiments that a differential probe locus could be synthesised from one logged (section 5.6.3) using an absolute probe, thus removing any advantage gained by using a differential probe whilst computerised data logging was taking place. Furthermore the differentiation is not restricted to the direction of probe travel, but can also be along the tube axis rather than along a circumference. In this way a scan with a differential bobbin probe can be simulated even when scanning was done originally with an absolute pancake probe. A further advantage is that in differentiating in the axial direction the signal differentiated pertains to a much narrower angular sector of tube (dictated by the effective field diameter, $D_{eff}$, eqn. 2.5) than for a real bobbin probe: in this way one of the limitations of bobbin probes is overcome, namely their inability to convey any angular positional

information. The most important factor in getting this synthesis to work correctly was separation of the sample points to be subracted one from the other. They had to be a distance apart in time equivalent to the pole separation distance in the real differential probe.

### 2.8 Rotating and Bobbin Probes

For ID tubing inspection, there are two main types of probe: the (non-rotating) bobbin coil (so called because it resembles a bobbin of thread), and the rotating pancake coil. The former has been established in use for tubing inspection since the late sixties, but is physically incapable of inspecting the expanded diameter tubing in the tubesheet region, or of being forced past a dent. In a rotating probe, the coil is mounted upon a leafspring (see fig 2.5), and the main body of the probe is smaller than the tube bore by a few millimetres. The probe should glide along the bore, centred by a ring at each end, covered with stiff nylon bristles, so that the probe tip is capable of taking up small variations in the tube diameter. A number of such probes were designed and built by the author, similar to (but mostly smaller than) that shown in approximately 1:1 scale in fig 2.5.

Since about 1978, **Brown** and various other authors (e.g. **Van Drunen, Sharp and Cecco**,1983) have asserted that the detection of low volume defects (ie cracks) in dented areas can only be accomplished with high resolution probes such as the one in fig. 2.5, whether it be wound as an absolute or a differential probe. The aim of this research is to improve defect detection and classification rates for smaller more localised defects than have been previously of concern, due to the increasing emphasis being placed upon the integrity of

primary circuit components, and so exclusive use has been made of small diameter (typically 2.5 mm.) pancake probes. Various pieces of practical work have been performed by the author and his sponsors to support this. Before this study started, Babcock Power performed some research to determine the most sensitive ferrite core cross sections for surface probes (**Farley, Hamilton and Little**,1980). The author accumulated considerable practical experience during the process of making a number of probes in attempts to detect the original '316' defect set, in machining small ferrite cores and winding coils with very fine enamalled copper. The work that was done is discussed further in section 2.12.

The advantages and disadvantages of the four possible combinations of probes by electrical and geometric categories are best summarised in tabular form. Table 2.1 summarises the differences between surface and bobbin probes and Table 2.2 compares absolute and differential probes.

### 2.8.1 Speed of Inspection

An axial inspection speed of 0.25m/s for a bobbin probe is widely quoted as possible. The limiting speed is the maximum at which a crack (the defect requiring the greatest bandwidth) can be traversed without there being any degradation of the details of the shape of its locus with respect to that obtained at a minimal drawing speed. Distortion can arise due to bandwidth limitations of the recording method, whether that be tape recording or digitisation.

**Russell and Richardson**(1984) reported drops in signal peak heights of 50% for increases of bobbin probe speed from 0 to 0.6 m/s, but showed that this was

due to the slewrate limit imposed by their chart recorder despite its being of the best quality. They noted a drop of only 5% over the same range of speeds when digital aquisition was used. **Selig and Neumaier**(1983) have demonstrated that the effect of not enough bandwidth in any stage between the probe and the display screen is slewrate limiting which compresses the phase vs. depth relationship in a manner such that an impedance plane phase spread of 30° can completely dissappear, making assessment of the throughwall depth highly inaccurate (see fig. 2.6). When the differential probe y-signal is plotted against time (assuming the ID and OD axes are aligned as for section 2.5) the effect of insufficient bandwidth can be seen to be a widening of the gap between the peaks of the characteristic swings in opposite directions associated with a defect. The deepest defects contain the biggest swings in the y direction. These are attenuated by the bandwidth limitation, and the phase of a severe defect appears the same as a moderate one, although the phase spread for ID defects to moderate OD remains unaffected because the bandwidth requirements for moderate to shallow defects are not so great.

For undegraded storage of a crack's impedance plane locus with respect to paper recordings made of the trace at logging time, a minimum spatial sampling rate of 1 sample/mm of tube was found necessary. For multifrequency eddy current equipment, as used here, the sampling rate associated with the excitation multiplexer is usually the limiting factor. The MIZ-12 multiplexing rate was such that the individual channel sampling rate was 400Hz, effectively imposing a maximum speed of just over 8 revolutions/second around the 15.5mm internal diameter tubing for a rotating probe at a sampling rate of 1 point/mm. The actual rate was 6.5 revolutions per second, so the signal was unaffected by sampling. The pitch of the helix

described by the probe was 1mm./revolution, so the axial rate of inspection was 6.5 mm./sec. For the same resolution using a bobbin probe, 0.4m/s would have been possible. Bobbin probes are still used for most work because of this 50-fold advantage in axial rate of progress and the consequent cost of inspection time compared to rotating probes. Refering to the tubeplots in Chapter 5, the first set of loggings (those done with the University computers - see appendix) had a much higher sampling rate than was subsequently decided necessary. In the later set of tubeplots (those done on a lower quality plotter as will be seen by comparing fig. 5.43 with fig. 5.23), the sampling rate was reduced and the horizontal and vertical scales have been adjusted so that the scales on the plots are equal.

A visual check was performed on the recorded loci for any directional effects due to velocity. The through wall holes for angular registration at either end of the IGA tube samples were known to be circular and so their image in the "unrolled" tubeplots was expected to be symmetrical about the centre of the hole. This was indeed so (see fig. 5.25), and from this it was concluded that there was no significant field distortion due to the speed of travel of the probe across the metal. Within the literature, there is almost no mention of the effect of probe velocity on EC testing of small diameter tubing probably due to the relatively low speeds and high inspection frequencies involved. No workers in other groups inspecting steam generator tubing have reported any effects due to deformation of the probe field at speeds of up to $0.5m^{-1}$. Morgan(1985) experienced a difference between the actual and sensed position of defects when inspecting 25mm. wall thickness, large bore gas mains tubing at $4ms^{-1}$ and 1kHz, due, he suspected to the magnetic field from the 100mm. diameter probe lagging behind because of the low frequency and high

speed. He confirmed this by calculating the field shape for the probe using a finite element analysis package and the effect was very apparent in the field plots.

## 2.9 Theoretical Studies

If complete simulation of the interaction of a field with a defect were possible, then it would not be necessary to approach this project using pattern recognition and methods based on a collection of prototype samples. Instead, a complete set of prototype signals could have been manufactured. Furthermore, if inversion of the signal back to a functional description of the defect itself were possible, the classification problem could have been attacked from a non-statistical point of view as well. What is currently possible and what has been done in the past in simulation, and why efforts to do more have been frustrated so far, will now be considered.

The history of simulation of the interaction of magnetic fields with conductors to produce eddy currents reflects the analytically insoluble nature of the governing equations for all but the simplest of geometries of coil and conductor. Early work centred on deriving an equation in terms of the vector magnetic potential (normally denoted A) from Maxwell's equations. This quantity is a mathematical convenience of tenuous physical significance, popular due to the ease with which probe current and impedance can be calculated from it. Only in the cases of a probe coil (single turn or rectangular cross section) standing end-on, above a homogeneous infinitely thick conducting slab, and an encircling coil around a long bar, have the necessary conditions of symmetry been present about the probe axis to allow an

analytical solution for **A**. Even then, the solutions are so complex that their evaluation by computer is almost essential to enable plotting of predicted eddy current magnitudes and phases.

**Bareham**(1960) was the first to use computational methods upon Maxwell's equations in the context of eddy current NDT. Starting with the much quoted derivation of the eddy current diffusion equation in terms of the vector magnetic potential, he solved it for for the particular case of an infinitely long encircling coil (the simplest inspection geometry to analyse), containing thin-walled tubing, in order to quantify the eddy current intensity and phase lag profiles with depth. He gives a formula for the required excitation frequency, to achieve one of four profiles. Even for this simple configuration, numerical evaluation of the closed form solutions to the vector magnetic potential differential equations was necessary.

The computation involved in solving the same basic equations, without the simplifying assumption of axial symmetry, only became feasible recently and were first applied to eddy currents when **Brudar** (1982) simulated a bobbin coil encircling tubes with a uniform ID slot running along the length of the tube. He did this using a 2-dimensional finite element simulation (see below).

Brudar's plots indicate the way in which the eddy currents avoid a defect by going underneath it but the simulation still assumes an infinitely long tube with an infinitely long crack, so it is still not possible to see to what extent the currents are deflected around the ends of a defect.

**Dodd and Deeds**(1968) published a very widely referenced paper, in which the same basic vector magnetic potential equations are solved for both

common inspection configurations: a solenoid above a conducting halfplane of infinite depth clad with a layer of a different conductor, and a rod similarly clad, surrounded by an encircling solenoid. The paper then went on to define coil impedance upon the solutions of these equations, and to demonstrate the correctness of the treatment by generating the classic impedance plane loci for liftoff, and for variations in excitation frequency, conductivity or permeability.

**Zaman**(1982) imitated the same method, and produced current intensity contours which explain the dip in the response as a surface probe passes over the centre of a hole: the current intensity peaks directly under windings of the coil(see fig. 2.7), at its mean radius, and decays exponentially outside this circle. It is zero everywhere within the hole.

**Betzold and Becker**(1976) calculated the current density under a pancake probe under various circumstances, as did **Dodd,Deeds,Luquire and Spoeri**(1969). By calculating the contours of a complex quantity with both magnitude and phase, the "defect sensitivity factor", they clearly showed the extent to which eddy currents from a pancake coil have only limited penetration.

As already mentioned in connection with Brudar's work, finite element (FE) simulation has recently become computationally practical. It works by iterative minimisation of a function describing the degree of discrepancy between the points in a mesh, given the governing field equations (Maxwell's equations expressed in a discrete form).

**Lord** has published extensively on the application of a 2-d axisymmetric

finite element simulation (described in detail only in **Lord**(1981)) for instance to design a probe to maximally separate the defect and tube roll (see fig 1.4) signals whilst inspecting the tube expansion region (**Ida** et al, 1983). He reported the time taken for a calculation of the signal from a bobbin probe passing axisymmterical pitting at the top of a tube support plate as 66 minutes of cpu time on a Cyber 205 supercomputer (**Brown**, 1984). A recent review of 2 and 3 dimensional FE simulation packages concluded that 3-D simulation of the time varying field due to the passage of a probe over a realistically thin crack is still prohibitively costly (**Becker**,1986) in computer time and requires enormous amounts of memory to store (and in particular invert) the model's nodal matrix which contains a real number for every node in the problem's mesh. **Lord** has suggested that processing this matrix in the case of defects requiring modelling at the granular level such as IGA and IGSCC will require the storage of $10^9$ real numbers, which is equivalent to the storage capacity of twenty 200 Megabyte Winchester disk drives, and is thus currently still a long way away from being feasible.

**Brown**(1983) has used a similar finite element model to investigate the shape of loci from a surface probe as it passes from good thin plate to a 2 layered plate, where one layer is given the resistive properties of intergranular attack (IGA) or thinning. Discrimination between IGA and thinning was predicted to be possible, using the magnitude and phase of the loci extrema, but became more difficult as the resistivity of the IGA approached an infinite value. This difficulty would lead to a practical problem, because infinitely resistive IGA has been observed: it is apparently due to loss of contact between the grains. Upon total grain boundary dissolution and consequent wall loss from erosion, there ceases to be any bulk difference

between IGA & thinning that is measurable with eddy currents, and their characteristic loci become identical.

The concern caused by the occurence of IGA in SG tubing is due to intergranular stress corrosion cracking (IGSCC) which tends to accompany it. A tube rupture can be initiated from IGA when of a lesser depth than would be necessary with thinning because the IGSCC provides crack initiation sites unlike smooth surfaced thinning which is caused by mechanical abrasion. Thus a more stringent plugging criteria has to be applied for IGA. American power generation companies have reported difficulty in telling IGA and thinning apart, and in determining the depth of penetration of IGA, because both depth of penetration and wall conductivity affect the locus size and deep IGA with little grain boundary dissolution has been observed (by tube sectioning) and produces a very similar locus to that of shallow IGA with heavily attacked grain boundaries. Some authors (e.g. **Brown**,1983) have suggested that it might be possible to tell IGA and thinning apart using multifrequency equipment, but in order to analyse several loci outside of a laboratory at speeds economic for on-site usage, automatic signal analysis will be necessary. This work was in part prompted by such remarks.

## 2.10 Transfer Function of the Probe

In eddy current NDT, the testing equipment can thought of as a system with an input and an output: the input is ultimately the crack in the metal wall even though sensed with a coupling field and currents and the output from the probe (or more specifically the probe signal demodulation circuitry: see fig. 2.4) is a deterministic output due to a defect, partially submerged in

background noise and probe wobble. In actuality, little is known about the input to the probe. Although an axial crack is a linear discontinuity in conductivity, there is little justification other than of an intuitive nature, in saying that the input to the probe is therefore an impulse. However, it is reasonable to postulate that as the field is passed over the defect, the discontinuity sequentially affects the parts of the field. The final ouput is analogous with the cross correlation of an increasingly lagging pulse with a function which defines a field distribution similar to that shown in fig.2.7. It seem would seem reasonable to apply parameter identification approach to the signal from the probe to identify the "defect-to-probe- output" transfer function, and then apply a matched filter to invert the probe transfer function, in order to reconstruct the original defect cross section.

Unfortunately, the determination of the transfer function of a system is achievable only if both the input and output are known.

Various methods of parameter identification including Kalman filtering were studied in an attempt to find a way around this; but ultimately there is no way to avoid the maxim that "either the transfer function of a system must be known, in order to estimate the input from a noisy output, or the input and output must both be known, in order to estimate an unknown transfer function". Any methods found to overcome this restriction ultimately involved enforcing some model upon the probe transfer function. Analytical methods of predicting the impedance of coil in the presence of a defect of a particular shape and position with respect to the probe (and the reverse operation popularly referred to as inversion in the literature) are still in their infancy, and require the imposition of gross assumptions on the form of the defect e.g.

that it consists of one vertical fissure of negligible width. Numerical modelling methods are also unsuitable for the production of a transfer function because they do not yield a functional relationship between the defect (however its shape might be functionally functionally) and the signal shape (similarly, however that might be defined).

**2.11 Mixing**

Suppose that an OD defect and an unwanted signal such as surface noise separate by 90° at some frequency. At another higher frequency, the phase separation will be substantially different. By inspecting an area of tube where only the undesirable locus occurs, and adjusting the size and degree of rotation (phase) of the higher frequency channel locus a subtractive mix insensitive to the undesirable effect can be found. The penalty paid for this is that the phase spread between phenomena of different depths is reduced, making assessment of depth more error prone; and that an extra channel of different frequency is required.

All multifrequency instrumentation has provision for mixing of channels, usually within special purpose mixer modules which have front panel phase adjustment controls in addition to the individual channel gain potentiometers normally found in such units. In essence what is done is that the source of an unwanted signal is scanned at two well separated frequencies. One locus is then rotated and scaled, and the phase and horizontal and vertical gains are adjusted so that a third locus (which is the difference between the first two) is optimised i.e. minimised in size. If this can be fully achieved, the mixed locus collapses to a stationary spot, though only for

variations in the tube wall parameter not of interest. In commercial multifrequency equipment there are currently two popular ways of presenting the operator with enough visual clues to enable him to "mix out" unwanted signals. The first of these is simply to display to him the mixed locus itself, and the second way is to present a locus derived from the two signals being matched prior to subtraction, to focus the adjustment of controls upon achieving a match.

The block diagram of fig. 2.8 shows a mixing system representative of the first type. This is exactly how the Hocking D5 and D6 instruments work (**Hocking**,1985) : by rotation of the locus of channel 2, equivalent to the standard rotational transformation of a point (x,y) by $\phi$ in a clockwise direction to form a new point (x',y'):

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \qquad (2.7)$$

Unfortunately, it is not particularly easy to obtain a good annulment with this method, although it is the simplest to visualise and describe. This is because there are three controls to be adjusted: phase, horizontal gain and vertical gain.

The mixing method employed in the Zetec MIZ series (see fig 2.9) tends to be preferred by NDT technicians because of the greater ease of setting up a mix. The phase rotation scheme is slightly different. The horizontal and vertical signals of the subtracted loci can be rotated by different angles, since a phase rotator is employed for each orthogonal component of one channel in a two channel mix. The equations describing the operation performed by the phase rotator are thought to be

$$\begin{bmatrix} x'_2 \\ y'_2 \end{bmatrix} = \begin{bmatrix} \cos\phi h & \sin\phi h \\ \sin\phi v & \cos\phi v \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \qquad (2.8)$$

since this was their form in the prototype developed by **Davis**(1980) although the manufacturer would not confirm this. $\phi_H$ and $\phi_V$ are angles specified on thumb wheel switches on the front panel.

The benefit bestowed by use of this form is that the mirror image of the figure about its major axis can be generated and the direction of description reversed. Variation of the angle specified to the $\phi_H$ phase rotator through $360^\circ$ has been simulated for a figure-of-eight (fig. 2.10), asymmetric about the major axis in order to highlight the flipping of the locus. It tilts from a gradient of +1 to -1, and back again, but during the second $180^\circ$, the figure is reflected in its major axis. The variation of $\phi$ for the more normal Hocking type equations is shown in fig. 2.11.

It seems unlikely that this ability to reverse the sense of description of the loci is really necessary since the Hocking equipment cannot do this and yet a mix can always be found, even if it is more difficult to do the adjustments to achieve this. It is more likely that the form of the rotation equations has been picked in order to facilitate the MIZ-12's rather clever operators aid for anullment, which is a display of either the two horizontal or two vertical components from the channels to be mixed, plotted against each other on cartesian axes. Elimination is done by adjusting the horizontal gain and the horizontal phase rotator whilst watching the horizontal axes figure. When they match perfectly, a straight line with gradient of +1 is displayed. The process is then repeated with the two vertical signals. On completion the undesirable locus in the derived channel will have been mixed out.

### 2.11.1 Choice of Frequency

The choice of frequency tends to be a compromise when only one frequency is allowed because the standard depth penetration/phase spread theory states that the standard depth of penetration, $d \propto 1/f^2$ and the phase lag, $B \propto f^2$

For a pancake probe, these relationships do not hold quite so rigorously, and the choice of frequency is not so paramount. The probe diameter sets the maximum penetration, and within that limit the frequency causes variation in the rate of change of phase with penetration depth. The behaviour of both current intensity and phase against depth as has already been explained, is not functionally defined since all plots of this have been produced computationally (see fig. 2.1). No studies of the variations in these distributions for frequency have yet been published. Thus there exist no quantitative rules for picking the frequency of inspection for a pancake probe, other than eqn. 2.6 for the $f_{90}$ frequency, for 90° phase spread between ID and OD defects. In other work, when 3 frequencies have been used, they have been picked thus:

(i) One, as high as possible (up to 1.6MHz), to sense only surface phenomena and liftoff.

(ii) Another to sense OD phenomena such as tube supports, the tube sheet, external magnetite and copper deposits, or corrosive ferromagnetic sludge in the tube roll/tube sheet crevice. This one is required to be as low as possible.

(iii) The third, normally chosen to be the $f_{90}$ frequency, for good phase spread between ID and OD.

A mix of such frequencies allows the elimination of the signal due to probe wobble and one OD phenomena, such as the tubesheet, or copper deposits.

The fourth channel in the MIZ-12 exists chiefly because at the time of its development, the benefit of multifrequency systems was perceived to be that they allowed "parallel acquisition": simultaneous scanning of tubing with frequencies optimised for the detection of several phenomena simultaneously, in order to reduce the time spent scanning.

For Inconel steam generator tubing, with 1mm wall thickness, the phenomena to be detected are typically

(i)tube sheet cracking,

(ii)the presence of magnetite in the tube sheet crevice and sludge profiling,

(iii)ID and OD defects and denting

and

(iv)denting and ID effects alone.

It was with these aims in mind that the frequencies for this study were chosen: 100,220,300 and 600kHz.

Brown's (1982) work was done using the Zetec MIZ12 prototype, which allowed him to use 1.6MHz in one channel to detect only Pilger noise. Unfortunately, the production version was limited to frequencies in the range 10kHz to 1MHz, so duplication of his mix for comparison of results was not possible. This was a pity since he was using the same type of tubing as has been studied here. Also, he complained of highly distorted loci bearing no resemblance to their unmixed equivalents (as have some other authors) when using 2 frequency Pilger noise elimination, and of the

undesirability of a frequency as high as 1.6MHz. **Brown**(1982) also demonstrated that it was possible to remove Pilger noise from a 200kHz locus using two lower frequency channels of 400 and 800kHz rather than the 1.6MHz channel, due to the additional redundancy of information in the 3rd channel. Surprisingly, the present author found it perfectly possible to remove Pilger noise from 220kHz with a 600kHz locus, as is explained in section 5.5. Unfortunately, **Brown**'s mixing technique is not explained quantitatively enough for it to be computationally performed.

## 2.12 Choice of Probe Type and Material

It had originally been intended to use **Macleod**'s defect set in 316 stainless steel tubing for the multifrequency work, expanding it as necessary by the manufacture of more defects in the same type of tubing. The defect signal loci from both the old and the new inspection systems could then have been compared, using the same frequency in one of the four channels and the same type of probe. This would have been desirable, to allow pattern recognition experiments to commence with the same feature set, as had been used by **Macleod**, and thus provide a means of validating the results.

Several factors prevented this, and eventually combined to force a change of tubing type, and of probe type.

The Circograph was restricted to inspection at 400kHz, because of the tuned rotating core transformer, which it uses to couple (the demodulation electronics) to the probe coils as well as to the mechanical shaft drive.

The sponsors became more interested in the determination of the limits of

defect detectability in Inconel 600 rather than in 316 stainless steel (and particularly in the IGA occuring in it) and so fabricated a set of defects in that material. The reason for the change in emphasis was that it had been established that nickel based alloys such as Inconel are more resistant to the types of corrosion likely to occur in a PWR than stainless steel alloys.

The Zetec equipment was found to be incapable of detecting defects in 316 tubing, regardless of probe type. The reason was never successfully determined, despite repeated communication with the manufacturer's British agents and many attempts with various probe core materials, geometries, winding configurations and types and gauges of wire.

In a final attempt to detect a defect in the original tubing set prior to the switch to Inconel 600, two pairs of rotating absolute probes, identical in every detail other than in diameter, were tested with the Zetec equipment, one pair in 316 stainless steel test and reference tubes, and the other pair in Inconel 600 tubes. The notches in both of the tubes under test were sufficiently severe (0.5 mm. deep) to be visible to the naked eye. The notch in the Inconel was detected, but a similar one in the 316 tubing was not, regardless of the frequency used over the range 10kHz to 1MHz. Particular note was taken of the non appearance of the defect at 400kHz, the same frequency at which the notch in the 316 tubing was apparent using the Circograph. Subsequent double checking, by inspecting the slightly larger bore Inconel tubing by hand with the probes designed for the smaller 316 tubing showed that they could detect the notch in the Inconel with ease. The materials were both nonferromagnetic, with similar conductivities, and the notches were produced by the same means i.e. electrode discharge machining.

No certain explanation for this anomaly was found, but it seems likely that it was due to a lack of sensitivity in the MIZ- 12. If this is true, the design fault may well have been rectified by the manufacturer in later models, since the system bought by Babcock Power Ltd. was one of the first.

## 2.13 Concluding Remarks

The complexities of eddy current testing have been reviewed in order to demonstrate the advantages of applying a computer to various tasks within the subject. In the next chapter, the tasks which have to be programmed in order to automate defect detection and classification are discussed in detail.

## 2.14 References

1. **Hagemaier,D.J.**,"Eddy Current Impedance Plane Analysis",M aterials Evaluation Vol.41, pp211-218,February, 1983.

2. **Cecco, V.S., Van Drunen, G., and Sharp F.L.**, "Eddy Current Manual", Vol. 1, Rev.1, Sept., 1983, AECL-7523,Chalk River Nuclear Laboratories

3. **Davis, T.J.**,"Multifrequency Eddy-Current System for Inspection of Steam Generator Tubing",EPRI report EP NP-1621,Nov. 1980.

4. **Brown, S.D.**,"An Evaluation of Eddy Current Inspection Methods for PWR Steam Generator Tubing", EPRI-NP-636,Oct. 1978.

5. **Farley, J.M., Hamilton, J.W. and Little, I.M.**, "Practical Experience of Eddy Current Tube Bore Inspection with a Rotating Probe System", Babcock

Research Centre internal report 832, 1980.

6. **Russell, M. and Richardson, K.P.**,"Automatic Inpection and Evaluation of Power Plant Condenser Tubing",25-1 -25-14, CEGB internal Report.

7.**Selig and Neumaier**,"Testing of Steam Generator Tubing by the Eddy Current Method - Improvement of Defect Sizing", Proc. 3rd ASME Conf. on Qual. Ass. in the Nuclear Industry, Dec.,1983.

8. **Morgan, L.L.**, "Pipeline Inspection On-line", pp248-261, 20th Ann. Brit. Conf. on NDT, Erskine, 1985.

9. **Brudar, B.**,"The Electromagnetic Field in the Neighbourhood of a Defect in a Material", from Research Techniques in NDT, Vol.5, pp169-229, Academic Press, 1982.

10. **Dodd,C.V. and Deeds,W.E.**,"Analytical Solutions to Eddy-Current Probe-Coil Problems", J.App. Phys.,39,(6),2829-2838. May,1968.

11. **Zaman,A.J.M.,Gardner,C.G.and Long, S.A.**,"Change in Impedance of a Single-Turn Coil Due to a Flaw in a Conducting Half Space",J. Nondestructive Eval., Vol. 3, No.1, 1982.

12. **Betzold, K. and Becker, R.**,"Numerical and Experimental Examinations for Eddy Current Testing", Materialprufung, 1976, 18,(9),345-348.

13. **Dodd, C.V, Deeds,W.E.,Luquire, J.W. and** Spoeri,W.G.,"Analysis of Eddy Current Problems wih a Time-Sharing Computer",Materials Evaluation, pp 165-168, July,1969.

14. **Lord, W.**,"Development of a Finite Element Model for Eddy- Current NDT Phenomena", EPRI-NP-2026, Sept.,1981.

15. **Ida,N.,Palanisamy,R. and Lord,W.**,"Eddy Current Probe Design Using Finite Element Analysis",Materials Evaluation/41, pp1389- 1394, November,1983.

16. **Becker, R.** et al, "The Modelling of Electrical Current NDT Methods: Its Application to Weld Testing (parts I and II)", Brit. J. NDT., pp.286-294, Sept. 1986 and pp 314-320, Nov. 1986.

17. **Brown, S.**, EPRI 1984 Annual Research Report, Section 19, EPRI NDE Research Centre, 1984.

18. **Brown, S.D.**,"Eddy current NDE for intergranular Attack",EPRI- NP-2862, Feb,1983.

19. **Brown, S.**, "Field Experiences With Multifrequency- Multiparameter Eddy Current Technology", EPRI-NP-2299, March 1982.

20. **Brown, S.D.**,"Multifrequency/Multiparameter Eddy Current Steam Generator NDE",pp99-105, 2nd ASME Conf. on QA in the Nuclear Industry, May 1982.

21. **Macleod, J.E.S.**, "Feasability Study on Computer Sorting of Flaws in Steam Generator Tubes Scanned by a Rotating Eddy Current Scanner", Babcock Power Research Report No. 07/82/46, July 1982.

22. **Hocking Associates** Private communication concerning D6 Specification, Hocking Associates, St. Albans, 1985.

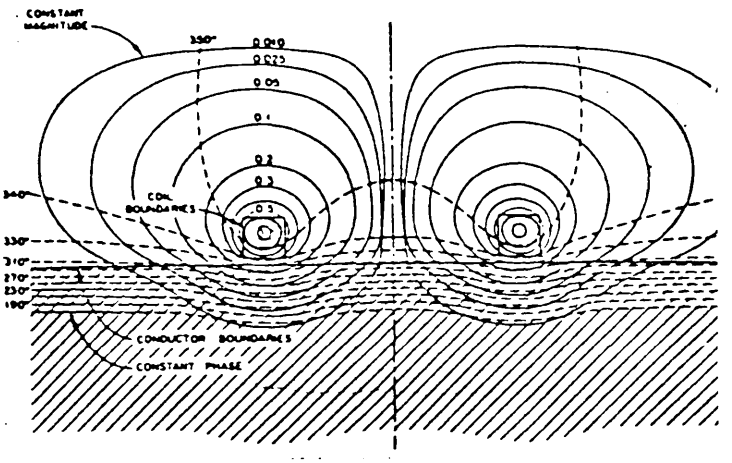



**fig 2.1** Circular magnetic flux contours and lines of equal phase shift with respect to the coil excitation current, from a numerical simulation by **Dodd** (1978), demonstrating the non-planarity of the field incident upon the air/metal boundary. The intensity of the eddy currents flowing anywhere in the metal are directly proportional to the intensity of the sinusoidally varying magnetic flux at that point.

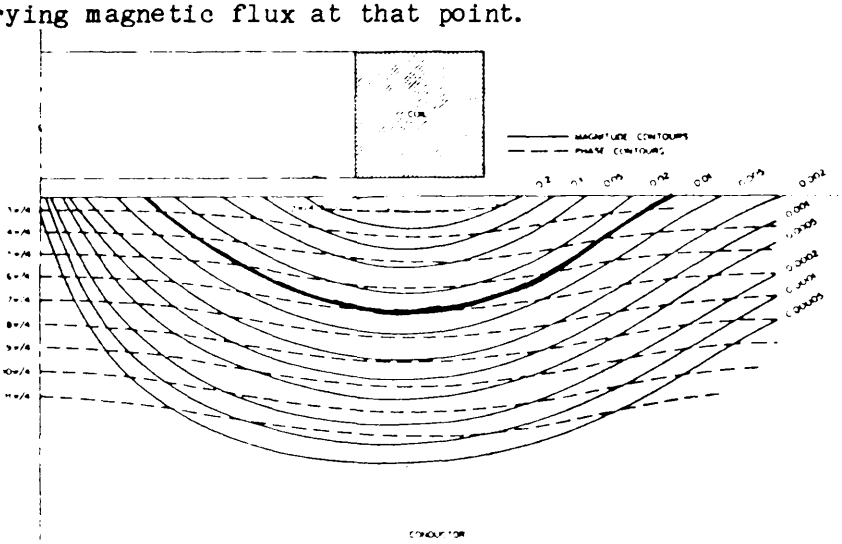


**fig. 2.2** Contours of defect sensitivity ie the normalised magnitude and phase change of the coil impedance for a small defect (from ref. 11)

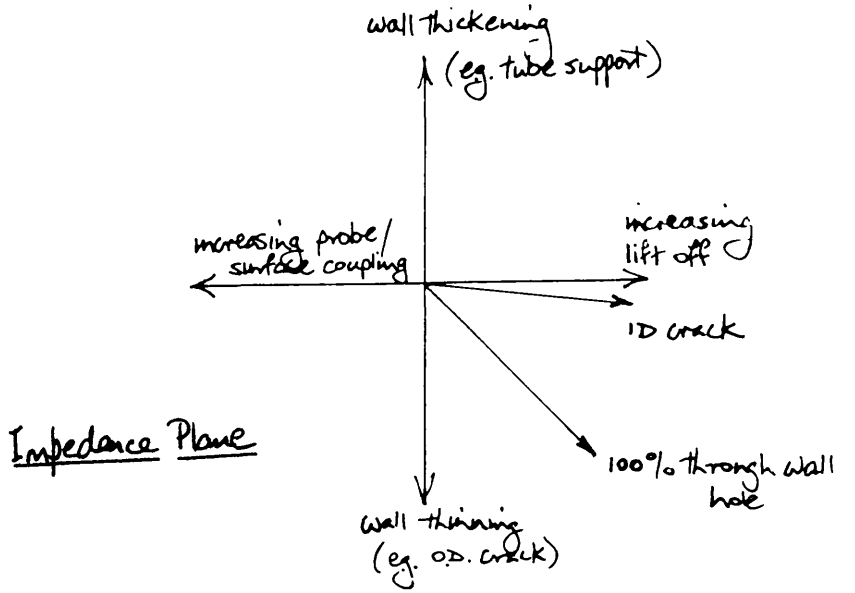



Fig. 2.3 The meanings associated with movements of the coil impedance away from the null point in the impedance plane, with excitation of the probe at the $f_{90}$ frequency.

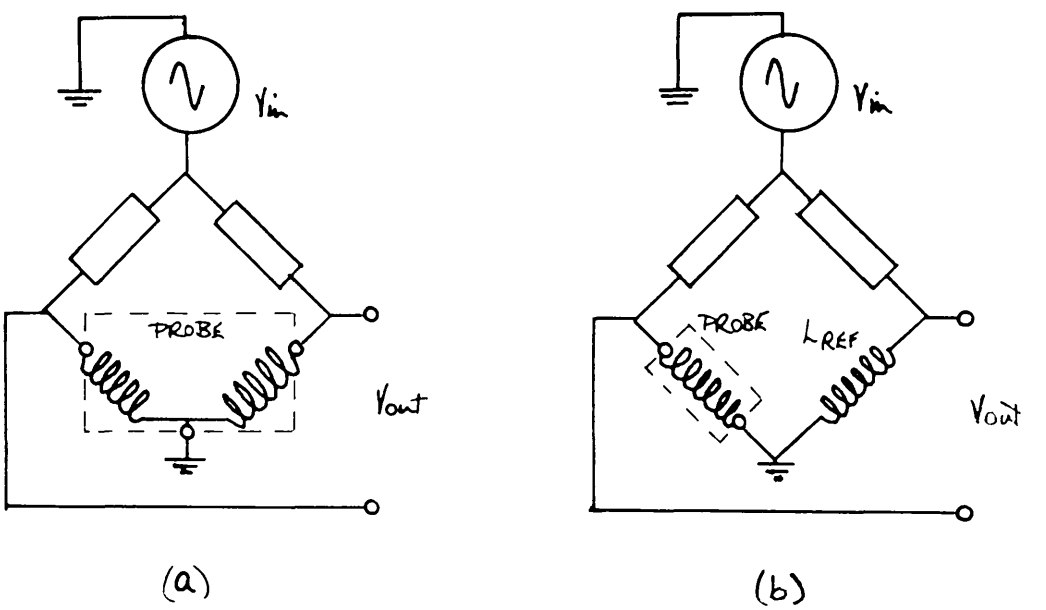



Fig. 2.4 The difference between (a) a differential and (b) an absolute probe when represented schematically in terms of the way in which they are connected into the bridge found in most instruments. Subsequently the output is mixed with the input sinusoid, and then passed through a low pass filter. The baseband signals which remain provide the beam deflections which produce the characteristic impedance plane loci.

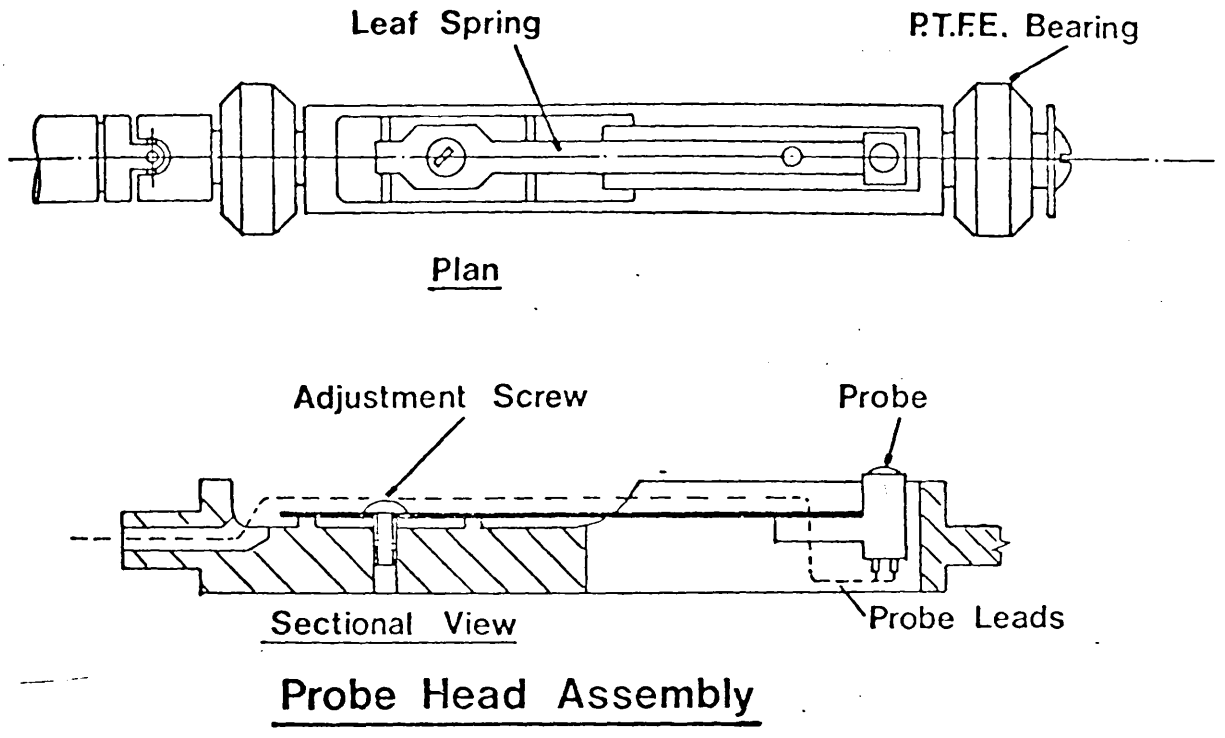



fig. 2.5 The Circograph rotating probe. The rotating probe used upon the Inconel 600 tubing in Chapter 5 was based upon this design.

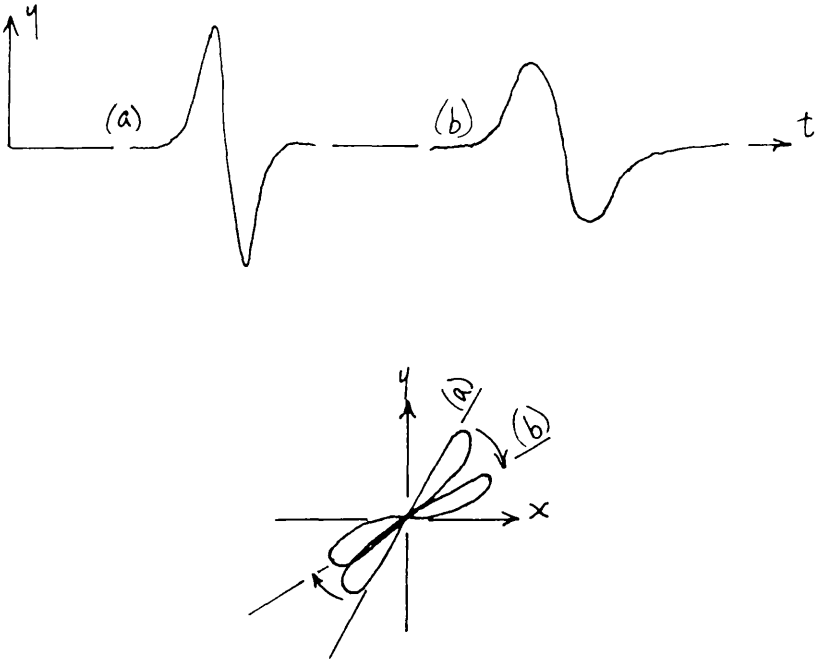



**fig 2.6** Bandwidth reduction in the y-channel and consequent slew rate limiting can cause reduction in the peak height: for example (a) is degraded to (b). Assuming that the x signal is already fairly rounded it is unaffected, so the bandwidth reduction is manifest as a rotation of the loci toward the x-axis, because large transient values of y are supressed.

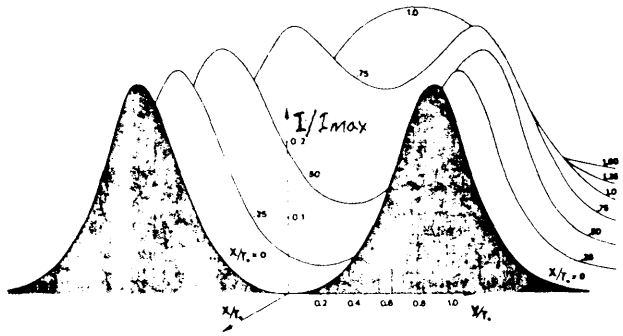



Fig. 2.7 The current induced in a surface probe as a raster scan is performed past a hole. X and Y refer to the displacement of the probe centre with respect to the hole centre, and $r_o$ is the hole diameter.

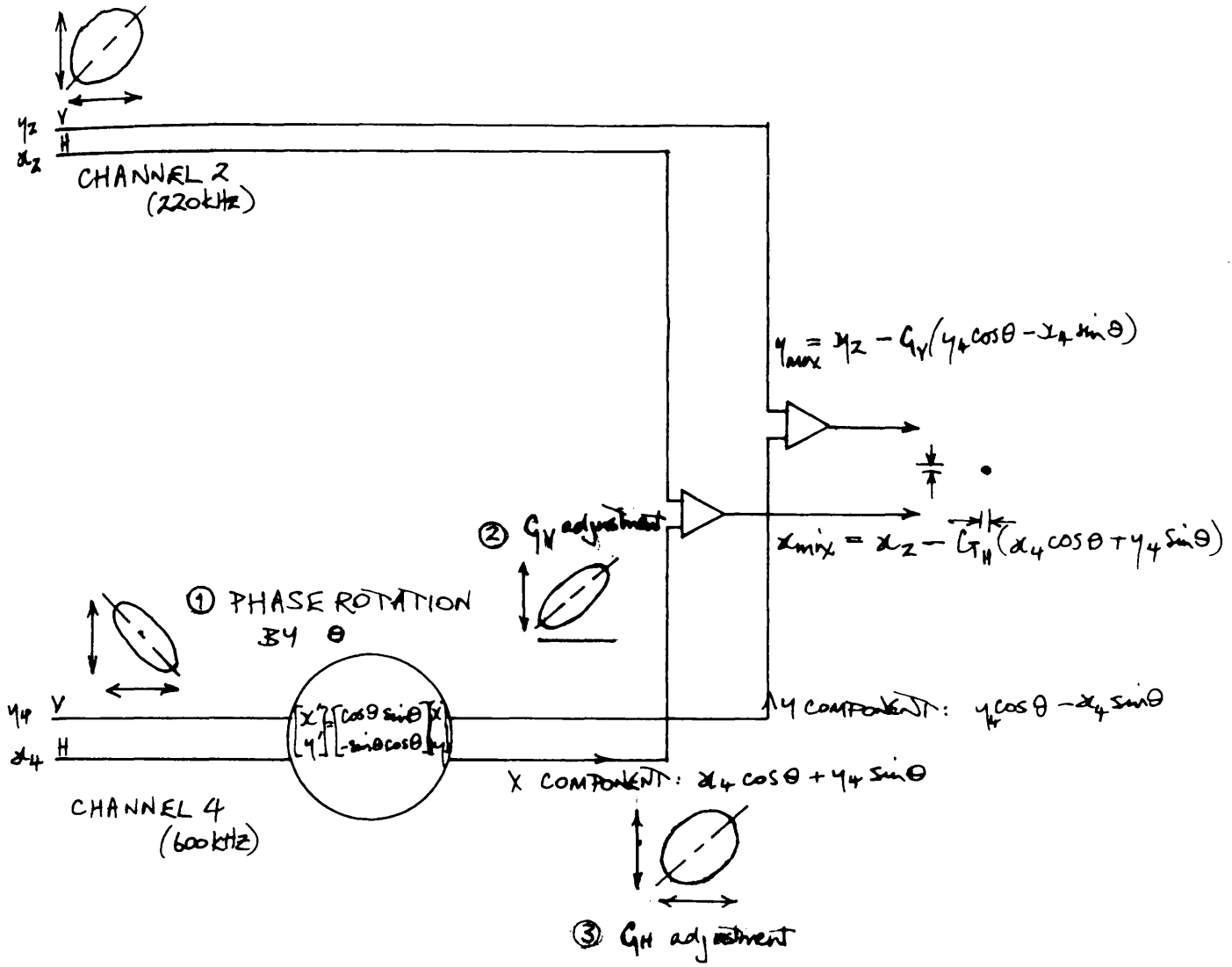



Fig. 2.8 A conventional mixing scheme (as employed in for example the Hocking instruments), where three controls have to be adjusted in order to cause the locus in channel 2 to subtractrively cancel the locus in channel 1. The three controls are the degree of rotation of the x/y axes about the null point, and the vertical and horizontal gains, $G_V$ and $G_H$ all of which are applied to channel 2 only.

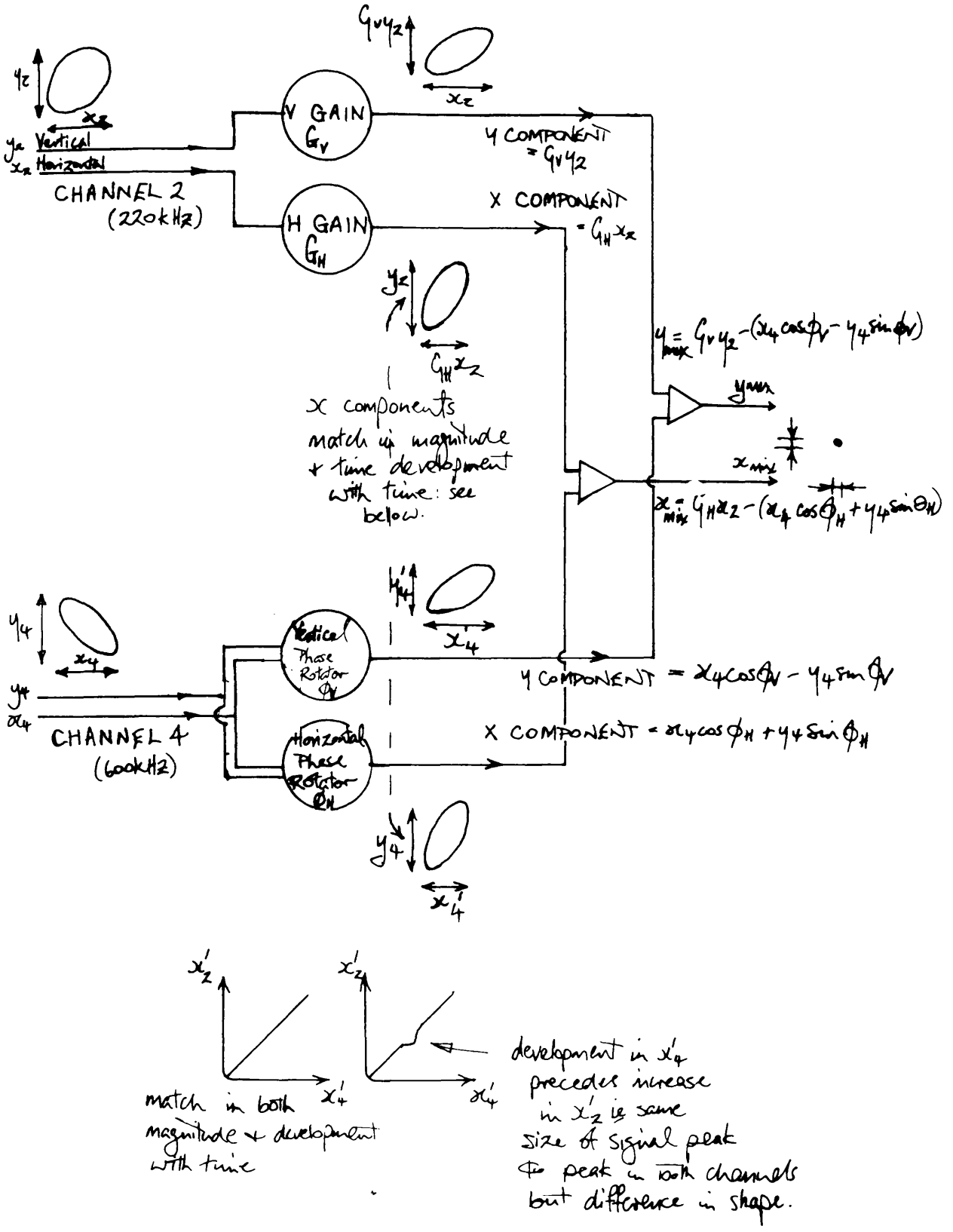



Fig. 2.9 The mixing scheme used by the Zetec MIZ-12. The X and Y signals are treated separately. For each of them individually, a phase rotator and a gain are adjusted in order that the rotated and scaled channel 2 signal matches its compatriot in channel 1. A match is indicated by the plot of the two x or two y components resulting in a 45° line. This is popularly believed to be easier to achieve in practice because there are only two controls to be adjusted to optimise each x/x' or y/y' plot, and because the x and y plots are treated in isolation from each other. The effect of the unusual phase rotators is shown separately in fig. 2.10.


Fig. 2.1Ø The action of the Zetec phase rotator upon a synthetic differential probe locus. It is shown enlarged in the first frame to make clear the asymmetry of the uppermost lobe about its major axis. $\phi_H$ is advanced from 0° to 36Ø°, in 16 equal steps of 22.5° whilst $\phi_V$ is held constant at Ø°. The effect of this is to swing the figure through 9Ø , and then back again, but in the last 8 frames ($\phi_H$ > 18Ø°), the direction of description of the locus is reversed, as shown in the comparison of $\phi_H$ = 22.5° and $\phi_H$ = 247.5°.

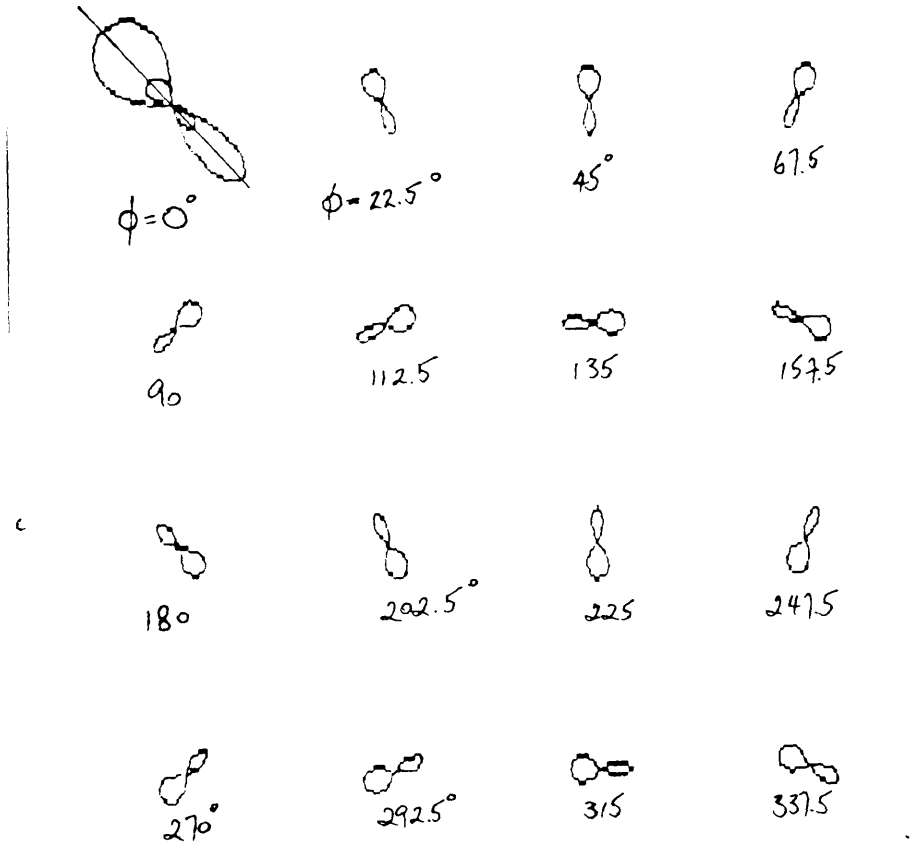



Fig. 2.11 The action of a more conventional phase rotator (e.g as found in a Hocking instrument) on the same figure as shown in fig. 2.1Ø. Here there is only one **θ** to adjust, and this has been made to advance from Ø° to 36Ø°, again in 16 equal steps of 22.5°. This time the figure rotates in equal increments, but always with the direction of description of the figure in the same sense.

**Table 2.1: Rotating vs.Bobbin Probe**

**Rotating Surface Probe**

**Advantages**

1. axial and circumferential information can be retrieved
2. higher resolution than bobbin type
3. sprung probe tip can negotiate dents and tube roll without liftoff

**Disadvantages**

1. need slip rings connect to probe drive circuitry (increased noise) or rotating core transformer (bulky)
2. axial inspection speed reduced by a factor of $\frac{\pi \cdot (\text{bore diam})}{(\text{helix pitch})}$
3. penetration limited by probe coil diameter
4. poor sensitivity to shallow surface defects
5. poor rejection of probe wobble/pilger noise due to smaller field extent

**Circumferential/Bobbin Probes**

**Advantages**

1. less sensitive to lift off due to averaging nature of coupling geometry
2. no slipsringsrequired
3. less wear, heating and thermal drift since coil does not require such tight coupling to surface
4. simpler drive mechanics: no need to rotate probe.
5. standard skin depth and phase lag equations hold because field is approximately planar

**Disadvantages**

1. No angular resolution around tube wall
2. periphery obstructed by dents
3. not good at inspecting tube roll

**Table 2.2: Absolute vs.Differential probe**

**Absolute Probe**

**Advantages**

1. respond to both sudden and gradual phenomena
2. simple interpretation of both impedance plane and tube plots
3. total extent of defect obvious: indicated by non-null impedance value
4. easier to manufacture than differential probes.
5. fewer wires connecting probe to outside world
6. equally sensitive in all directions around probe tip

**Disadvantages**

1. null point prone to drift, due to friction and/or coil resistive heating

**Differential Probes**

**Advantages**

1. Immune to thermal drift
2. insensitive to probe wobble

**Disadvantages**

1. sensitive only to beginnings and ends of cracks: insensitive to gradual changes
2. only detects ends of long defects
3. more difficult to interpret, especially around superimposed defects and phenomena
4. only sensitive along direction of differentiation

# 3 Chapter 3: Pattern Recognition Methods

In this chapter, the structure of a pattern recognition system appropriate to the task in question is defined, and the capabilites of the constituent elements required are discussed in detail.

Also within this chapter, experimental work performed on feature extractors based upon the fast Fourier transform, Fourier descriptors, and AR spectral estimation is described, along with work on branch and bound feature selection, criterion functions, and adaptive learning networks.

## 3.1 Introduction

The structure is based upon the classic statistical pattern recognition system of fig. 3.1, where the flow of information is from sensor (the eddy current system) through preprocessor (the defect detector i.e. segmentation program) to feature extractor, and finally to classifier. In this section, the function of each of these component parts is defined in more detail.

The block diagram of fig. 3.1 represents the operation of a pattern recognition system in its final operational form. During the development stage, more stages are required, and these are shown in fig. 3.2. Essentially, this diagram shows how much more emphasis there is put upon displaying the data at every point in the investigative stages than in the final "production" system. Also, it can be seen that the data set is split into a training set and a testing set. This is characteristic of a statistical pattern recognition system, in that the difference between classes of pattern points is described by the statistics of each class of points (or their class probability

distribution functions) and the use of a "statistical" classifier must therefore include training upon a large set of points representing the full spread of variation within each of the classes. There is an alternative type of pattern recognition system, syntactic or contextual pattern recognition, but this type has not been used in this thesis. The reasons for this are discussed later.

The form of the diagram in fig. 3.1 implies that such a system would immediately process and classify any defect signal passed to the feature extractor, whereas in fig. 3.2, the mention of files at all intermediate points suggests that the system in that case would be operated off-line, i.e. not synchronised to the probe or producing classifications in real time, but processing the signals in batches. The ultimate aim of this project is the development of a site portable unit, which will probably have to be capable of being trained by staff not expert in the design of pattern recognition systems. It will thus have to run a reduced form of the procedural block diagram in fig. 3.2, and then adapt to the production system of fig. 3.1. The original work of this chapter centres upon the development of methods capable of use in such an automatic feature set development process (like the one in fig. 3.3, which illustrates the manner in which the system of fig. 3.2 would be transformed to that of fig. 3.1). The ability to adapt from one type of system to the other is essential when the signals for each class of defect can change between one set of tubes and another, and so on site training becomes necessary.

### 3.1.1 Some Definitions

In this problem, the preprocessor is responsible for passing only those signal segments which represent defects to the feature extractor, and for any signal

processing which may be necessary upon the raw input signal to prepare it for the extraction of features.

The feature extractor makes a number of measurements upon the signal segment passed from the preprocessor, forming a vector of real-valued features, known as a pattern point or sample vector. Often, the features extracted are ad-hoc or heuristic, in that they are chosen by the designer in a non-systematic manner, based upon past experience. In view of the necessity that a working system be trainable on site by people unversed in featureset design, effort has been directed towards the design of feature vectors which encode all of the information about the signal.

After using such a scheme, it will be necessary to reduce the dimensionality of the feature vector, since often the essential differences between different classes of signal can be adequately described by a subset of the full featureset.

Feature selection reduces the dimensionality of the measurement vector. It is done by investigating the ability of different subsets of the set of features to descriminate betweeen pattern classes. Various increases of class separation exist in order to find the best subset. Most measures assume normal class distributions and so are not suitable for multimodal distributions, or situations where nothing is known about the form of the distributions. None of the measures commonly discussed in the classic pattern recognition texts are suitable in practice in these cases. A new measure, which makes no assumptions about the form of the class distributions, has been devised by the author, and is described in section 3.4.3

To finish the process, the classifier must attribute class membership of one particular class to a new pattern of unknown class, based either upon the proximity of stored patterns points (nearest neighbour classification), or upon class boundaries calculated from a training set (decision function-based classification). The former method is more robust under conditions of unknown class distribution (see section 3.5.2), and so this type of classifier finds favour in the conclusion sections of the main experimental chapters (4 & 5).

### 3.1.2 Practical Considerations

This project is intended to spawn a field portable instrument in the near future. This puts a special emphasis upon the robustness of the algorithms used in each of the four areas. They must all be capable of operating satisfactorily regardless of the types of defect present in the tubing set to be inspected.

For instance the shape of the loci associated with each class of defect depend to some extent upon factors other than the shape of the defects themselves, and the relationship between the loci for different defect types and depths. In particular, the tubing's conductivity and permeability vary with time, due to the high temperatures and pressures experienced, and there may be differences between steam generators: different tube wall dimensions will make the shape of the locus change. Hence it is necessary that, from the feature extractor onwards in the system of both figs. 3.1 & 3.2, each of the sections be capable of coping with variations in the relative significance of individual features in the featureset, and variations in the form of the defect class distributions.

This implies that

(i) the segmentation threshold must depend upon and adapt to the background noise,

(ii) the feature extraction scheme cannot be ad hoc, since by the nature of an ad hoc system, it will be attuned to only the training data upon which it was developed.

and

(iii) neither the feature selection criterion function, nor any feature space rotation or

(iv) any classifier must make any assumptions about the form of the class distributions.

Each one of these 4 topics will now be discussed in more detail.

### 3.2 Preprocessing

The preprocessor is needed in order to isolate defects from clean tubing (in order to prevent confusion in the later stages by requiring classification of signals which do not represent defects).

For the particular problem of EC signals from helically scanned defects, using a non-directionally sensitive probe (see section 2.4 and chapter 5 upon absolute pancake probes) the preprocessor must also re-orientate the direction of scanning by means of a buffer store, from circumferential scan to axial scan, in order to get the same type of signal from axially symmetric defects as for the other types, so that a common mix and/or threshold and/or feature extractor may be used over all types of defect.

Certain types of feature extractor (such as that used in section 3.5.1), use an FFT and treat the signal as a series of complex valued samples in the impedance plane, a constant spacing apart in time. They require an input series of fixed length i.e. with the number of points a (fixed) integer power of two. The differing characteristic widths of the various defect types cause different lengths of time series to be passed by the segmentor. Thus the preprocessor may be required to equalise the length of time series fed to the feature extractor, either by widening the limits of the section which was segmented, or by "zero-padding" both ends of the sample sequence.

In the feasibility study preceding this work, **Macleod**(1982) used hand segmentation, feeding the angular and axial extent of defects to the feature extractor as a data file, to be read by the feature extraction program at run time whilst the tube data files were being processed. He suggested a simple segmentation technique which used an initially high amplitude threshold, which having been exceeded, was reduced exponentially to allow all of the defect signal to pass.

The scheme was tested upon the set of defects of Chapter 4 by this author, but the scheme was found to work on pronounced defects, but not sufficiently well to catch small defects, since lowering the initial amplitude threshold tended to pass patches of noise as well. In retrospect, this problem might have been alleviated by the inclusion of a median filter prior to the thresholding stage (an n-sample median filter is similar to an n-sample moving average filter except that the n previous input samples are sorted into ascending order by value and the output is the median valued input sample). This would have prevented the spikes in the noise from triggering the

exponential decay, as opposed to the genuine rise in signal level over tens of sample points associated with a true defect. A more robust ad-hoc algorithm, developed by the author to suit the bipolar signal from a differential probe, is given in Chapter 4.

Two segmentation algorithms have been developed and tested. One is given in Chapter 4, designed for use on the signal from a single frequency differential probe. The other is given in Chapter 5 for a helical scan absolute probe with multifrequency excitation.

### 3.2.1 Published Work

In most papers published upon automatic defect classification, no mention is made of the manner in which automatic defect detection was achieved. This tends to imply that other authors have segmented defects out by hand, by visual inspection of tube and scan plots, and that they were more concerned with classification than with automation of the inspection and data storage process. The systems of CNRL, Vincotte and the CEGB (referred to in section 1.9) are exceptions to this. The detection criteria used by these systems (previously described in section 1.9) are amplitude-based, rather than z-plane based. They make no use of phase information to discriminate between the signals due to defects and other phenomena in the impedance plane.

In the more general pattern recognition and signal processing literature (not specific to automated NDT) not specific to automated NDT, much more work has been published upon image segmentation than on time series

segmentation (see for example, **Rosenfeld and Kak**(1977)). No single technique seems pre-eminent in time series segmentation, since most authors have adopted an ad-hoc method. This is not surprising, since a good ad-hoc solution is only as complicated as it needs to be. Simplicity is important, since in systems operating in real time it is important that the processor be able to apply the segmentation criteria fast enough to be able to handle data at the same rate as the sensor producing it. Unfortunately, few simple algorithms for the segmentation of phenomena in time series have been published in the open literature, perhaps to maintain commercial confidentiality, but the few that have been are now considered briefly.

**Bodenstein and Praetorius**(1977) have segmented abnormalities in EEG waveforms using changes in spectral energy computed over short time segments, when the signal has already exceeded an amplitude threshold. **Mottl and Muchnik**(1979) developed an algorithm for computing the complexity of a segment from the number of zero crossings within it. These papers are typical of the complexity of published time series segmentation techniques, and it was decided that such methods were overcomplicated for this work and also inappropriate because they would not have been suited to the eventually intended real time implementation. Furthermore they do not take advantage of any characteristics (such as the phase characteristics) specific to the types of signal involved.

One simple method considered was that of **Logan**(1973), which treated the sampled reflectivity values from a raster scan of sheet metal as a 2 dimensional image. If while advancing through the scans during segmentation, a point exceeded the amplitude threshold, every point within a rectangular

boundary was tested, and if any any other point was set, the whole area within the boundary was deemed to belong to the defect.

### 3.2.2 Feasibility of Image Processing

An image processing approach to the tube plot was considered, treating the real and imaginary amplitudes as two separate images, with amplitude linearly related to intensity. The results of preliminary work on a PDP11/45 at the MIRU in Edinburgh are included in Chapter 4, in fig. 4.1., where the plots inspired the segmentation algorithm used there.

Ultimately, image processing was deemed to be unsuitable for this project, for 2 reasons:

(a) the cost of special purpose image processing hardware or the processing time requirements for a fully serial approach peformed in software, and

(b) the successful development of a segmentation scheme which worked acceptably upon individual scans, without needing to refer to adjacent ones. (The successful development of a featureset and classifier which worked acceptably well upon single scans also contributed to this decision).

Some software was developed in order to assess the feasibility of image processing on the computer hardware available in the department. It was experimentally determined that the 32 kwords of memory on the PDP11/45 could not accomodate more than 12 single frequency scans (assuming 400points/channel/scan and 1 byte /sample) because of the

size of the Fortran object code, and of the Teklib graphics library available. Overlaying caused bugs to appear in the graphics due to the overwriting of code with data which were in practice not traceable given the software development facilities available.

The feasibility of using the departmental GEC4180 was also investigated. It had virtual memory management and thus practically no limit to memory and supposedly one of its graphics libraries, FINGS, allowed 128 level variable intensity drawing upon a T4014 terminal. Experimentally it actually appeared to have only about 6 discernible intensity levels (i.e. brightness level 110 looked like 120, but level 100 was less bright than 120) and FINGS was no longer supported (having been replaced by GINO). Perhaps of greater significance was the fact that a move onto the 4180, which was deemed strictly not for real time or data logging use, would have been a move away from a site portable system.

### 3.3 Feature Extraction

For all of the feature extractors in this section, it is tacitly assumed that enough information is contained in the impedance plane locus from one single scan across a defect to discriminate between different types of defect, provided that the orientation of the scan is sympathetic to the type of defect. A preprocessor conforming to requirements of the previous section would ensure that a null locus was never passed to the feature extractor i.e. a locus that has degenerated to a dot, as it would if the scan represented one revolution over the centre of an axi-symmetric band-type defect. It is also assumed that the shape of the whole locus is available, not just the magnitude and phase of the

outermost point in the impedance plane, and (in the case of the parametrically based feature extractors) that if the locus cannot be reconstructed, the features do not hold enough power.

### 3.3.1 Existing Work

**Doctor and Harrington**(1980) have classified 3 types of defect in 304 austenitic stainless steel tubing from differential bobbin coil signals, with an 86% correct classification rate. The featureset used was an ad-hoc one: selected ordinates in the autocorrelation series, and power spectral density features describing the spread of power across the spectrum in terms of area under the power spectral density curve.

**Macleod** (1982) used a similarly ad hoc featureset: geometric features describing the length, tilt and fatness of the characteristic figure-of-eight, and spectral features, to describe the distribution of spectral energy in the frequency domain. These features are discussed in detail in chapter 4 and also listed in the feasibility study.

**Satish and Lord** (1984) have attempted to describe impedance plane loci with a Fourier Descriptor-based featureset and have applied the technique to the loci from both differential, and absolute bobbin probes in Inconel 600 tubing. Some experiments with Fourier descriptors synthesised defect loci have been performed. These experiments are discussed in section 3.5.2.

**Satish** (1983) has reported the use of type 2 Fourier Descriptor coefficients on ID differential bobbin probe loci. He succeeded in grouping a total of 30 defects due to 10 different hole/tube plate configurations into 10

"clusters", using a feature vector comprised of the shape rotation, translation and dilation invariants type 2 coefficients, using a k-means clustering algorithm.

### 3.3.2 Parametric Feature Extraction Methods

The description of the perimeter of two dimensional planar objects has merited considerable interest in pattern recognition(**Pavlidis**,1980), particularly in character recognition.

Parametrically based descriptions are the basis of several possible methods for doing this. In essence, a function is chosen which is capable of describing any of the possible boundaries by variation of the coefficients in that function. A very crude scheme might use a circle to describe the boundaries of shapes in the x-y plane in terms of the radius and centre of the circle which had the best fit. This would obviously lead to considerable errors of fit for shapes other than circular ones. The Fourier descriptors mentioned above (and in more detail in section 3.3.7) are an extension of the simple technique above, in that they attempt to describe boundaries by a series of circles of variable radius and starting point, of ascending rotational frequency.

Parametric descriptions in general have found favour in many applications because of their ability to represent a data series with some prespecified function, with significantly fewer parameters than the number of samples originally defining the shape, whilst retaining the ability to reproduce the essential form of the shape.

### 3.3.3 Polynomial fitting-based Feature Extractors

The parametrically based feature extraction methods in this sub-section are all based upon polynomial fitting. In the following sub-sections experiments with transform based descriptions are discussed.

One simple parametric description which was amongst the first considered, describes the position of a point tracing out the impedance plane locus as a function the phase angle, :

$$X(i)=a_o+a_1\phi(i)+a_2[\phi(i)]^2 \qquad (3.1)$$

$$Y(i)=b_o+b_1\phi(i)+b_2[\phi(i)]^2 \qquad (3.2)$$

where the x and y components were treated as separate polynomial functions of $\phi$, and i denotes the sample number.

Unfortunately the parameters are sensitive to the position of the locus with respect to the origin, and to its size, its orientation with respect to the horizontal, and also to the starting point within the locus. The first and last of these shortcomings are not really a problem in this application, because with a differential probe a defect locus will always start and finish at the origin, and with an absolute probe the end points may well not be coincident with the origin but the segmentation algorithm developed in Chapter 5 will relocate the locus accordingly. The problems regarding the scale and angle cannot be so easily solved, and the functions cannot cope with curves which are a non-monotonically increasing function of $\phi$ (e.g. loci which fold back upon themselves or contain loops).

A method which overcomes the objection regarding phase angle has been used experimentally and is described in Chapter 5. Each locus to be analysed is first

rotated clockwise about its segment origin (ie the first sample), until its extremal point lies along the x-axis. A polynomial fit of the form

$$Y=\sum_{n=0}^{5} a_n X^n \qquad (3.3)$$

is then fitted to the outgoing trajectory only, (so that the difference between it and inbound one can later be tested for potential usefulness as a feature). A fifth order polynomial in x was thought to reflect the most complex curvature that the curve might attain whilst giving a measure of smoothing.

The experimental results of this approach are given in Chapter 5. They suggest that normalising all loci to have unit length from one extremity to the other in order to eliminate the confusing effects of defect volume upon locus size, would increase the power of this featureset. Also, in Chapter 5, any width possesed by the locus is shown to be spurious from a discriminatory viewpoint; thus the application of the above approach to loci mixed for the elimination of Pilger noise would probably have given a more distinct class structure to the clusters of samples in the featurespace.

The simple methods given above all lack the ability to describe a locus which contains more than one lobe. Such loci always result from the use of a differential probe, but would also be encountered with an absolute probe when a defect is close to a tube plate edge. Under such circumstances, the lobes become coupled and the locus becomes multilobed. **Brown**'s (1982) major report upon multifrequency studies contains many examples of 4 lobed loci from a differential probe scanning holes near a tube support place (see for example **Brown**'s fig 4.12).

### 3.3.4 Feature Extractor Testing

For any parametric description of a locus, the parameters should be unique for one particular shape, insensitive to any discrepancies in size, orientation or translation between two versions of the same shape, and should be capable of recontructing the locus.

Computationally generated synthetic loci are useful in testing whether or not a proposed parameter set satisfies these criteria, because shifted, rotated and dilated versions of the same locus can easily be generated. Ideally, the behaviour of the generated shape with variations in the parameters of the generating function should be known quantitatively, along with the relationship between those parameters and the ones in the parametric description being tested, in order to be able to predict which ordinates in the parametric description will make the best features for given variations in shape.

Only one other author has considered functional approximation of defect loci. **Satish**(1983) has attempted to simulate defect loci from a differential probe with the lemniscate of Bernoulli because it is possible to use it to generate a figure of eight similar to that produced by a differential probe (fig 3.4). It is normally expressed as a function where 0 is allowed to increase monotonically over the range 0 to 360°

$$r^2=a^2\cos 2\theta, \{\theta:0^\circ \leqslant \theta \leqslant 360^\circ\} \qquad (3.4)$$

where a is the half length of the major axis, as shown in figure 3.4. In order for the function to describe a figure of eight in the same direction as the signal from a differential probe does, one lobe must be desribed clockwise and the

other anticlockwise, and $\theta$ must be made to vary over the ranges $[-45^{\bullet},45^{\bullet}]$ and $[135^{\bullet},225^{\bullet}]$, increasing in one of these intervals and decreasing in the other, though this is not mentioned by **Satish.**

### 3.3.5 Transform-Based Feature Extractors

Methods of this type are capable of describing a curve in which the locus crosses over its previous path i.e. where the locus is multilobed unlike the methods already mentioned. They are for example the Discrete Fourier Transform (DFT), Fourier Descriptors (FD's) and Autoregressive (AR) Spectral Estimation These are methods which start from the set of points in a locus, and transform them into another set of points (which are usually ordinates in the frequency domain). The previous methods operated by fitting a function to the points in the original impedance plane. Each of the three methods is now discussed.

### 3.3.6 Discrete Fourier Transforms

Methods based upon the DFT model a closed curve as the sum of a set of phasors rotating at a series of rates which evenly span the range of rotational frequencies from $-f_{sampling/2}$ to $f_{sampling/2}$, equally spaced across this range at intervals of $f_{sampling/N}$ where N is the number of samples in the time domain. There are thus N/2 phasors rotating clockwise and a further N/2 going anticlockwise, and any shape can be parameterised into this form, providing that it has been sampled finely enough. The parameters in such a description are the radius of, and the angular starting point within each

circle in the series, which are more commonly called the magnitude and phase of each ordinate in the DFT.

By treating a locus as a series of points, (x(t),y(t)), representing samples taken at equal spacing in time, a set of parameters can be generated, using the DFT

$$F(n\Omega) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} f(kT) e^{-j\Omega T nk} \qquad n = -\frac{N}{2}, \ldots, \frac{N}{2}, \; T = \frac{1}{\Omega} \quad (3.5)$$

and the locus can be reconstructed using the inverse transform

$$f(kT) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} F(n\Omega) e^{j\Omega T nk} \qquad k = 0, \ldots, N-1 \quad (3.6)$$

The number of parameters pairs, $\{Re(F(n\Omega)), Im(F(n\Omega))\}$, $n=0,\ldots,(N-1)$, generated is the same as the number of sample point pairs, but the tendency of the spectral energy to concentrate in the lower frequency ordinates (ie the tendency of the magnitudes of the low order ordinates to be largest), means that the figure can be approximately described with just a few of them. Also, the ordinate magnitudes are unaffected by the direction of rotation around the locus (although this makes little difference in this application since the segmented loci always start and finish at the signal origin), and are unaffected by the starting point (i.e. they are invariant under phase shifting). It is this invariance property, however, which accounts for the immense popularity of the DFT or FFT in signal analysis.

There are several problems with using the DFT. The DFT is slow to perform. In order to speed up the production of the transform, the FFT is used. This restricts the length of the input time series to some particular length or number of points, which is usually an integer power of 2. The segments passed from the segmentor vary in width and so the series must usually be zero

padded in order to fit the transform length.

Experimentally, the first thing to be done with the DFT (i.e. the FFT) was to check that it did work. This was done be performing a forward transform on an array containing a few cycles of real valued sine wave, and then applying the inverse transfrom to reconstruct the wave. The frequency domain representation was as expected: two positive going peaks at a position corresponding to the period of the sine wave in terms of the sampling period, and two antisymmetrical peaks along the imaginary axis. The lemniscate of Bernoulli in fig 3.5 was also correctly reconstructed, using the full 64 frequency domain ordinates generated by the transform. In order to test the compression of shape information into the lower frequency ordinates of the DFT, an attempt was made to reconstruct a 64 point lemniscate of Bernoulli (modified to be asymmetrical as shown in fig. 3.6) from a truncated version of the frequency domain ordinates. The reconstructed form degenerated into the form illustrated in fig. 3.7, even when as many as the 16 lowest frequency ordinates from the original 64 point FFT were used. The resemblance of the reconstructed figure to the original was dissapointing in view of the small magnitudes of the high frequency ordinates (see fig. 3.6(b)), especially given that the DFT is optimal in the sense that the reconstruction error (ie the mean squared error over the whole locus) is minimal for a reconstruction from a truncated frequency domain with respect to any other truncated parameterisation of the same order.

In the later set of logged tubes, in Chapter 5, the scan length is typically 55 samples. This sampling rate is the minimum necessary to retain the detail in the signal (i.e. to avoid aliasing). Some defects, for example patches of

IGA, last for more than 32 points, so a DFT/FFT transform length of 64 points was necessary (there was not enough time available for experimentation with FFT's whose length is not an integer power of 2). The possibility of compressing or expanding each individual segment to fit a 32 point transform was considered, but since the width and rise times for certain defects were thought to have discriminatory potential, this was decided against. No further work with the DFT was done, in favour of concentrating upon finding a method which would yield a feature vector of lower dimensionality - a factor which was of considerable importance given the small defect sets in both Chapters 4 and 5.

### 3.3.7 Fourier Descriptors

Fourier descriptors are a variation on the discrete Fourier transform. Essentially they transform from arc length, instead of time, to frequency.

As explained in section 3.3.6 problems have been experienced with truncating the FFT ordinate series into manageably short feature vectors. Several authors have claimed that considering the instantaneous position of the point tracing out a boundary (or impedance plane locus) as a function of arc length (as opposed a function of time) helps to compress the spectral power into the low frequency ordinates of the transform. Consequently, Fourier Descriptors were chosen as a candidate feature set, and an investigation into the details of the method was undertaken in order to implement forward and inverse Fourier Descriptor transforms in the form of subroutines.

The work in this section was originally intended to replicate that of **Satish**, and then apply it to the larger and more realistically generated defect set in hand. **Satish's** thesis (1983) contained some intriguing claims for the reconstructability of impedance plane loci (which would allow checking of both the transform generating program and of the parametric description itself by inversion and comparison with the original), and for the insensitivity of the Fourier descriptors to nonconstant sample spacing of the samples along the length of the arc (or locus). Essentially, this means that FD's are claimed to be insensitive to the probe speed. However, in attempting to get the requisite programs to work, both of the above claims were found to be unjustified. This section now considers why the technique did not work.

### 3.3.7.1 Type 1 Fourier Descriptors

There are two types of transform, known as type 1 and 2. In Type 1 Fourier Descriptors, originally devised by **Cosgriff**(1960), but first described in the open literature by **Zahn and Roskies**(1972), the forward transform acts upon the tangential angle, treating it as a function of arc length. The Fourier descriptors are defined as a set of complex valued coefficients, $S_n$, defined in most of the literature in terms of t, where t is the cumulative arc length. Thus

$$S_n = \frac{1}{2\pi} \int_0^{2\pi} \phi(t)\, e^{-jnt}\, dt \qquad (3.6)$$

where $\phi(t)$ is tangential angle, and t is arc length. Since $\phi(t)$ is a real valued series, there will be only N/2 distinct coefficients, when there are N samples in the input polygon, due to antisymmetry of the Fourier transform of a real-valued time series. Practical computation of the coefficients in the

discrete form (**Zahn & Roskies**, 1972) is achieved by defining a function $\phi^*(t)$, based upon the tangential angle at each point t, which is normalised to vary only between 0 and $2\pi$, defined as

$$\phi^*(t)=\phi(\frac{Lt}{2\pi})+t \qquad (3.7)$$

($L$ = total arc length)

$\phi^*(t)$ essentially measures the difference between the figure and a perfect circle, being zero everywhere for a circle. $\phi^*(t)$ can be expanded as a Fourier series (shown here truncated to length 2m):

$$\phi^*(t)=a_o+\sum_{k=-m}^{m} a_k\cos(k\omega t)+b_k\sin(k\omega t) \qquad (3.8)$$

where the coefficients are defined

$$a_k = -\frac{1}{k\pi}\sum_{i=1}^{m-1}\Delta\phi_i \sin\left(\frac{2\pi k L_i}{L}\right) \qquad (3.9)$$

$$b_k = \frac{1}{k\pi}\sum_{i=1}^{m}\Delta\phi_i \cos\left(\frac{2\pi k L_i}{L}\right) \qquad (3.10)$$

$$a_o = -\pi - \frac{1}{L}\sum_{i=1}^{m-1}\Delta\phi_i L_i \qquad (3.11)$$

or using the polar form,

$$\phi^* = a_o+\sum_{k=-m}^{m} A_k\cos(k\omega t-\alpha_k) \qquad (3.12)$$

where

$$A_k = \left(a_k^2+b_k^2\right)^{\frac{1}{2}} \qquad (3.13)$$

and

$$\alpha_k = \tan^{-1}\frac{b_k}{a_k} \qquad (3.14)$$

The chief advantage of this approach is that the set of coefficients $\{A_k,\alpha_k\}$ is unique for a given shape, regardless of the scale, angular orientation, or distance of that shape from the origin. Morover, the magnitudes, $A_k$, are independent of the starting point within a locus.

### 3.3.7.2 Criticisms of the Method

The disadvantages of this method are that the full reconstruction integral is complicated and that reconstruction of a closed locus (which tends to imply that it is approximately similar to the original) is not guaranteed with a truncated parameter set. The definition of the tangential angle $\phi^*(t)$ gives rise to a function like a sawtooth wave which contains frequent discontinuities (see fig.2, **Zahn and Roskies**), and so the spectrum has a substantial high frequency content, making truncation inadvisable because of the amount of information that is likely to be lost.

### 3.3.7.3 Type 2 Fourier Descriptors

In type 2 Fourier descriptors (**Granlund**, 1972), the position of a point tracing out the boundary is defined as a function of the arc length. The Fourier descriptors can be expressed as a series of N parameters,

$$T_n = \frac{1}{L}\int_0^L b(t)\, e^{-j(2\pi n/L)}\, dt, \quad n=1,\cdots,N \qquad (3.15)$$

where b(t) is some function of the arc length t, usually the co-ordinates of the boundary position itself (i.e. of the complex-valued signal) U(t)=x(t)+iy(t), and L is the total arc length.

Twice as many coefficients are produced by the type2 method as with the type 1 method, because U(t) is a complex valued series and thus there is no antisymmetry of the coefficients. This ought to be more than compensated for by the low frequency concentration of the spectral power and consequent ease of truncation, due to the smoothness of the locus time series U(t) compared with $\phi^*(t)$.

Granlund gave a semidiscrete form of the transform pair (i.e. where an integer number of ordinates in the spatial frequency domain are derived from a continuous function of cumulative arc length, t) as

$$C_n = \frac{1}{L}\int_0^L u(t)\, e^{-i\frac{2\pi n t}{L}}\, dt \qquad (3.16)$$

$$U(t) = \sum_{n=-\infty}^{\infty} C_n e^{i\frac{2\pi n t}{L}} \qquad (3.17)$$

The continous variable U(t)=x(t)+iy(t) is the boundary position itself, t is cumulative arc length, L the total arc length, and n the order of the frequency domain ordinate. The parameters obtained are not insensitive to scaling, rotation or translation, and so he also devised a set which were. **Persoon and Fu** (1977) devised a method of calculating the coefficients for a polygonal curve. After defining $l_k$ as the accumulated arc length up to the $k^{th}$ vertex $V_k$, and setting $l_o$=0, they defined the set of incremental unit position vectors as

$$b_k = \frac{V_{k+1} - V_k}{|V_{k+1} - V_k|} \qquad (3.18)$$

and the forward transform as

$$C_n = \frac{1}{L\left(\frac{2\pi n}{L}\right)^2}\sum (b_{k-1} - b_k) e^{-i\frac{2\pi n l_k}{L}} \qquad (3.19)$$

whilst claiming that the inverse transform continued to be eqn. 3.17.

A subroutine was coded to calculate the Fourier Descriptors according to eqns. 3.18 and 3.19. It was tested by feeding the results back into eqn. 3.17, in the form of a test procedure. Results typical of the many attempts made to get the procedures to produce the results which were to be expected according to **Satish**, are shown in fig. 3.8. The reason than the results were as shown seem to be because of a number of flaws in the technique, rather than any flaw with the software, which was extensively tested. These flaws are now

explained.

### 3.3.7.4 Criticisms of the Method

**Satish**(1983) and before him **Persoon and Fu**(1977) advocate the use of the type 2 method because of its supposed ability to reconstruct a closed version of the original locus regardless of the severity with which the full set of coefficients is truncated. With the equations given by these authors, this appears not to be so, for the following reasons:

Firstly, it is implied by both Satish and by Persoon and Fu that eqn. 3.17 and 3.19 form a Fourier pair. For this to be true the term $(b_{k-1}-b_k)$ would have to equal U(t). In fact

$$b_{k-1}-b_k = \frac{dU(t)}{dt}$$

that is the discrete derivative of U(t):

$$\frac{\Delta U(l_k, l_{k-1})}{\Delta(l_k, l_{k-1})} = -\frac{[u(l_k)-u(l_{k-1})]}{(l_k-l_{k-1})} \qquad (3.20)$$

Secondly, the appearance of the $n^2$ weighting in the denominator in eqn. 3.20 is not justified when compared with the original transform, nor is this a valid way of reducing the effect of the high frequency ordinates' contribution to the shape of the reconstructed figure (if that was originally the reason for its inclusion), since this means of de-emphasising the upper ordinates does not mean that the ordinate sequence can then be legitimately truncated. In any case, a similar but inverted term would have to appear inside the argument of the summation operator in eqn. 3.18 for parity of the transform pair to be maintained.

Thirdly, reconstruction of a locus is not possible without knowledge of the incremental arc length series, $l_k$, $k=1,..,n_{points},0$. The effect of non constant sampling along a locus is to produce a nonlinearly scaled frequency axis (provided frequency is defined in the conventional way as proportional to inverse real time), ie the height of the ordinates is correct but the spacing of the ordinates along the f-axis is nonconstant and not simply related to the arc sample spacing. This is an important objection in the present application, because (since the intention of applying Fourier descriptors is to discriminate between locus shapes), the ordinates in the frequency domain must be genuinely comparable: differences in the number of ordinates, and particularly in their spacing, are not tolerable. Even interpolation to synthesise a set of evenly spaced frequency domain ordinates would not be easily accomplished since the relationship between random sample spacing in time and the consequent frequency domain ordinate spacing is not well understood (**Bartlett**,1972). In practice, resampling the locus with a fixed arc length sampling interval, after total arc length normalisation, and then applying a standard FFT subroutine to the point set would be a much simpler task than attempting to apply the rather tortuous theory of point processes.

### 3.3.7.5 Suggested Modifications

In order to make Fourier descriptors really usable, reconstructability of the impedance plane loci, and comparability in the transformed domain must be achieved. Achieving a transform pair which operates directly upon the impedance plane, and can reconstruct a locus, would be best done with an

unmodified FFT, so that the body of well understood theory remains applicable. However, first of all the number of samples has to be equalised in every locus to fit the transform length.

Differences in the number of samples in a locus are bound to occur due to the differing widths of, for example axial notches and circumferential notches. For comparability of transformed coefficients, the number of samples must be the same, in order to produce the same number of coefficients, and the easiest way to achieve this is to resample the locus by linear interpolation.

In addition, for greatest truncatability for an acceptable loss of shape information, a smoothly varying function of arc length is required, which is real valued in order to keep the number of coefficients down as a result of antisymmetry. This excludes the possibility of using the boundary position itself.

It might be argued that the distribution of the samples along the locus for any particular type of defect is as important a function of the type of defect as the shape of the locus. For example, although an axial notch has a locus shape not dissimilar to that of intergranular attack, the concentration of samples within the locus at the extrema will be less pronounced in the case of the IGA. It is not obvious how the sample spacing can be encoded in a Fourier Descriptor feature extractor while at the same time satisfying the requirement to fix the number of Fourier descriptor ordinates and have the arc length samples uniformly spaced. The ideal solution might be to use features from both the time and arc length domain based transforms, for example the ordinates from an FFT plus the coefficients of a polynomial fitted to a graph of the incremental sample spacing versus sample number. However, such a scheme

which will produce a large number of features requiring a systematic feature selection scheme which can cope with a large number of features, and also with class distributions of an unknown nature. This problem has been addressed in section 3.6.

### 3.3.8 Autoregressive Spectral Estimation

All discrete Fourier transforms impose a fixed length upon the data series, and also require windowing, of the data series (premultiplication by a symmetrically tapering function to de- emphasize the ends of the time series), to reduce leakage. Leakage is the migration of high frequency spectral energy to lower frequencies, due to the abruptness of the ends of a finite length data series. Leakage reduction is achieved at the expense of spectral resolution.

**Burg** in his seminal work (1967) upon the maximum entropy method (a particular autoregressive spectral estimation method) advocated his approach because it did not impose a fixed length sample series upon the input data and did not require the data to be windowed (and hence be corrupted before even reaching the transform).

Autoregressive spectral estimation attempts to model a signal as a linear combination of previous outputs, where the driving input is a white noise source: $\phi_n$, and

$$y_k = \sum_{k=1}^{m} a_k y_{n-k} + \phi_n \quad n=1,..,m \qquad (3.21)$$

The spectrum G(f) of such a process can be approximated by

$$G_m(f) = \frac{2\sigma_\varepsilon^2}{\left|1-\sum_{k=1}^{m} a_k \exp(-2nikf)\right|^2} \qquad (3.22)$$

The parameter set $\{a_o,..a_m\}$ describe the form of the power spectral density, and so can be used as a feature vector. The method is popular with pattern recognition workers in geophysics (**Tjostheim**,1977).

Calculation of the set of coefficients can be a problem, given that the order of the model, m, is not usually known in advance in most systems, although **Akaike**'s Final Prediction Error (1969) and other similar measures claim to be able to determine it.

Two experiments were performed, one with the method of **Ulrich and Bishop** (1975) (their paper contained a Fortran implementation), and the other using two subroutines only recently introduced in the NAG mathematical subroutine library, G13ABF and G13ACF. The first of these 2 subroutines calculates the autocorrelation series of the data, and the second uses this series to obtain the AR parameters, for a given model order, which must be supplied by the calling program. Both methods were successfully tested upon a simple sine wave and upon the simple 6th order autoregressive process given by **Ulrich and Bishop** in their paper as a test procedure.

However, when tested upon scans containing intergranular stress corrosion cracking and axial cracking defects from Chapter 4 (see figs. 4.7 and 4.8), both methods gave spectra which decreased smoothly over the range of frequencies from d.c. to the Nyquist limit, regardless of the order of the model employed. All orders of m were tried, from 1 to 20. There was little difference in the observed spectra between the two types of defects. This was discouraging

since the FFT spectra for the same scans were clearly different. Since the results from the two versions of the technique concerned concurred, it was concluded that the fault was not one of programming logic but rather that the technique was not applicable to the defect signal waveforms. This may have been because within one scan, there is only "one cycle" of defect signal. In **Ulrich and Bishop**'s paper, the test waveforms which were for analysis by the AR method in each case were of several cycles duration.

Alternatively, the shape may possibly not be capable of being modelled by an AR process though it might be by one of the other related methods, such as an autoregressive moving average (ARMA) model, where the present output is the weighted sum both of previous outputs, and of previous process inputs, from the noise source.

**3.3.9 Criticism of the lemniscate of Bernoulli** Bernoulli's lemniscate cannot fully test a parameterisation scheme, such as one based on Fourier Descriptors, because it is symmetrical about its major axis and thus does not test for the ability to represent figures which are not symmetrical about the x or y axes. The manner in which the two lobes of the figure of eight are described in opposite senses is, however, of benefit in replicating the signal for differential probes, so the function was modified to be asymmetrical about the y-axis (see fig. 3.5(a)).

For loci produced with an absolute probe, as used in Chapter 5, a better functional approximation is produced by a Joukowski aerofoil, since it is single lobed, and asymmetrical. The Joukowski aerofoil is produced by a mapping a circle in the complex plane, onto the w-plane using the mapping

$$w=z+(1/z) \qquad (3.24)$$

where z is traditionally a point on the generating function: a unit circle (as opposed to a point in the eddy current probe impedance plane, lest there be any confusion). An aerofoil is produced by the mapping, providing that the generating circle passes through either the point x=+1 or x=-1 and also encloses the one which it does not pass through.

The Joukowski aerofoil resembles the comma shaped loci associated with OD and ID notches and IGA, although simulation of the more complicated loci due to holes (hook shaped) and thinning (s-shaped) could not be achieved.

The shape is better for testing the effect of truncation on a series of Fourier descriptor, since the sharpness of the change of direction at the extremity implies that the figure will contain a greater concentration of high frequencies. Only the accuracy of the low frequency terms of the Fourier descriptor is significant with a lemniscate of Bernoulli. An even distribution of points in the z-plane around the generating circle gives rise to a set of points in the w-plane which mimics the distribution of samples in a real defect locus.

The choice of a comma shaped synthetic locus with non-zero thickness was deliberate, since it was thought that the parametric scheme might be required to operate upon loci which had not had the Pilger noise removed from them, and hence would have non-zero width.

As a consequence of testing with a more complicated shape than the lemniscate of Bernoulli, the Fourier descriptor as described by Satish was found not to be capable of reproducing figures symmetrical about their major

axes.

### 3.3.10 Closing Remarks

Apart from the assumptions given at the beginning of this section, it has also been assumed (in the case of the parametrically based feature extractors anyway) that if the locus cannot be reconstructed, the features do not hold enough power to allow discrimination between classes. This may well not be so, but it has been the only way to validate parametric featuresets available so far. In order to relax the assumption, it is necessary to have a means of testing subsets of the large systematically generated featuresets which the parameteric methods tend to yield. This is addressed in the next section. It has not yet been possible to apply the products of the work described there to the output from feature extractors described here, for reasons given at the end of section 4.4.

## 3.4 Feature Selection

Dimensionality reduction is usually necessary in order to avoid "the curse of dimensionality", which limits the number of measurements for a given sample size. Specifically, if the dimensionality, d, exceeds (n-2)/2, where n represents the number of pattern points per class, then the probability of there existing an arbitrary hyperplane satisfying the desired dichotomisation exceeds 0.5. If this limit is violated, the solution obtained is overfitted to the training points used.

A feature selection scheme uses a class separability measure (or criterion

function), to evaluate subsets of the full feature set produced by a subset generating algorithm, until it is maximised for a subset of the desired (reduced) dimensionality.

### 3.4.1 The Branch & Bound Algorithm

The best subset is not necessarily made up of the best features when those features are considered individually. Thus evaluation of subsets rather than individual features is necessary. Exhausive evaluation of all of the subsets of rank n for an initial dimensionality of N would require

$$\binom{N}{n} = \frac{N!}{n!(N-n)!} \qquad (3.25)$$

evaluations of the criterion function, a massive amount of computation for most realistic situations.

The only existing scheme which implicitly performs an exhaustive search is **Fukunaga and Narendra's** Branch and Bound algorithm (1977). **Roberts** (1985) published an implementation of the algorithm in Fortran. The program has been tested successfully upon 2 class 6-dimensional data with Gaussianly distributed classes, and has been found to select the features in the order expected.

### 3.4.2 Criterion Functions

The most common criterion functions (the Mahalanobis, Bhattacharrya and Divergence measures, for instance) have the form:

$$J = X^t C^{-1} X \qquad (3.26)$$

where J is a scalar quantity representing the distance between the classes, and for a two class situation, typically, X represents the class intermean vector and C is the average class covariance matrix (the precise definitions of X and C depend upon exactly which measure is being used). The popularity of this form of criterion function stems from the fact that J can be recursively recalculated as features are deleted from the subset without the inverse of the matrix C having to be recalculated from scratch after every reduction of rank, thus saving considerable computation **Fukunaga and Narendra**(1977). The proof of the monotonicity of the increase in J with increasing dimensionality of the feature subset is a consequence of the recursive recalculability of J. The function used by Roberts in her implementation of the branch and bound algorithm (the Mahalanobis distance) was also of this form and her program calculates J recursively with reductions in the rank of C and X.

Criterion functions described in the classic pattern recognition texts tend to be defined in terms of class conditional probability distribution functions, and measures of the distance between them involving integrating the difference in these densities over the complete featurespace. The practical versions of these functions are defined in terms of class sample covariance matrices. These typically assume parametrically defined distributions, usually normally distributed classes with equal class covariance matrices.

### 3.4.3 A New Nonparametric Criterion function

The measure given here uses the concept of the local k-NN mean as developed by **Fukunaga and Mantock**(1983), who used it for generation of linear transformations for feature extraction, though only for the two class case. The most popular linear transformations start with the training set and calculating the eigenvectors of the sample based covariance matrix (the eigenvectors represent a set of orthogonal axes aligned with the directions of maximum variance, calculation of which is made simpler by the assumption of normally distributed classes).

The well known Fisher Discriminant Ratio is defined

$$J.=\frac{d^t S_b d}{d^t S_w d}$$

where $S_b$ and $S_w$ are the between and within class scatter matrices respectively, and d, the inter-class-mean vector. The from of this measure has been used as the basis for a new parametrically based measure.

When it is desired to measure the separability of awkwardly distributed classes, ideally one would want to use a criterion function which is an integral of the difference between the class probabilities at every point. A method of doing this is as following.

The class conditional probibility density $p(x/w)$ at any point $(X_i)$ can be deduced from the vector from $X_i$ to the geometric mean of the k nearest neighbours from the appropriate class. Specifically the vector points in the direction that the probability density function rises most steeply, and the length of the vector is proportional to how steeply it rises. This distance

can be represented as inner product of the form

$$(X_i-\mu_{knn})^t(X_i-\mu_{knn})$$

where $\mu_{knn}$ is the geometric mean of the k nearest neighbours. By reversing the order of the multiplication, and pre- and postmultiplying by the vector joining the two local class means from classes j and l, to form:

$$(\mu^j_{knn}-\mu^l_{knn})^t(X_i-\mu_{knn})(X_i-\mu_{knn})^t(\mu^j_{knn}-\mu^l_{knn})$$

the point is projected down onto a line joining the two class means, and the expression represents the distances along this line from the image of the point projected onto the line to the class mean. If these distances are then summed to form a scalar measure in an analogous manner. The summation is

$$J_{knn} = \frac{1}{n_{points}} \sum_{j=1}^{2} \sum_{all\, i \in j} \frac{w_{Xi}(\mu^j_{knn}-\mu^l_{knn})^t(X_i-\mu^l_{knn})(X_i-\mu^l_{knn})^t(\mu_{knn}-\mu^l_{knn})}{(\mu^j_{knn}-\mu^l_{knn})^t(X_i-\mu^j_{knn})(X_i-\mu^j_{knn})^t(\mu^j_{knn}-\mu^l_{knn})} \qquad (3.27)$$

where $\mu_{knn}$ is the local mean for point $X_i$ from class j, and i is the index of that point within class j. A weighting factor $w_{Xi}$ has been introduced that has yet to be explained. The numerator and denominator in eqn. 3.27, when summed across the entire point set in a 2 class problem are analogous to $S_b$ and $S_w$, but the difference lies in the fact that $S_b$ and $S_w$ are defined in terms of whole class sample-based means and not in terms of local means.

A criterion function must monotonically increase with increasing separability of the classes, and with increasing dimensionality of nested feature sets. In order to test that this was so, the function was tested upon 2-d distributions analogous to those shown projected into one dimension in fig. 3.9. Testing could only be done computationally because an analytical treatment of a k nearest neighbour scheme under small sample conditions is very difficult

(**Kittler and Devijver**, 1983) and has not yet been successfully attempted.

The use of a scatter matrix which is defined using local means rather than using means based upon the whole sample set as a basis for the matrix from which to extract the eigenvectors which best display class separability has been shown to be advantageous. The scheme of **Fukunaga and Mantock**(1983) was tested upon the single frequency data set in Chapter 4, and worked better than the **Kittler Young** transformation (1973), which tended to be confused by the closeness of the ZCK and IGSCC classes (Chapter 4) compared to the larger distance between both of these classes and the GCO class. The projection with the new transformation is shown in fig. 4.18. The comparable plot of the points projected into the best two feature axes produced by the Kittler Young transformation is shown in fig. 4.17.

Without the weighting function, the contribution from each point to the total value of the scalar separability measure, is accepted regardless of whether that point falls near to the interclass boundary which would be imposed by a nearest neighbour classifier, or far way from it. A weighting function originally suggested by **Fukunaga & Mantock**

$$w_{Xi} = \frac{\min(d_1,d_2)}{d_1+d_2} \qquad \begin{matrix} d_1=|(X_i-\mu^j_{knn})| \\ d_2=|(X_i-\mu_{knn})| \end{matrix} \qquad (3.28)$$

(and also studied by **Luk**(1986)) de-emphasises the contribution from points near to the boundary. Although admittedly the value of the ratio does fall as a point is moved from a position coincident with its own class's local mean across the boundary towards the other mean (from infinity to one and then to zero), the ratio does still increase the value of the criterion function, even although the point would be wrongly classified by an NN classifier. By

redefining the weighting function

$$w_{Xi} = 0.5 - \frac{d_1}{d_1+d_2} \qquad (3.29)$$

is is possible to actively discrimate against points which occur on the wrong side of the k nearest neighbour classification boundary. This weighting function (eqn. 3.29) has been incorporated in the criterion function (eqn. 3.27) as $w_{Xi}$.

The criterion function $J_{nonparametric}$ has thus been designed to be sensitive to local separation of the classes, and to measure the class separability which would be experienced by a p.d.f.- based classifier if the class conditional p.d.f's were known explicitly. It seems reasonable to suggest that any subspace selected for a given level of performance with this criterion function will perform similarly well with a k-NN classifier with the k set to the same value.

The main disadvantage of the method is that the recursively computatable property of the criterion function has been lost because a reduction of dimensionality is capable of changing exactly which points are the k nearest neighbours to any other point, so that each vector $(X_i-\mu_{knn})$ might change under a reduction of rank as the feature set is pruned, and thus require recalcul ation. Consequently the main terms in eqn. 3.27 are not preserved under a reduction of rank.

Unfortunately the power of the technique could not be properly tested because of the limited size of the PDP11 memory, and so it could not be run on the tube defect feature vectors because of the large number of raw dimensions

available from the various feature extractors (for comparison of the various feature sets attempted as part of the work for Chapter 5) This was because of

(a) the dimensionality, and consequent memory which would be required for work space for the criterion function and feature selection (branch and bound) algorithm

and

(b) a lack of samples to ensure statistical validity of the results even had the memory been available on the pdp11.

When these restrictions are overcome, it should be possible to evaluate very large sets of features to produce a good small subset without having to first go through the feature axis plotting and transformation stage in order to understand the data structure which is necessary in order to interpret the meaning of results from class sample mean-based measures such as the Mahalanobis distance and the Fisher discriminant ratio.

### 3.5 Classification Methods

Basically two types of classifier have been used for the experimental results:

(i) linear and

(ii) nearest neighbour classifiers,

although quadratic and scan cross correlation classifiers have been successfully applied and attempts have also been made to use an N-tuple based classifier and an adaptive learning network (Chapter 4). This section considers three of these types of classifier in more detail. The remainder are explained in Chapter 4.

### 3.5.1 Linear Classification

The operation of a linear classifier is based upon the position of a pattern point of unknown class with respect to a hyperplane, which defines the position of the decision boundary between two classes. When $[w_1,...,w_{n+1}]$ are regarded as the entries of a weights vector W, class membership of a candidate point $X=[x_1,...,x_n]$ is dependent upon the sign of $W^tX$, the weighted sum of all of the feature values ( n is the dimensionality of the feature space). The decision function, $w_1x_1+...w_nx_n+w_{n+1}=0$ at any point on the boundary. The weights $[w_1,..,w_{n+1}]$ in the weghts vector are optimised or trained to divide (or dichotomise) a two class training set into its two classes. Two class linear classifiers are readily generalised to handle c-class (i.e. multiclass) situations by the use of c such weight vectors, providing that the classes are unimodal and the system designer knows enough about the class distributions to be able to program the additional logic necessary to deduce the class to which a point belongs from the signs of each of the weighted sums.

In iterative or "perceptron" methods training proceeds repeatedly through the training set until the number of points being misclassified is minimised. With simple schemes, where the coefficient $\alpha$ is fixed, and the adjustment due to misclassification is of the form $W_{k+1}=W_k+\alpha X$ for the $k^{th}$ misclassified point, instability of the boundary position can be caused by rogue points, resulting in the boundary not settling in the optimum position.

The technique used here, due to **Duda & Fossum** (1966), gradually reduces $\alpha$ with k, but disregards points which are seriously misclassified (by reference to the magnitude of $W^tX$), and only includes points within a dead

band close to the boundary after $\alpha$ has dropped to below a threshold value. **Logan** (1973) found that this technique converged faster upon an acceptable boundary than the simpler techniques already mentioned. This is because simple perceptron training schemes depend crucially upon the assumption that classes are separable by some plane in the feature space. If this assumption is not satisfied, the points which are repeatedly misclassified dominate the convergence behaviour of the dichotomising hyperplane so that it may not settle in the best possible place.

### 3.5.2 Nearest Neighbour Classification

The 1-NN classifier classifies a new pattern as belonging to the same class as that of its nearest neighbour in a stored training set. The k-NN assigns a new pattern to the most heavily represented class amongst its k nearest neighbours. The (k,l)-NN rule further includes a rejection threshold requiring that at least $l$ votes are cast in favour of the chosen class from the k nearest neigbours.

For a given set of features, the Bayesian classifier's misclassification rate, $E^*$, cannot be bettered. Therefore, estimating the performance of a featureset via the Bayesian classifier error rate is equivalent to assessing the featureset's effectiveness alone, untainted by the suitability of the classifier chosen for the task. A major attraction of NN classifiers is that their misclassification error rates can be used to estimate $E^*$, as is now explained.

### 3.5.3 Bayesian Error Rate Estimation

The ideal classifier is the Bayesian classifier. It classifies a point X in featurespace, as belonging to the class which has the greatest probability of occuring at point X. For each class, this probability is the product of the class conditional probability density function $p(X/w_i)$ and the a priori probability $P_i$ that class i (i=1,..,c) occurs. $P_i$ is usually deemed to be either 1/c for all i, or the fraction of the sample set belonging to class $w_i$. Expressed formally, a Bayesian classifier attributes X to $w_i$ from a set of classes $\{w_j, j=1,...,k\}$ if

$$p(X/w_i)P_i > p(X/w_j)P_j, \quad j \neq i \qquad (3.30)$$

Although theoretically capable of coping with complicated distributions, where for instance one class has a cluster of members inside a hole in another class's pdf, this type of classifier is not usually practical to implement, due to the difficulty involved in functionally determining class conditional p.d.f.'s.

Given that a few assumptions are valid (i.e. that a large training set exists, statistically identical to the testing set; that the NN classifications based upon a finite size training set agree with those that would be obtained with an infinite set), bounds upon the Bayesian error, $E^*$, can be derived from nearest neighbour misclassification rates, denoted $E_{k,l}$ where $l$ is the minimum number of votes that any class has to pole in order for a point to be associated with it. If $l$ votes are not poled by any class, then the point is rejected.

A simple bound on $E^*$ is contained in the result that half of the

classification information available in an infinite training set is contained in the first nearest neighbour and therefor that

$$\frac{(E_1)}{2} < E^* \qquad (3.31)$$

i.e. that the best achieveable error rate will be at least as bad as half the error rate achieved with a 1-NN rule classifier. The particular selection of values of k and $\ell$ for which miscellaneous rates are shown, is due to two rules: "that when the last nearest neighbour is of even rank, it does not contribute any additional classification information", and that in order that the k nearest neighbours be genuinely within the vicinity of the point being classified, $k \simeq \sqrt{n_{sample}}$.

**Kittler and Devijver** book (1983) considers the upper and lower bounds upon $E^*$, in terms of the $(k,\ell)$-NN, and the tightest set of bounds which they give are:

$$E_{(2k'-1,k'+1)} < E_{(2k',k'+1)} < E^* < E_{(2k'-1)} \qquad (3.32)$$

where k' is the least integer which satisfies $k'>k/2$.

These bounds have been used in Chapter 4 in an attempt to measure the scope for improvement still remaining with the featuresets and classifiers tested there. However, it should be pointed out that the bounds are only strictly applicable when the sample set is infinitely large.

### 3.5.4 Adaptive Learning Networks

The use of ALN's as a classifier training technique has been much advocated by certain workers in automated N.D.E. An ALN functions as follows. An attempt is made to map each pattern point in each of the possible d(d-1)/2 pairs of feature axes onto a target value (see fig. 3.10) representing its class using a function of the form

$$y(x_i,x_j)=a_0+a_1x_i+a_2x_j+a_3x_ix_j+a_4x_i^2+a_5x_j^2 \qquad (3.33)$$

where $(x_i,x_j)$ are features within $\underset{\sim}{X}$, $\underset{\sim}{X}=(x_1,\ldots,x_{nfeatures})$

The coefficients $\{a_0,\ldots,a_5\}$ are optimised to minimise the mean squared error $E_{mse}$ between $y_{ij}$ and the target value t,

$$E_{x_i,x_j} = \frac{1}{n_{points\ total}} \sum_{k=1}^{n_{classes}} \sum_{all\ x \in class\ k} [y(x_i,x_j)-t]^2 \qquad (3.34)$$

over the entire point set. Rejection criteria based upon this measure applied to the learning and checking datasets eliminate unpromising feature pairs (hence greatly reducing the number of terms in the final polynomial decision function). Usually, the best 20 pairs according to the root mean square error criterion are then passed on to the next stage where the process is repeated, This time the points are defined in terms of the sets of $y_{ij}$ values, and another attempt is made to map the points onto the class target values. The process terminates when the mapping is good enough, or has ceased to improve (according to the values of the $E_{mse}$ yielded by successive generations).

The technique has been used with ultrasonic inspection of pipe welds, in order to determine the best amongst a very large number of ad-hoc spectral, cepstral and time domain features (**Shankar et al**,1978), most notably by **Mucciardi et al**(1975,1978). who has also applied the method to EC

NDT(**Mucciardi,** 1979), along with **Brown**(1979) who did so whilst inspecting fuel rod cladding.

An ALN was coded in Fortran, according to the method as specified by **Carter**(1980). This was the only paper which explained the technique in sufficient detail for it to be coded. Several problems and shortcomings with the technique became apparent. It was found that the PDP11 was too small to link in the large optimisation routine (NAG EO4FDF), the graphics library and the array storage required for workspace, even using overlaying. During each stage, the boundary optimisation algorithm is run for each possible pairing, and the best 20 pairings and the accompanying 20 set of target values for each of the pattern vector feature pairs must be retained at the end of each generation level. Because of this, both the storage and CPU time requirements were high. The experiment was transferred to the departmental GEC4070, which had virtual memory.

The test set was the 9 dimensional 3 class single frequency data set of **Macleod**(1982), described in Chapter 4. In none of the feature axis pairings did the optimisation routine converge upon a solution set. This was concluded to be because the form of the function cannot dichotomise a multiclass set, except under certain circumstances where the boundaries can be represented by one or more of the terms in eqn 3.35. The technique is limited in most cases to 2 class situations, in view of the functional form of the dichotomisation. In each of the 2-spaces, the boundary must be a conical section of some sort, i.e. either an ellipse, a straight line, a parabola or a hyperbola. If the form of the afunction of $x_i$ and $x_j$ given in eqn. 3.33 was to be restated more conventionally as a function of x and y which defines a surface of height z

above the x-y plane, then it might be more apparent that the function can at most describe a hill with an elliptical or hyperbolic plan form at a series of contours of constant z. Thus the discriminant function in the 2-space cannot separate more than two classes, unless the boundaries separating the classes are all of the same geometric form. Furthermore, the progression of the target values must reflect the relative positions of the classes.

Many of Mucciardi's group's publications are based upon work originally published upon the fitting of non linear discriminant functions to patterns derived from ultrasonic signal time series(**Shankar**,1978). Although there were only 34 data points in the training set, a 16th order incomplete polynomial in 7 features was fitted to these points using the ALN algorithm to generate and prune the feature pairs. Normally, a large data set must be available to train a highly non linear surface if overfitting is not to occur. It has been claimed that a much smaller set is required with an ALN, since only the weights for two features and their associated squared terms are trained at each stage. However, the final polynomial must still observe the limits placed upon the number of terms trainable for a given number of parameters. The number of terms being trained here was clearly much greater than the number of points merited, and the surface fitted was greatly overtrained upon the data, so the results upon the test set ought to have been poor. However, the defect set was only split into two classes and these were defects and non-defects. The ALN was not actually being asked to find a surface which could discriminate between defect types, and it is quite likely that most of the features varied considerably between defects and non-defects though they may not have varied nearly so much within those categories, and so the technique was not tested very severely. In view of the

method's shortcomings in dealing with more than two classes, had it been tested upon discrimination of 3 or more types of defect, it would probably not have performed so well.

### 3.6 References

1(a). **Macleod,J.E.S.,**"Automatic Inspection of Tubes Scanned By A Rotating Eddy Current Probe", Signal Processing, Vol. 1., No.5,pp445-450,1983.

1(b). **Macleod, J.E.S.,** "Feasability Study on Computer Sorting of Flaws in Steam Generator Tubes Scanned by a Rotating Eddy Current Scanner", Babcock Power Research Report No. 07/82/46, July 1982.

1(c). **Macleod, J.E.S.,** "Pattern Classification in the Automatic Inspection of Tubes Scanned by a Rotating Eddy Current Probe", Proc. Int. Conf. on Patt. Rec., Munich, Oct. 1982.

2. **Rosenfeld, A. and Kak, A.,** "Digital Picture Processing", Chapter 8, New York, Academic Press, 1976.

3. **Bodenstein, G. and Praetorius, H.M.,**"Feature Extraction from the Electroencephalogram by Adaptive Segmentation", Proc. IEEE, Vol.65,No.5,pp.642-652,May 1977.

4. **Mottl, V.V. and Muchnik, I.B.,**"Linguistic Analysis of Experimental Curves", Proc. IEEE, Vol.67,No.5,pp.714-736, May 1979.

5. **Logan, I.G. and Macleod, J.E.S.,** "An Application of Pattern Recognition Algorithms to the Automatic Inspection of Steel Strip Surfaces"

6. **Doctor, P.G. and Harrington, T.P.**, "Analysis of Eddy Current Data Using Pattern Recognition Methods", IEEE 5th Conf, on Patt. Rec., Miami, 1980.

7. **Udpa, S.S. and Lord, W.**,"A Fourier Descriptor Classification Scheme for Differential Probe Signals",Materials Evaluation/42, pp1136-1141, Aug. 1984.

8. **Satish, S.R.**,"Parametric signal Processing for Eddy Current NDT", Ph.D. thesis, Colorado State Univ.,1983.

9. **Pavlidis, T.**, "Algorithms for Shape Analysis of Contours and Waveforms", IEEE PAMI-2, No.4, pp. 301-312, July 1980.

10. **Cosgriff, R.L.**, "Identification of Shape", Ohio State University Res. Foundation, Columbus Rep. 820 - 11, ASTIA AD 254 792, Dec. 1960.

11. **Zahn, C.T. and Roskies, R.Z.**,"Fourier Descriptors for Plane Closed Curves",IEEE Trans. Comp., Vol C-21, No.3, pp. 269-281, March 1972.

12. **Granlund, G.H.**,"Fourier Preprocessing for Hand Print Character Recognition", IEEE Trans. Comp., pp. 195-201, Feb. 1972.

13. **Persoon, E. and Fu, K.S.**, "Shape Discrimination Using Fourier Descriptors", IEEE Trans Sys. Man. Cybern., Vol. SMC-7, No.3, pp171-179, March 1977.

14. **Bartlett, M.S.**, "The Spectral Analysis of Point Process Functions",pp264-296,1963,

15. **Burg**,"Maximum Entropy Spectral Analysis",reproduced in "Modern Spectrum Analysis",IEEE Press, 1978.

16. **Tjostheim, D.**,"Recognition of Waveforms Using Autoregressive Feature Extraction",IEEE Trans.Comp, pp268-270, March 1977.

16.**Akaike**,"A New Means of Determining the Order of an AR Process", reproduced in "Modern Spectrum Analysis",IEEE Press, 1978.

17.**Ulrich and Bishop**, "A Fortran Implementation of the Maximum Entropy Method" reproduced in "Modern Spectrum Analysis",IEEE Press, 1978.

18. **Fukunaga, K. and Narendra, P.M.**,"A Branch and Bound Algorithm for Feature Subset Selection",IEEE Trans. Comp., Vol. C-26, No. 9, pp917-922, Sept. 1977.

19. **Roberts, S.J.**,"A Branch and Bound Algorithm for Determining the Optimal Feature Subset of Given Size",J. Royal Stat. Soc.,pp236-241,1984.

20. **Fukunaga, K. and Mantock, J.M.**,"Nonparametric Discriminant Analysis", IEEE PAMI-5, No.6, pp671-678,Nov.,1983.

21. **Kittler, J., and Young, P.C.**, "A new approach to feature selection based on the Karhunen-Loeve expansion", Pattern Recognition, Vol. 5, pp. 335-352, 1973.

22. **Devijver, P.A. and Kittler, J.**, "Pattern Recognition: A Statistical Approach", Prentice-Hall,1982.

23. **Duda, R.O. and Fossum, H.**, "Pattern Classification by Iteratively Determined Linear and Piecewise Linear Discriminant Functions", IEEE Trans. Comput., vol. EC-15, no.2, pp 220-232, April, 1966.

24. **Logan, I.G.**, "The Application of Pattern Recognition Algorithms to

the Automatic Inspection of Steel Strip Surfaces", Ph.D. Thesis, Univ. of Glasgow, 1973.

25. **Shankar,R.,Mucciardi,A.N.,Lawrie,W.E.,Stein,R.N.,**"Development of Adaptive Learning Networks for Pipe Inspection",EPRI-NP- 688,Mar.,1978.

26. **Mucciardi, A.N., Shankar, R. and Buckley, M.J.,**"Applications of Adaptive Learning Networks to NDE Technology",Proc. IEEE Nat. Aerospace & electronics Conference,pp460-469,1975.

27. **Mucciardi, A.N. and Dau, G.J.,** "Development of Adaptive Learning Network for Eddy Current Signal Analysis". "Nondestructive Evaluation Program: Progress in 1979", EPRI NP- 1234-SR, Special Report, December, 1979.

28.**Brown, R.L.,**"Investigating the Computer analysis of Current NDT Data.", HEDL SA 1721, Hanford Engineering Development Lab., Richland, WA.,1979.

29. **Carter, P.,**"Comment on Adaptive Learning Networks in Pattern Recognition",Harwell Report AERE-R 10021, Dec.,1980.

30. Luk, A.,"Some New Results in Nearest Neighbour Classification and Lung Sound Analysis", Ph.D. thesis, University of Glasgow, 1987.

31. **Logan, I, and Macleod, J.E.S.,** "An Application of Pattern Recognition Algorithms to the Automatic Inspection of Steel Strip Surface",Proc 2nd international conference on Pattern Recognition, Copenhagen, 1974, pp286-290.

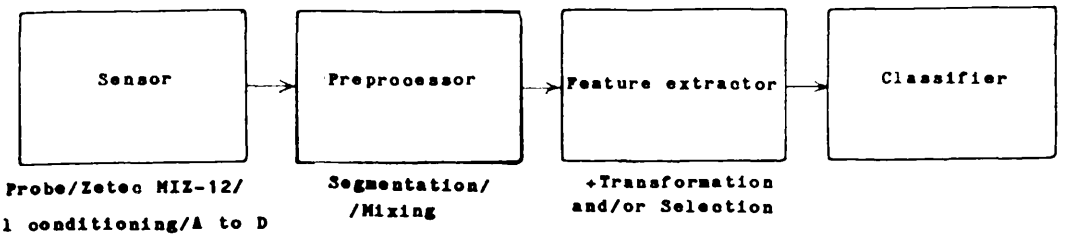



fig 3.1 A production or operational pattern recognition system

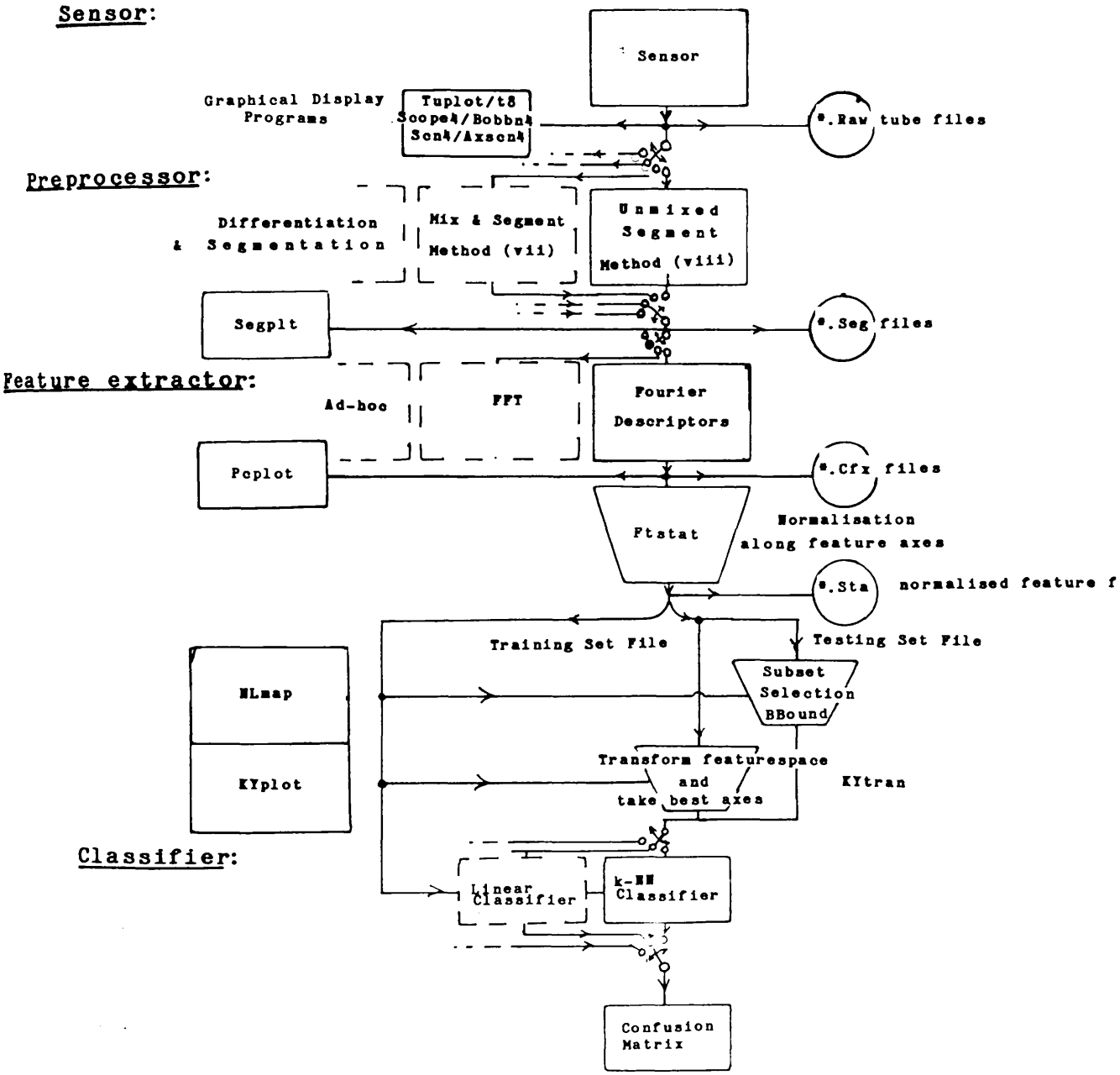



fig 3.2 An experimental system. This is a simplified representation of the structure of the pattern recognition suite developed during the work done for Chapters 3 and 5. The switches shown were in effect set at link time, when one particular module at each level was included in preference to others.

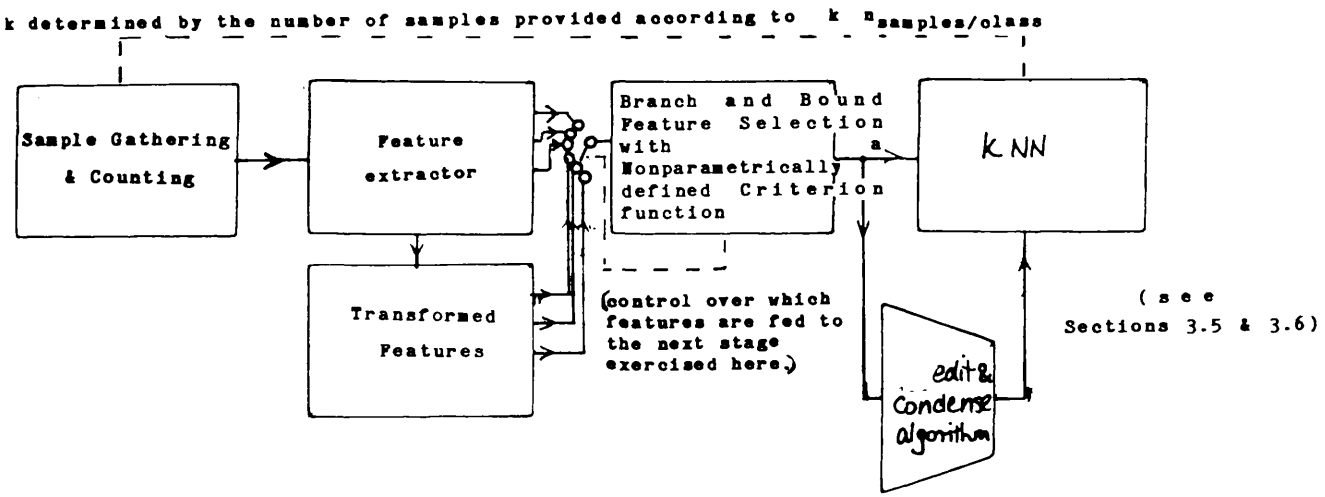



**fig 3.3** An Adaptive system, which tailors the featureset fed to the classifier according upon the results achieved with various features, and adjusts the number of points k used in the nearest neighbour classifier.

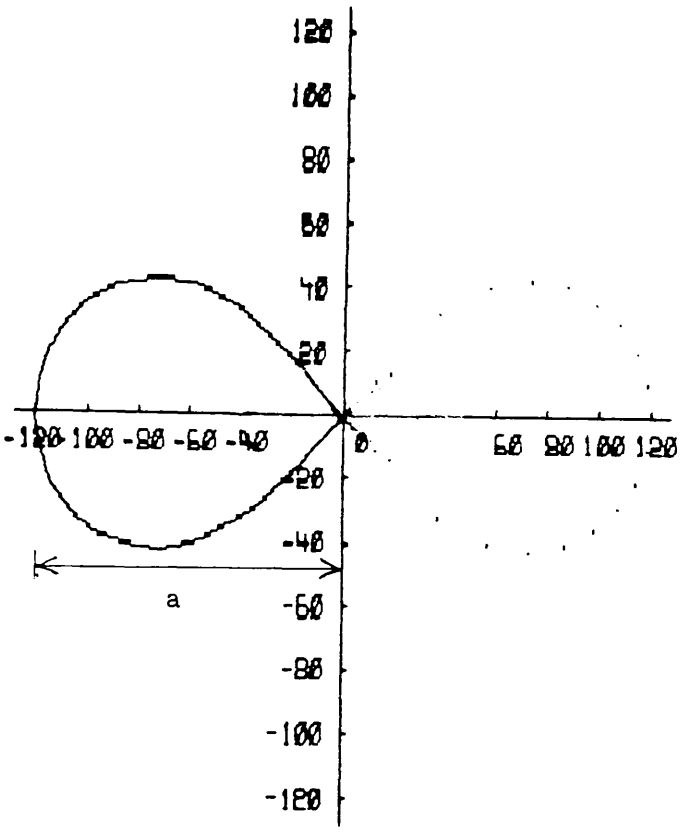



**fig 3.4** A 38-point Lemniscate of Bernoulli, with the samples joined only in the left hand plane. The non-constant arc length sample spacing is clearly visible in the right hand plane. The figure was generated by substitution of $\theta$ over the ranges [-45,45] & [225,135] in increments of a constant size of $5^{o}$ into the right hand side of

$$r^2 = a^2 \cos 2\theta,$$

where a is the major axis half length as shown.

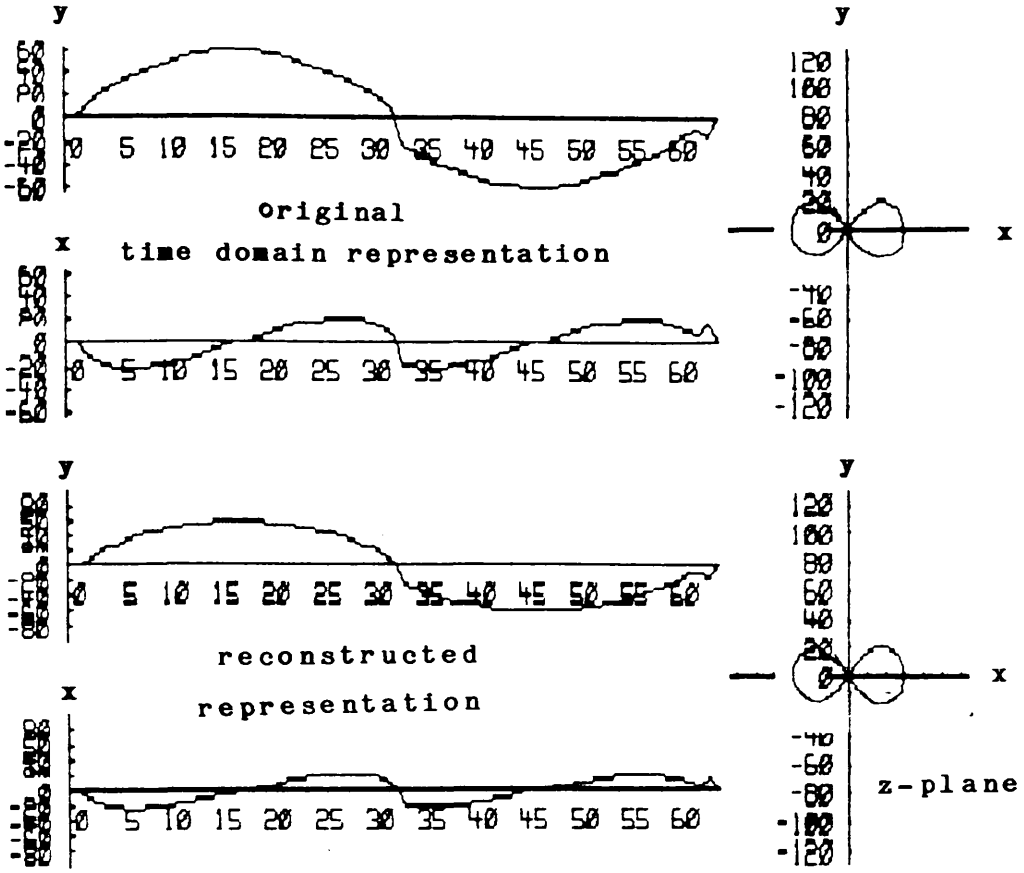



**fig 3.**5 Both the times series and z-plane forms of the original and reconstructed versions of a 64 point lemniscate, which was parameterised with a 64 point FFT. Note that there was no truncation of the f-domain ordinate sequence before the inverse transform in order to get such complete reconstruction.

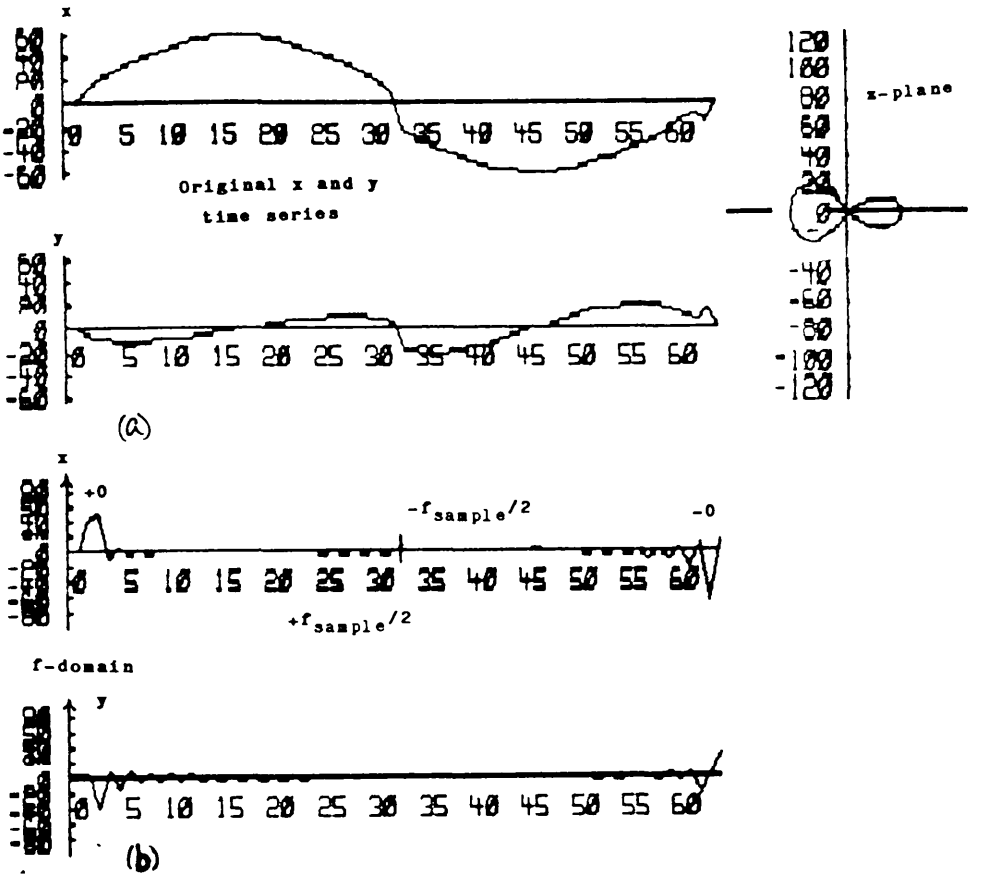



**fig. 3.6**(a) A slightly customised lemniscate which was altered to be assymetrical about the y-axis, to test the parameterisation schemes harder than the original lemnsicate.
(b) the frequency domain for (a), which tends to suggest that the figure should reconstruct with the lowest frequency (outer) 3 or 4 terms on either side, especially since the vertical scales are amplitude rather than power

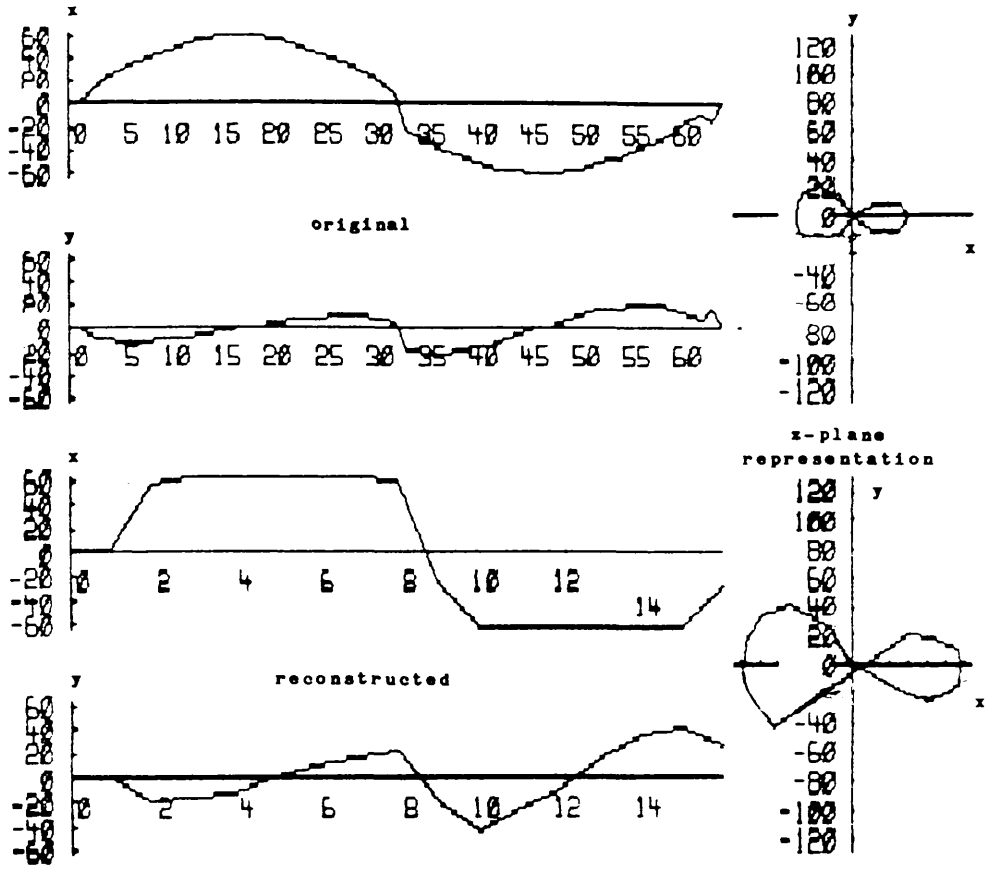



**fig 3.7** An attempt to reconstruct the lemnsicate of Bernoulli from the 16 lowest frequency ordinates: 8 on either side. The non-closure of the figure became even worse with further truncation, indicating that this number of ordinate pairs was not capable of describing the bulk shape of even such a simple and evenly rounded shape such as this. Results with sharper lobes such as the Joukowski aerofoil which bore more resemplance to the eddy current loci of interest were similarly dissapointing.

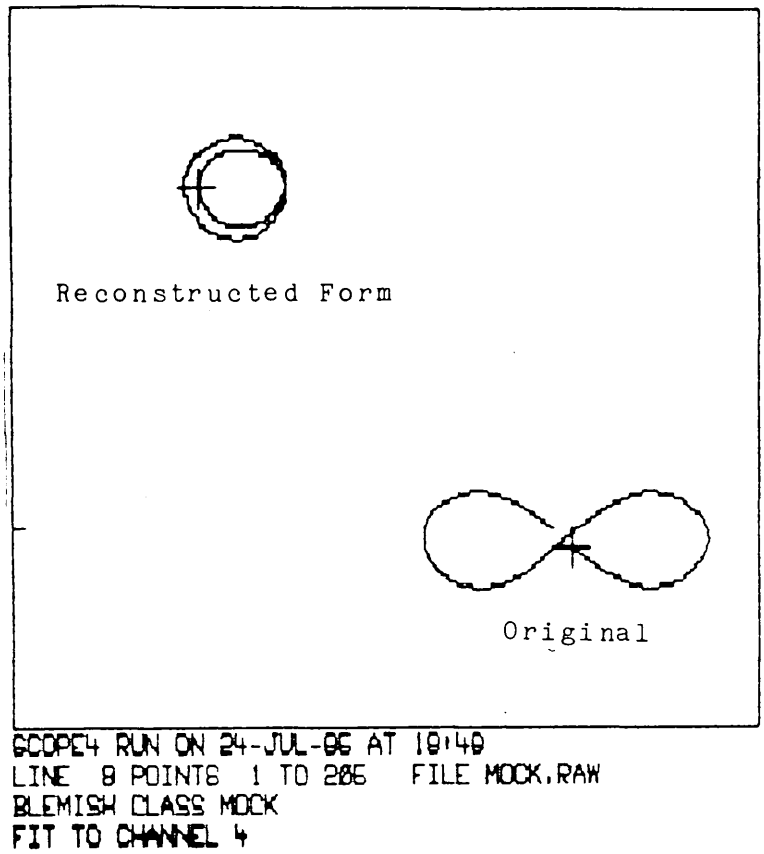



**fig 3.8** Fourier Descriptors failed to reconstruct even a simple 38 point lemniscate in a 64 point zero padded input series, with no frequency domain truncation. The use of a zero padded 38 point lemnsicate was intentional, in order to check that the FD's were unaltered compared to those for a 64 point lemniscate in a 64 point long input series.

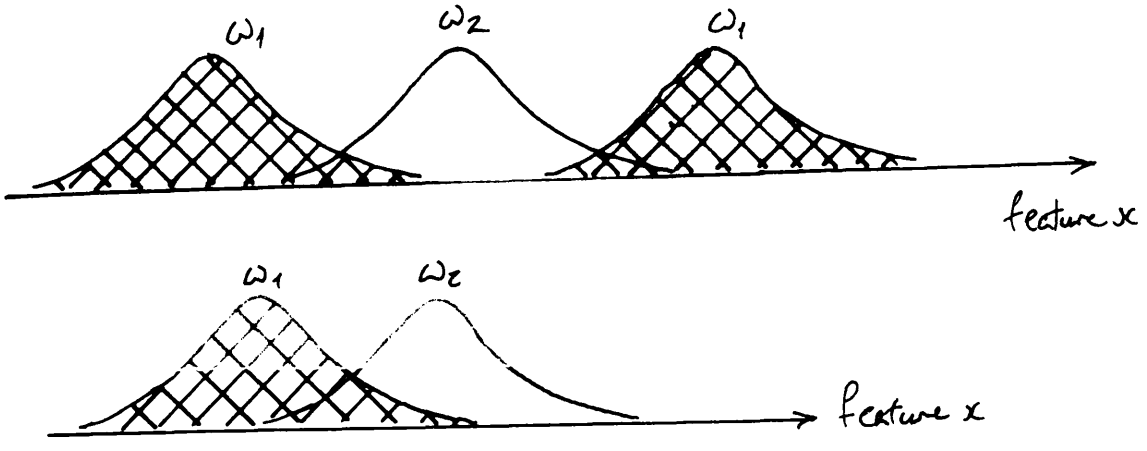



Fig. 3.9(a) The form of the distribution used to test the new nonparametrically-defined criterion function, projected into one dimension (see Section 3.4.3). The function had a higher value for this distribution than for one in which the classes were unimodal and normally distributed with coincident class 1 and 2 means, despite the fact that the statistically defined class means were apperently also coincident in this distribution.

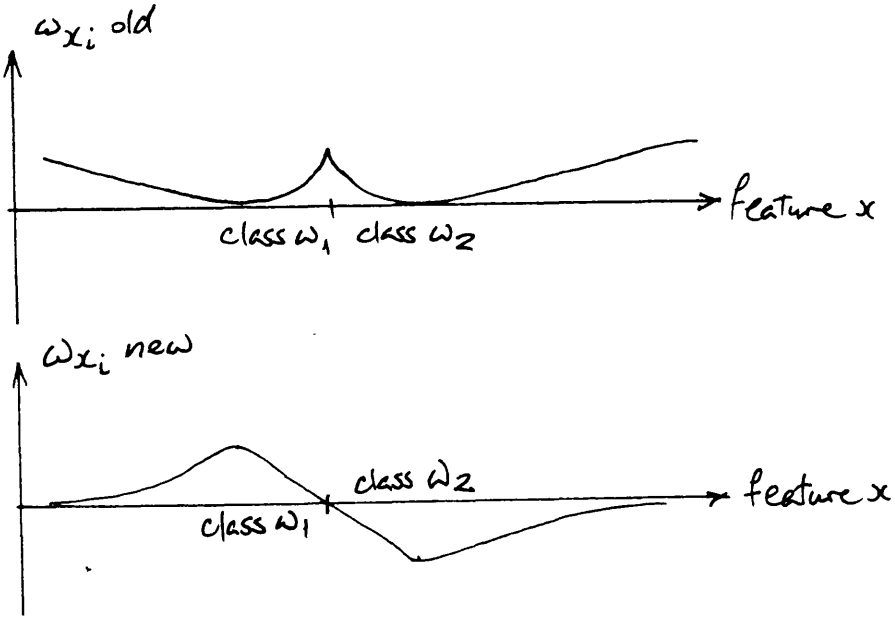



Fig. 3.9(b) The behaviour of the weighting functions given in eqns. 3.28 and 3.29 as a test point $X_i$ is moved along a line joining two local class means. The point belongs to the class to the left of the decision boundary.

The modification to Fukunaga nd Mantock's weighting function shown in graph (b) actively discriminates against points which misclassifiy by a large margin. The original form in graph (a) emphasises points close to the boundary, regardless of whether they are misclassified or not, which tends to make the total criterion function unduly biased towards the points which are most troublesome to dichotomise correctly.

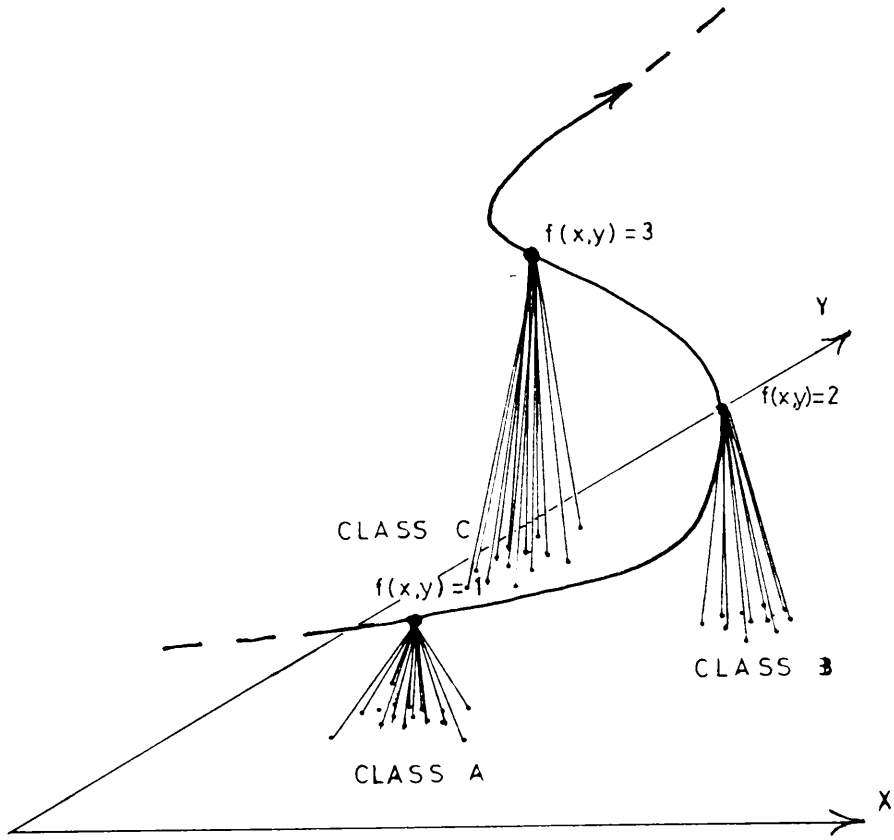



Fig. 3.10 Points in a 2 dimensional featurespace defined in terms of features x and y, being mapped by one stage of an adaptive learning network onto target values of the function f(x,y) which can be represented as heights upon the z-axis (see eqns. 3.33 and 3.34).

# 4 Chapter 4: Further Work Upon the 316 Tubing

## 4.1 Introduction

Initially, the major intention of this project was to extend the use of pattern recognition on single frequency eddy current loci to the processing of multifrequency loci. However, after the initial feasibility study, there still remained some "unknowns" such as the upper limit of performance possible in conjunction with the ID defect set in the 316 tubing. This has been determined both by trial and error and systematically, firstly by attempting to make improvements in the misclassification rate by (i) using different sets of features and classifiers;
(ii) assessing the extent of the marginal improvement obtained from the addition of further features to the featureset (by measuring their correlation coefficients with the existing ones),
and (iii) estimating the maximum possible correct classification rate, by estimating the rate that would be achieved with a Bayesian classifier (see Chapter 3).

Single frequency testing is still of considerable interest to the NDT community regardless of the trends towards multifrequency testing discussed in the chapter 2. In routine inspection of tubing, single frequency testing is still the norm, so although the results in this chapter are specific to defects in finned 316 stainless steel tubing, the segmentation technique and the feature set that have been developed are more generally applicable.

### 4.1.1 Summary of the Feasibility Study

At the end of Macleod's feasibility study, all of the available twenty one sections of tubing had been logged. These contained a total of 38 defects with approximately equal numbers of each of the three types, which were axial cracks (ZCK), intergranular stress corrosion cracking (IGSCC), and general corrosion (GCO). These were logged onto tape using a conventional amplitude modulation-type tape recorder (recording a d.c. signal component was unnecessary since a differential probe was being used). The in-phase and quadrature signals, and a once-per- revolution, position-synchronising pulse were recorded. The signals were then digitised and stored in RAM in a small 8-bit S100-based microcomputer system (see the mark 1 system in the appendix), and when the user indicated at the terminal that the logging should be terminated, dumped onto floppy disk.

The datafiles representing the tubes were then transferred to a larger computer, the PDP11/45 in the Department of Electrical Engineering at the University of Glasgow. Some preprocessing of the files had to be done to produce scans of a uniform length because the rotation synchronising pulse was occasionally missed by the logging program, and also because the length of the drive shaft between probe and drive motor caused some shuddering of the probe.

One scan through a defect was treated as a raw pattern and features were extracted only from scans containing defects. The feature vectors were normalised such that the overall feature distributions had a zero mean and unit variance along all of the feature axes. A feature space rotation was then performed and the best 2-space plotted in order to demonstate the extent

of class separability. Finally, a two class linear classifier using the 2-class training method of **Duda and Fossum**(1966), was used in order to discriminate between the 3 defect classes. This was done sequentially, first by separating the GCO defects from the union of the ZCK and IGSCC classes, and then by separating those two classes.

### 4.1.2 Initial Aims

**Macleod's** suggestions for further work (**Macleod**, 1982) were treated as the initial aims for this part of the work and may be summarised as follows:

(i) an enlargement of the defect set by manufacture of accurately sized notches and cracks.

(ii) more sophisticated segmentation of defect signals from the background noise, (rather than by hand or by the methods developed previously (**Macleod**, 1982)) including the use of linear predictive coding to detect the ends of cracks;

(iii) to assess the applicability of image processing or recognition

(iv) apply digital filtering to remove noise within the defect signal;

(v) attempt probe transfer function modelling with a view to matched filtering as a preprocessing stage;

(vi) experiment with further previously untried features;

(vii) further classification methods.

All of these topics have been addressed with the exception of linear predictive coding and the results obtained for each of them are now given in detail.

### 4.2 The Defect Set

Macleod's logged defect set consisted of an approximately equal mix of three defect types. The training set file (referred to as CP7.dat, see fig. 3.2) contained 50 scans through each of the 3 types, and the testing set (CP9) contained all of the remaining scans for each type i.e. 119 axial crack scans, 100 intergranular crack scans and 94 general corrosion scans.

#### 4.2.1 Manufacture of New Defects

Considerable effort was expended in an attempt to manufacture genuine cracks in 316 stainless steel 25mm. bar, by welding a fillet onto the bar, notching it, and cyclicly bending the bar with a three point loader. However the remnant signal from the fillet even after it was milled off swamped the crack signal, and because of problems with logging the cracks, they could not be added to the defect set for this study. Even if the logging problems had been solved, the production of cracks would have been a very time consuming and expensive process.

The production of intergranular attack (IGA) was also attempted, but controlling the extent of the IGA, or limiting it to anything other than the whole of the inside of a tube was beyond the capabilities of Babcock's Research Centre's Chemistry Department at the time. The tubes which were

produced with IGA were either too brittle for inspection with a rotating probe, or the inner diameter had decreased because of some swelling of the metal due to the attack of the surface and it was no longer possible to get the probe into them.

It would have been desirable to have been able to quantify the difference between the signal from for instance a spark notch and a crack of the same depth, but since this would have meant sacrificing defects for sectioning, it was not possible. Consequently, the degree of similarity between the defects in the defect set (which have all occured during manufacture) and those which occur during service must now be considered.

### 4.2.2 Manufactured vs. In-Service Defects

Very little has been published upon this subject, probably because usually the stress is upon the detection of defects, and also because manufacturers do not like admitting that such defects occur. However, some doubt surrounds the degree of similarity between manufacturing defects and those which occur under service conditions. For instance IGSCC, which occurs during the manufacturing process, is thought to be due to the inclusion of some foreign matter, perhaps grease, during the hot extrusion process. On contact with the hot metal, the grease gasifies and forces paths into the metal surface. Removal of the grease by acid washing in the later stages of manufacture leaves the metal surface laced with small fissures. IGSCC which occurs during the operational life of the tubing is thought to be due to the combination of high pressures and temperatures present in a heat exchanger, and also the vibration due to the turbulent flow of coolant over the tube exterior surfaces.

It is also worthy of notice that anything within the primary or secondary cooling circuits, i.e., the insides and outsides of the steam generator tubing, acquires a coating of magnetite during service. Magnetite is a water borne corrosion by-product which occurs as a consequence of the corrosion of steel in other parts of the circuit. This coating was not present on the tubes inspected, nor would it have been easy to generate synthetically. Were it to be thick enough, it would greatly reduce the penetration of the eddy currents into the tube wall because it is magnetic. However, fortunately the inspection of steam generator tubing in service is not hampered much by magnetite deposits, because the coating tends to be quite thin, and pressurised water washing of the tube interiors prior to inspection removes the more troublesome internal layers.

### 4.3 Segmentation

During his study, **Macleod** experimented with 3 simple methods of segmentation (see **Macleod** (1982), Section 4, Figs. 2, 3a and 3b). These were:

**Method (i)** an amplitude threshold applied to the y channel magnitude alone.

**Method (ii)** an extension of method (i) which allowed the amplitude threshold level to exponentially decay to include the tail of the defect signal.

**Method (iii)** the threshold criterion of method (i), combined with a logical test on the adjacent points in the previous and current scans to find out if they had passed the threshold test and were part of the same

defect.

These methods worked, but method (ii) was insensitive to the start of the signal due to a defect, whilst method (iii) required additional storage to be provided for the previous scan to be held in memory. A new method was devised, and this has the advantages of both of Macleod's methods (ii) and (iii), i.e. speed, and equal sensitivity to the signal both before and after the defect peak, but without the disadvantage of the requirement to store two scans simultaneously.

There would in fact have been enough memory available to have stored two single frequency scans simultaneously, but it was envisaged at the time that this work was done that the single frequency and multifrequency segmentation would be done by similar methods, and since it was not feasible to store two scans worth of 8 channels at 400 points/scan/channel with the memory space available on the PDP11/45, further development of method (iii) was not favoured.

The algorithm utilises a few properties peculiar to these particular types of defect signals which became apparent as a result of some work at the Machine Intelligence Research Unit, at Edinburgh University to investigate the feasibility of image processing (see fig. 4.1).

From inspection of a selection of tube signal plots, it became apparent that the background noise was of the same amplitude in all files i.e. about 4 adc units (although this tended to be disguised by the different scaling used to accommodate various defect signal amplitudes in different plots). In any area where fluctuations became larger than this, if the x and y signals deflected

in the same sense from the datum, the "defect" was observed to be always general surface corrosion. It was undesirable that this be passed on to the feature extraction stage because it is harmless i.e. not associated with tube rupturing.

In any area in which an axial crack or IGSCC was present , the x and y signals always diverged in opposite directions. These defect types always had a width of at least 20 points. Furthermore the peaks and troughs in the characteristic response (see figs. 4.7 & 8) were never separated by more than 20 points and the defect signal always decayed back into the background within 20 points on either side of the two outer minima adjacent to the peak in the y-signal. These "rules of thumb" were combined into a segmentation algorithm as follows:

Let a binary valued list, kset (400), denote the presence or non-presence of a defect at each point in a scan of length 400 points (representing 1 complete revolution of the probe around the tube), where all elements in the list are initialised as false, indicating the non-presence of a defect at all points in the scan under consideration. Then

(i) for each point i, i=1,..,400, set kset(i) =true if $|x(i)| > 4$ and $|y(i)| > 4$ and the polarity of x and y are opposite.

(ii) for each point (i), if there is another point k where kset(k)=true and $|(k - i)| > 20$, then set kset = true for all intervening points.

(iii) delete all sequences of points already set true where the group is less than 20 points in length.

(iv) set a margin of 20 points true on either side of each remaining

sequence of points.

The list kset(i), i = 1,...,400 now represents a mask, which is set true for all sequences of points corresponding to a defect.

Part (iii) removes indications which are too short to be normal cracks or IGSCC. The value of 20 strikes a balance between accepting noise and rejecting the tail-ends of cracks. This also prevents feature extractor stage from being presented with the low amplitude scans at the ends of defects, which tend to be the scans which are misclassified, and thus overly pessimistic misclassification rates are prevented. Somewhat obviously the axial lengths of defects found by this method were a little conservative. From visual inspection of figs. 4.2 and 4.3 it would seem that 1 or 2 scans at each end of the blemishes discernibly contained a defect signal, and that these scans were not being picked up by the segmentor. The effect was much less noticeable with larger amplitude defects. The tubes shown in figs. 4.2 and 4.3 are worst case examples because they both contain particularly low amplitude defects. The algorithm had no difficulty in resolving the ends of more severe defects.

The segmentor program could have been modified to pick up these tail end indications by including an extra two scans at each end of each defect and using the mask in the kset array as it was for the first and last segment and maintaining a two scan buffer between the segmentor and the feature extractor. However other means have been found to detect the ends of a defect using cross correlation with an idealised crack signal (see Section 4.4.2).

The detection rate for the segmentation algorithm was 79.7% for the ZCK class, and 98% for the IGSCC class in the training set, (representing 47 scans from a possible 59, and 57 from 58 for each class respectively) and 60.3% (73 from 121) for the ZCK class and 87% (87 from 100) for the IGSCC class in the testing set. This was an overall scan detection rate of 66.6% for axial cracks, and 91.1% for intergranular cracking, over the entire data set. There were no instances of noise or surface corrosion being detected as defects, except in one experiment when the amplitude threshold was lowered from 4.0 adc units to 2.5 out of curiosity.

The true level of noise and corrosion was exacerbated by the differential nature of the probe because it is particularly sensitive to any sharp transition along its path whether it was a crack or surface roughness.

Using the technique developed here, it was found possible to segment the signal due to a defect without recourse to more involved methods requiring the storage of previous scans. The only effect upon the signal in a production system would be the incorporation of a time delay equal to the segment association distance (the minimum width of any defect: in this case 20 samples with 400 sample/scan). Use of the algorithm typically cut the amount of disk space necessary for the storage of the defect set by 80% for tube files containing IGSCC and 90% for axial cracks.

### 4.4 Cross Correlation and Digital Filtering

### 4.4.1 Digital Filtering

The objective in introducing filtering was to reduce the high frequency noise content of the signal whilst leaving the defect signal unaltered, in an attempt to reduce the variability within the classes. Here, this was more necessary than in the case of the subsequent multifrequency tube loggings, since the differential probe tended to exaggerate any small and rapid (i.e. high frequency) signal changes due to surface roughness.

One dimensional filtering (i.e. of the time series rather than of the whole tube signal plot) has been used, in order to keep the whole system as simple as possible, with a view towards future practical implementation. The alternative, filtering along the longitudinal axis of the tube in order to eliminate uncorrelated noise between adjacent scans, would have required a lot of storage and processing time. In any case, since the probe was only sensitive to the component of the magnetic field variations along the scanned path, the axial defect characteristics were masked. Thus it seemed that there was little practical advantage to be gained in reordering the scanning pattern into a lengthwise one via a large buffer.

The first priority in the choice of filter was that it have linear phase change with frequency to prevent radical changes in the shape of the defect signals and also preserve the x/y phase relationship which is so central to the discrimination of GCO from cracking defects. Although nonrecusive filters exhibit good linear phase responses, they need many more previous input terms to achieve a sharp cut off than do comparable recursive ones and they have objectionably poor stop band characteristics due to Gibbs oscillations. Recursive filters achieve a much improved stop band response with relatively

few previous input and output terms, so a recursive filter was chosen. Usually these advantages just mentioned are obtained in exchange for a non linear phase response however a recursive filter whose design is based on the analog (Bessel) approximation shows the same linear (phase,frequency) relationship characteristic exhibited by the analogue version, so a Bessel filter was the one used.

The 2nd order (analog Bessel transfer function (giving 40 dB/decade roll off above 1 rad/sec) is

$$H(s) = \frac{3}{s^2+3s+3} \qquad (4.1)$$

Applying the bilinear tranformation

$$s = \frac{\beta\,(z-1)}{(z+1)} \qquad \text{where } \beta = \frac{2}{(w_{c/o}T)} \qquad (4.2)$$

changes this low pass filter in the continuous (s) domain into one in the sampled (z) domain with cut off frequency $w_{c/o}$ rad/sec and sampling interval T seconds, and so H(s) becomes

$$H(z) = \frac{1}{(\beta^2+3\beta+3)z^2+(6-2\beta^2)z+(\beta^2-3\beta+3)} \qquad (4.3)$$

which stated alternately in terms of $Y_k$ and $X_k$, the kth filter outputs and inputs is

$$Y_k = \frac{X_k+2X_{k-1}+X_{k-2}-(6-2\beta^2)Y_{k-1}-(\beta^2-3\beta+3)Y_{k-2}}{(\beta^2+3\beta+3)} \qquad (4.4)$$

The bilinear transformation does not preserve the linear phase shift /frequency characteristic sought with the use of a Bessel filter, but the nonlinearity only becomes significant as frequency approaches the Nyquist value. Hence in this low pass filter, the distortion of the linear phase shift/frequency relationship will be largely confined to the stop band, and the phase response will remain substantially linear throughout the frequency range of interest.

The cut off frequency was arrived at by trial and error, by inspection of the resulting signal, to ensure maximum high frequency noise elimination without allowing the shape of the defect signals to be distorted.

With a low pass cut off frequency corresponding to a period of 2 mm on the tube surface very little change was visible in the tube plots (see figure 4.4). Decreasing the cut off frequency to one corresponding to a $T_{c/o}$ of 5.2 mm (see figs. 4.4(b) and 4.5(b)) improved the signal by removing all visible noise, whilst leaving the defects and the surface corrosion signals' form unchanged. At $T_{c/o}$=7.8 mm, the signal, rather than merely the noise began to be affected (see figs 4.4(c) and 4.5(c)), and below this the signal was extinguished. In all of these plots the spatial sampling interval was 0.1mm (measured along the probe path). At the best compromise between surface noise attenuation and defect signal shape preservation (i.e. with $T_{c/o}$ = 5.2mm: compare figs.4.4.(a) and 4.4.(b), and 4.5(a) and (b)), surface corrosion was still obvious, so the filter order was increased by cascading two 2nd order Bessel filters of the type described above, which caused a rather oscillatory output in the vicinity of any defect.

The 4th order filter did improve the rejection of the high frequency noise over

that of the 2nd order but at the cost of markedly altering the defect signal, so the 2nd order Bessel filter with with $T_{c/o}$ = 5.2mm was chosen for use in tests upon a featureset with defect signals preprocessed by digital filter. The results of this are given in Table 4.2. Unfortunately, the results were actually poorer than the misclassification results from the unfiltered comparable experiment (**Macleod**,1982). This implies either that the high frequency content of the defect signal does carry discriminatory information, or that even though visible distortion of the defect signals was only just discernable, it was sufficient to degrade the dixrimatory information contained in the signals.

It would have been desirable to have been able to experiment further with much higher order filters with a much sharper cut-off but the 4th order proved difficult to produce in a form with an output not prone to instability, and so since the basic aim was improve classifier performance and since this had not been achieved, no further experimentation was carried out upon this.

Cross Correlation of Scans with Prototype Defects

An approach to noise reduction that is well known in communication theory is that of cross correlation of an input signal with the signal to be detected. Correlation can normally only be used under circumstances where the form of the desired input signal is known in advance but it was of interest in this project because the number of different signals to be detected was quite small, and all but the GCO class could be represented by idealised waveforms.

The correlation of two time series is expressed as:

$$R_{xy}(k)=\sum_{i=1}^{n} x(k)y(i),\ k=1,...,n \qquad (4.4)$$

where x(i) and y(i), i=1,...,n are the candidate scan and the prototype scan respectively; in this application n=400 points. The operation is usually quoted in the form of eqn. 4.4, but here what was required is more correctly called circular correlation, where a prototype scan is correlated with 1 complete scan at a time, rather than with a signal which continuously streams out of the probe. In circular correlation the series to be multiplied are held in two circular arrays, one containing the prototype scan, and the other, the latest probe scan. The k values in the correlation series $R_{xy}(k)$ are formed by stepping one array past the other by one element at a time, and for each step (or unit increment in k) repeating the summation of products operation. Circular correlation was preferred here over the more common linear correlation because of memory requirement considerations. Correlation was most easily accomplished using real number arithmetic because some normalisation of the output arrays was required. The use of circular correlation meant that only one scan had to be held in memory at once (rather than two with normal correlation because of the format of the data storage files), and this was critical to the feasability of the correlation task because of the the small amount of memory leftover for data when this program was loaded.

In order to generate the idealised forms of the signals due to defects or "prototype" scans, typical ZCK and IGSCC scans were hand-picked, and then an ad-hoc windowing function was applied to them (a 4th order Blackman-Harris window with a central plateau inserted to allow only the defect signal to pass unattenuated), to eliminate all but the signal segment deemed to be the defect signal. Low pass filtering and then normalisation (see section 4.4.3)

were then applied to produce the prototype for correlation. This rather ad hoc method worked better than the more obvious box-car averaging method of producing an averaged typical prototype from a number of superimposed scans. Box car averaging was tried experimentally, but was found unworkable due to the difficulty experienced in aligning the centres of defects taken from different scans.

As can be seen from figs. 4.7 and 4.8, the prototype defect signals were almost symmetrical about the central peak, which suggests that the results of correlation and convolution would be very similar. Accordingly, only one method of the two correlation methods, was pursued with to a conclusion. In addition, the definitions of the correlation operation are rather similar to those of convolution in both the time and frequency domains (**Brigham**,1974):

<u>correlation</u> <u>convolution</u>

$$Rxy(k)=\sum_{i=0}^{N-1} x(i)y(k-i) \quad (4.5) \qquad Cxy(k)=\sum_{i=0}^{N-1} x(i)y(k+i),\ k=0,..,N-1 \quad (4.7)$$

$$R_{xy}(t) = F^{-1}[X^{*}(f)Y(f)] \quad (4.6) \qquad C_{xy}(t) = F^{-1}[X(f)Y(f)] \quad (4.8)$$

From the time domain definitions, it can be seen that they differ only in that one function is time reversed before the sliding summation of products operation takes place.

Initially the correlation was performed in the spectral domain using eqn. 4.6, in view of the speed advantages expected with an FFT. However, this method would not work (despite a lengthy series of experiments to determine if the cause was related to the windowing of the data series, or the position

of the zero padding within the transform). Although a peak would occur when a defect was present, the peak always occurred close to the middle of the correlated time series and the positional information was lost. Zero padding the 400 point time series to fit a 512 point transform seemed to be the cause of the problem. The alternative approach, compressing the 400 point series to fit a 256 point transform was not possible because the resultant program approached the PDP11's space limitations to the extent that there was no core space left for graphical output or even for the arrays holding the correlated signals to be plotted of interest to be written to an output file for subsequent plotting.

Preservation of the peak positional information within scans could only be achieved by resorting to the simple "brute force" method of eqn. 4.5. Although this was slow at 45 sec/scan compared with 1 second/scan for an FFT-based method, it worked well and also avoided the uncertainties of whether or not to smooth, window, zero pad or compress the time series (or otherwise corrupt the data in order to use a transform of an unsuitable length).

### 4.4.2 Classification by Correlation

The key factor in getting the technique to work as a classifier was the method of scaling the result. Scaling of the correlation series is a topic seldom discussed in either communication theory text books or applications papers using correlation, probably because it seems trivial. Perhaps this is because most authors are interested in noise reduction, or in signal detection or image matching, rather than classification upon the basis of the comparative sizes of the correlation peaks from several prototypes. The scaling scheme used

was to prescale the prototype time series to have unit energy in both the y-signal and x-signals individually:

$$\sum_{k=1}^{nsamples/scan} y_{prototype}(k)=1 \quad \text{and} \quad \sum_{k=1}^{nsamples/scan} x_{prototype}(k)=1 \qquad (4.6)$$

whilst leaving the test scan and the correlation series itself unmodified. Normalisation with respect to the product of the two signal segment energies was also tried, but tended to produce spurious peaks the correlation series for scans containing no defect, due to division by zero (representing an almost zero signal energy). In other words, where there was a noisy line, devoid of a defect signal but containing a few small peaks, these would be artificially emphasised by the effect of the normalisation procedure which acted to so as to produce unit energy in all correlated scans. This effect is unavoidable whilst the normalisation coefficient is calculated individually for each scan, so the scheme finally chosen was that of normalisation of the prototype alone.

The results shown in figs. 4.9-14 show graphically how the height of the peaks in the two correlated y-signal plots for each tube correspond to the type of defect present. By classifying the defect as that of the prototype which produced the strongest correlation signal peak a 90% correct classfication rate was achieved. The correlated x-amplitude was always smaller than that of the corresponding y-channel because the eddy current equipment had been set up to maximise the signal in the y-direction by rotation of the phase control (at the expense of the the x-axis signal size).

For segmentation purposes the correlated y-signal obtained when the prototype was an axial crack was best because it was the most accurate indicator of defect position, and it would have been used for segmentation prior to the

classifier in the experiments in section 4.7 had it not been that the segmentation algorithm of Section 4.3 was much quicker and also blocked the GCO class signals from being passed to the classifier.

The extent of defects, particularly the detection of their ends, which are normally submerged in noise, is obvious in the y- signal correlations. Thus the correlation performs better than the segmentation algorithm in section 4.3 for detecting the ends of cracks. For correlation to be used in a system operating in real time, one of two methods would have to be implemented. Either, the scan lengths would have to be made to fit an FFT transform by altering the sampling rate, or a fast hardware time series correlator would have to be devised. The latter could be arranged to place no restrictions on the number of sample points per scan, as long as all of the scans matched the prototypes in length. In both cases, preservation of defect position information would be possible whilst still retaining processing speed.

It has thus been shown that correlation will highlight the position and axial extent of a defect, performing best when the prototype is an axial crack. The discriminatory capability of correlation is discussed in Section 4.8, where a practical implementation of a correlation-based classifier is suggested.

### 4.4.3 Matched Filtering

It is a standard result in communication theory (**Carlson**,1975) that if a pulse of known shape x(t), is contaminated by additive noise with spectral density $G_n(f)$, then the optimal filter for detection of that pulse is one with transfer function

$$H_{opt}(f)=\frac{X^{*}(f)e^{-jwt}}{G_{n}(f)} \qquad (4.9)$$

where t is the time lag after reception of the pulse at which the filter output peaks, w is the angular frequency in radians/sec., X(f) is the complex amplitude frequency spectrum of the pulse and * denotes complex conjugation. On closer inspection of eqn. 4.9, if the exponential term due to the time lag is neglected, since it is merely a phase rotation, and if the noise power spectral density is uniform within the region of the spectrum in which the defect pulse is non-zero, and negligible above that, then $H_{opt}(f)$ simplifies to $X^{*}(f)$. In other words, if the noise is approximately white, then matched filtering is equivalent to correlation of the input signal with a function or series which has the same shape as the desired input, x(t). The operation is correlation rather than convolution in view of the conjugation operation acting upon $X^{*}(f)$, (see eqns. 4.6 and 4.8). The similarities between correlation and matched filtering are particularly obvious in the analogue delay line realisation suggested by **Stremler** (1977).

In order to ensure the validity of this simplification, ensemble averages of the power spectral densities were estimated, from tests run upon 10 samples of each of the axial type defects, and from ordinarily noisy scans. From Fig. 4.15, it can be seen that the spectral power is approximately white in the upper part of the spectra, where all of the spectral energy of the defect signals is concentrated. Thus, it seems reasonable to say that matched filtering would have produced the same results as correlation, since division of the defect signal spectra by the background spectra as per eqn. 4.9 would have made little difference. The results were, then, that correlation was doing a sufficiently good imitation of matched filtering already. The correlation

results in figs. 10, 11, 13 and 14 4 support this proposition since there is little noise present after correlation, thus indicating that in this case correlation is equivalent to matched filtering.

### 4.5 Feature Set Development

In Section 4.3, a segmentation algorithm was described which considerably simplifies the classification problem for the set of single frequency defects, by reducing it to a two class one. The class segmented out (GCO) was also the class which had previously been the easiest to separate. This can be seen from the overall Fisher discriminant ratios computed between the 3 pairs of classes (using the 9-d featureset from training set cp7):

ZCK vs IGSCC 1.097

ZCK vs GCO 2.965

IGSCC vs GCO 3.289

These figures were calculated along the Fisher direction between each pair of classes using the sum of the class standard deviations calculated along the line joining the class means. The two classes remaining after the segmentation algorithm, ZCK and IGSCC, were the pair between which there was most similarity, and the highest pairwise class misclassification rate. It is possible to distiguish between the two types visually using only one scan, so it was reasoned that it ought to be possible to improve the feature set to make the machine classification rate match that of a human operator.

Referring to Table 4.1, features 1-9 were those originally used by **Macleod**(1982) and features 10-14 have been added in order to further differentiate between ZCK and IGSCC defects. The segmentation algorithm

allows legitimate use of the defect signal width as a feature, since previously, the segment start and end were determined somewhat arbitrarily over the length of a whole defect by eye.

The new features were evaluated in two ways: according to their individual Fisher discriminant powers and upon the basis of inter-feature correlation using the standard formula for correlation of two variables, x and y:

$$R_{xy} = \frac{\sum_{k=1}^{n_{samples}} (x_k - \bar{x})(y_k - \bar{y})}{n S_x S_y} \qquad (4.10)$$

The individual Fisher discriminant powers have only been shown between classes ZCK and IGSCC, since it was between these classes that all of the misclassification was occuring. A program which tested the features for cross correlation according to eqn. 4.10 found that features 5 and 7 were 95% correlated, 7 and 8 were -92% correlated, and 10, 11 and 12 were 97% correlated. Contrary to expectations, defect width did not display a particularly high Fisher discriminant power, but since it did not correlate strongly with any other feature, it was added to the featureset.

### 4.6 Visual Analysis of Class Structures

Some of the more common graphical tools used in pattern recognition have been used in this section, to try to analyse the relationship between the classes, and to compare the interclass relationships in the newly extended featurespace with the structure of **Macleod's** data. In order to see the true structure of Macleod's 9-dimensional feature vector set, a nonlinear mapping was coded on the PDP11/45. **Fig. 4.16** shows the points, mapped

according to the distance preserving scheme of **Sammon**(1969). This technique iteratively moves a first approximation of the set of points in 2-dimensions to maximally preserve the interpoint distances as measured in the original feature space. The clusters overlap only slightly, indicating that they are, at worst, almost completely separable with some nonlinear decision boundary.

**Fig. 4.17** shows ~~poor~~ separation of the two least separable classes in the first two dimensions of the feature space transformed according to the Kittler-Young version of the Karhunen-Loeve transformation (**Kittler and Young**,1973). THe Karhunen Loeve transformation takes as its input the covariance matrix of the complete set of feature vectors regardless of class and produces another matrix of orthoganol eigenvectors which represent the directions of maximum variance, which are ranked according to the eigenvalue associated with each of them (eigenvalues are equivalent to the standard deviation of the complete set of points along the direction of the associated eigenvector). The Kittler Young transformation is a modification of the the Karhunen Loeve transformation which emphasises class separation rather than just the directions of maximum variance (which may not be the same as the directions of greatest class separability). The classes are almost separable by a straight line in the full dimensionality of the featurespace they are 94% separable (table 4.2(c)) using a linear decision boundary, compared with the 88.4% achieved by Macleod using his 9-dimensional featureset. From inspection of the eigenvalues of the eigenvector matrix obtained from the Kittler Young transformation, the disriminatory power was not as heavily concentrated in the best ranked feature in the transformed space as it had been in **Macleod's** data. Instead, the first 4 of the 14 eigenvalues were between 50 and 15 in value, within a coordinate system with all feature axes scaled such

that the data has zero mean and variance 100.0. Experience has shown this to be a small spread, indicating substantial discriminatory power in the first four transformed dimensions. Also, from inspection of the mix of features in the highest ranked axis in the transformed space, no one original feature was particularly dominant, in contrast to **Macleod**'s previous findings when separating GCO from the two classes above.

**Fig. 4.18** shows the 3 class data of Macleod transformed according to the method of **Fukunaga and Mantock**(1983). This method, mentioned in Chapter 3, extracts the eigenvalues from a matrix which is calculated in an analogous way to the covariance matrix used in the Karhunen Loeve transformation, except that the geometric mean of the k nearest neighbours (i.e. the local mean) is used in the formulation of the matrix from which the eigenvectors are extracted instead of the sample based mean. Consequently the method makes no assumptions about the nature of the class distributions and so tends to produce a rather more accurate indication of the best 2-space for class separation than the Karhunen-Loeve transformation, since the latter's behaviour is rather subjective when there are more than 2 classes. The figure also shows that there exists a plane in which the 3 classes in the training set completely separate, and in that same plane (i.e. as calculated from the training set) the testing set classes almost linearly separate. This agrees well with the classification results of the next section.

## 4.7 Classification Results

The aims in this section are twofold: firstly, to describe the application of other classification methods over and above the original method of **Duda and**

**Fossum** (1966) used by **Macleod** (1982), and secondly, to find the upper limit on the classification rates given the best set of features thus far found. The links between the optimum achievable misclassification error rate and the results from nearest neighbour rule classifiers were introduced in section 3.5.3. Using the inequalities given there the performance of all the classifiers tried can be assessed more realistically, with estimates of the maximum performance attainable with the feature vector set available, rather than with respect to 100% correct classification.

Since there was a capacity for up to twenty features in all of the programs involved, and none of the features seemed to be actually detrimental to the featureset, all of the old ones and all of the new ones were used at the classification stage, except where it was desirable to use the same featureset as **Macleod** had done for comparison of a type of classifier with his, or to compare the effect of some preprocessing such digital filtering.

### 4.7.1 Previous Work

During the feasibility study only one type of classification was used: 2 stage sequential classification of the 3 classes (tables 1-3,**Macleod**,1982), as already described in section 4.1.

Multiclass classification by sequential 2 class classification requires the system designer to know which classes can be grouped together in order to reduce the effective class count to two at each stage. This would clearly be rather a serious limitation if on site training of the system were to be required. In addition, the problem in Chapter 5 was expected to contain considerably

more classes than that concerning the 316 tubing. Consequently some experiments into multiclass classification have been undertaken.

### 4.7.2 Multiclass Linear Classification

**Duda & Fossum** (1966) give a multiclass linear classifier weights training algorithm in the same paper that they describe their two class scheme. It was coded and tested upon the same 3 class 9-dimensional data set as used by Macleod and the results are shown in Table 4.2.

In some multiclass schemes, c(c-1)/2 discriminant functions must be trained in order to decide between each possible class pair, where c is the number of classes. **Duda and Fossum's** method represents a most economical method of training weight vectors (only one has to be trained per class). The training method is very similar to the 2-class one, the difference being that only the weight vectors associated with the misclassified and misassigned class pair are adjusted after a misclassification.

The misclassification rate was found to increase due to the pre-eminence of the phase feature, which had been so powerful in the discrimination of the GCO class from the union of the ZCK and IGSCC classes. The phase feature attained a high weighting as a result of the aforementioned dichotomy, but the same high weighting confused the classifier when discriminating between ZCK and IGSCC pattern vectors, due to significant overlapping of their class distributions along the phase feature axis.

Upon detailed examination of the misclassified scans, it was again found that the misclassified axial cracks were all of low amplitude. Most were the

first or last scans associated with a particular crack. Almost all of the misclassified IGSCC samples came from one tube data file, which contained a low amplitude intergranular crack. In subsequent tests upon the 14 dimensional data, (see Table 4.2(c)) the misclassification rate was reduced by redistributing the scans from this tube so that half of them were now in the training set. This unfortunately violated an earlier resolution, which was to prevent scans from the same parent blemish appearing in both the testing and training sets even although the scans themselves were different. However, there was no alternative in this case, for this was the lowest amplitude IGSCC specimen in the entire data set, and because there was no other IGSCC quite as low an amplitude by some margin judging by the number of misclassifications due to scans from it, so scans from it had to be redistributed and put into new versions of both the testing and the training sets in order to restore statistical equality to the sets. These new versions, cp7mk2 and cp9mk2, were used to produce **table 4.2(c)**. The number of samples in the IGSCC class is slightly less in this table than in the others because of the redistribution of the scans from this defect.

For feature sets where one feature is significantly more powerful than all of the others (as was the case here, according to individual feature Fisher discriminant ratios), then sequential linear classification is probably preferable to single stage multiclass classification, unless enough is known about the class distributions to train the (c-1)c/2 weights necessary for a full pairwise multiclass classification scheme.

### 4.7.3 Non Linear Classification

A non-linear classifier with a quadratic decision function was written in the early stages of the single frequency experimentation, before it became clear that there were major practical difficulties involved in extending the defect set to a sufficient extent to allow the statistically legitimate use of this classifier. For a non-linear classifier of the polynomial decision function type, the minimum training set class size for a given dimensionality is now governed by the the number of terms in the decision function.

A quadratic decision function can be expressed by

$$G(\underset{\sim}{x}) = \sum_{i,j=0}^{n_{features}} \sum w_k x_i x_j, \quad \text{where } k = i.n_{features} + j, \quad W = [w_o, \ldots, w_{((n+1)(n+2)-1)}] \quad \text{and } X_o = 1 \qquad (4.11)$$

$X \in$ class 1 if $G(x) > 0$

$X \in$ class 2 otherwise

where $x_i$ and $x_j$ are the ith and jth features of the pattern vector $X$, and $w_k$ is the weight associated with that term. For a quadratic classifier, if there are d original features, then there will be

$$\frac{(d+1)(d+2)}{2} \qquad (4.12)$$

terms to train in the augmented weight vector. For the 9 dimensional featurespace in data sets CP7 and CP9, there were 55 weights to train for each 2 class pair. According to standard decision theory, the number of samples per class required must satisfy the condition

$$n_{samples} \geq 2(d+1) \qquad (4.13)$$

where d is the feature space dimensionality. According to this, 112 samples per class are required for the 55 weights necessary to be trained without

underdetermination of the weights occuring (i.e. overfitting of the surface to the point set). However, purely out of curiosity (since the minimum class size conditions could not be fulfilled) the quadratic classifier was tested upon CP9 after having been trained upon CP7, and not surprisingly the results were poorer than those obtained with a linear classifier, due presumably to overtraining of the weights vector, and thus of the decision surface upon the training set.

### 4.7.4 Nearest Mean Classifier

At the opposite extreme as far as the number of weights to be trained is concerned, there is a type of classifier the action of which can be described within the general form of the decision function

$$G(\underset{\sim}{x}) = \underset{\sim}{w}^t\underset{\sim}{x} + v_{d+1},\ \underset{\sim}{x} \in \omega_1 \text{ iff } G(x) > 0 \qquad (4.14)$$

providing that only one training point is allowed per class. This is the nearest mean classifier. By splitting a class into k clusters, and using the mean of each cluster as a prototype and thus treating each cluster as a separate class, a k-means clustering classifier of the type used by **Doctor & Harrington**(1980) (mentioned in section 1.9) can be obtained.

In the two class case, the decision function for the one- training-point-per-class "nearest mean classifier" is

$$G(\underset{\sim}{x}) = \underset{\sim}{x}^t(\underset{\sim}{X}_1-\underset{\sim}{X}_2)-0.5(\underset{\sim}{X}_1^t\underset{\sim}{X}_1-\underset{\sim}{X}_2^t\underset{\sim}{X}_2) \qquad (4.15)$$

where $\underset{\sim}{X}_1$ and $\underset{\sim}{X}_2$ are the class conditional means of classes $\omega_1$ and $\omega_2$ respectively, and x is the new point to be classified.

The method was coded and extended to work on the 3 class data of the

CP7/9 training and testing sets. This 3 class problem was solved with the fundamentally 2-class decision function of eqn. 4.14 by using the three decision functions $G_{12}(x)$, $G_{23}(x)$ and $G_{31}(x)$ as shown in fig. 4.19, between each 2 of the 3 classes, and applying some simple logic to the signs of the functions when evaluated in order to determine in which sector of space a new point had fallen. A nearest mean classifier could have been coded directly, but by calculating the weights vectors first and then modifying the multiclass classifier used for the **Duda and Fossum** type, a result exactly comparable was obtained more easily.

Surprisingly little performance was lost due to the simplicity of this method compared with the iterative method of **Duda and Fossum**. The results in **Table 4.3** were obtained. The overall rate (83%) was slightly lower than that obtained by Macleod (88%) upon the same data, mostly due to IGSCC samples being misclassified as ZCK. This seemed reasonable because of the shorter distance between ZCK and IGSCC classes compared with that of the other two class pairings, and also because the ZCK class was more tightly clustered than the IGSCC class. The results could perhaps have been improved with the use of a weighted nearest mean classifier.

### 4.7.5 Nearest Neighbour Results

A selection of nearest neighbour classifiers were tested upon Macleod's data, for comparison with the results he obtained using the 2 stage linear classifer. The results with k between 1 and 5 were approximately as successful. It was obvious from the list of scan being misclassified given by the classifer at run time that attention should be focused on the dichotomisation of the ZCK and

IGSCC classes, and the overall error rates shown in Table 4.4 are probably erratic for this reason, so no further nearest nighbour experiments were conducted on that data.

### 4.7.6 Bounds on the Bayesian Rate

The misclassification rate was not zero entirely because of the slight interpenetration of the ZCK and IGSCC classes: the GCO class did not contribute to any confusion of the classifier because it was removed completely (either by segmentation or by the action of a linear discriminant boundary).

Using the ZCK and IGSCC classes alone with the 14 dimensional featureset (see Table 4.5) and the same set of scans from the CP7 and CP9 training and testing sets as in **Table 4.4**, nearest neighbour classifiers yielded the improved results given in **Table 4.5**, and consequently the following bounds on the Bayesian error rate:

(i) $E^* > \frac{3.09\%}{2}$ (i.e $E_1/2$) from eqn. 3.31: a not particularly tight bound,

and

(ii) $3.70\% \leq E^* \leq 3.45\%$ i.e $E_{13,7} \leq E^* \leq E_{13}$

from eqn. 3.32: tighter bounds.

The seemingly nonsensible ordering of the bounds is a consequence of the small size of the training set. The value of k was picked according to the rule of thumb that k should be approximately the same size as the sqaure root of the total set size (160, therefor k= 12 or 13), and $l$ was picked for the tightest bound upon $E^*$.

### 4.8 Suggestions for Further Work

Both the segmentation and correlation algorithms are amenable to hardware implementation, using CCD analogue shift registers and multipliers to process the prototype scan. A completely hardwired classifier based upon correlation could be manufactured. It would take the Circograph x and y outputs as its input, and its output could perhaps be a simple visual indicator, e.g., a LED to indicate the presence, and others to indicate the type. A meter might indicate the probability attached to the attributed class of a whole blemish based upon the relative sizes of the peaks from correlation with the prototypes.

In order to partially demonstrate the feasibility of the above, a classifier was implemented by first using the segmentation algorithm from Section 4.3, then masking off the non segmented areas by resetting the signal level to zero (leaving only the ZCK and IGSCC type defects, since the segmentation algorithm deliberately disregarded harmless general surface corrosion). The y-signal alone was then correlated with prototype ZCK and IGSCC y signals, in the manner of Section 4.4.2, in order to classify the defect according to type.

The results were as follows. Over the complete automatically detected defect set i.e. the union of the training and testing sets, 30 from 169 ZCK scans and 30 from 146 IGSCC scans were misclassified, i.e. 1 ZCK blemish from 17 and 2 IGSCC blemishes from 13. As was the case for all the other types of classifier, the misclassified scans tended to be from uniformly low amplitude blemishes or from scans near to blemish ends, which were also of a low amplitude. Specifically, defects with peak heights less than 15 adc units were at risk, especially intergranular cracks with a rather sharper than average

central peak, thus making them resemble an axial crack.

The most interesting features of this classifier were its simplicity (since it effectively extracted only one feature) and its lack of a training set requirement, other than 2 prototype scans. In a practical application, these and the background noise level would be interactively determined on site prior to its use. This would tailor the system to the particular surface conditions for a given set of tubes. The method, when implemented in software was quite slow in execution (45 seconds/scan), but it could be made to operate much faster in a purely hardware form, since the multiplications involved in the correlation could be performed with fast analogue multipliers.

Only the y signal has been used, since the x signal correlation tended to misclassify more often. However, the misclassification rate might be lowered by AND gating the classifications from the y signal with those from the x-signal, supplemented by an amplitude threshold of 15 adc units on the minimum peak height acceptable. All of this depends upon the phase shift applied to the signal being calibrated such that the proportion of y-signal to x-signal is maximised for an axial crack.

In conclusion then, this classifier/detector would both detect and classify defects at a speed sufficient to keep up with the circograph's probe (3 revolutions/second, sampled at 400 points/scan) were it to be built. It would need to store a separate prototype scan for each type of defect anticipated, and have either another set of segmentation and correlation circuitry for each additional defect class, or facilities to multiplex the first hardware correlator.

The correlation —based classifer was not the only novel classifier evaluated. Another classifier (normally hardware- based) which treated the impedance plane diagram as a bit map and used these bits in groups of 8-tuples as feature vectors was coded and tested, but the results were poorer (75% correct classification) than merit its inclusion here.

### 4.9 Concluding Remarks

The same basic data sets as used by **Macleod** in his reports have been adhered to for comparability, except in one respect. The scans from the most extreme example of a low amplitude IGSCC crack had to be split between the testing set CP9 and the training set CP7, because an imbalance had been caused previously due to the inclusion of such a defect in one set only. The misclassification results improved when the scans were split between the training set and the testing set.

If the weight vector for the IGSCC class was changed, so that the misclassification rates for each of the classes was approximately equal, then majority voting might be powerful enough to bring the rates up to 100% over complete blemishes. Regardless of the type of classifier employed, the misclassified scans tended to be low amplitude ZCKs misclassified as IGSCC, and vice versa. The GCO samples did not misclassify nearly as often as did the other two classes, but when they did, they tended to be deemed IGSCC rather than ZCK. The misclassified ZCK samples tended to come from the lowest amplitude axial cracks. The misclassified IGSCC samples similarly were almost invariably those of lowest amplitude.

## 4.10 Conclusions

The confidence limits imposed upon all of the performance figures by the sample set size were 5% with respect to the measured results for 3 classes each of size 50, though to calculate this the usual assumptions had to be made concerning class distributions. The scatter plots in fig 4.18 suggest that the class distributions are at least approximately unimodal if not exactly Gaussian. Thus the experimental results were insignificantly different from the theoretically achievable maximum with the 14 dimensional featureset.

The results from the nearest neighbour estimates of the Bayesian misclassification error rate indicate that the best results obtained from the linear and nearest neighbour rule based classifiers tested could not have been improved much further, if at all.

## 4.11 References

1. **Macleod, J.E.S.**, "Feasibility Study on Computer Sorting of Flaws in Steam Generator Tubes Scanned by a Rotating Eddy Current Scanner", Babcock Power Research Report No. 07/82/46, July 1982.

2.**Duda, R.O. and Fossum, H.**, "Pattern Classification by Iteratively Determined Linear and Piecewise Linear Discriminant Functions", IEEE Trans. Comput., vol. EC-15, no.2, pp 220-232, April, 1966.

3. **Brigham, E.O.**, "The Fast Fourier Transform", New Jersey, Prentice-Hall, 1974.

4. **Carlson, B.A.**, "Communication Systems", pp156-158, Kogakusha, Mcgraw-Hill, 1975.

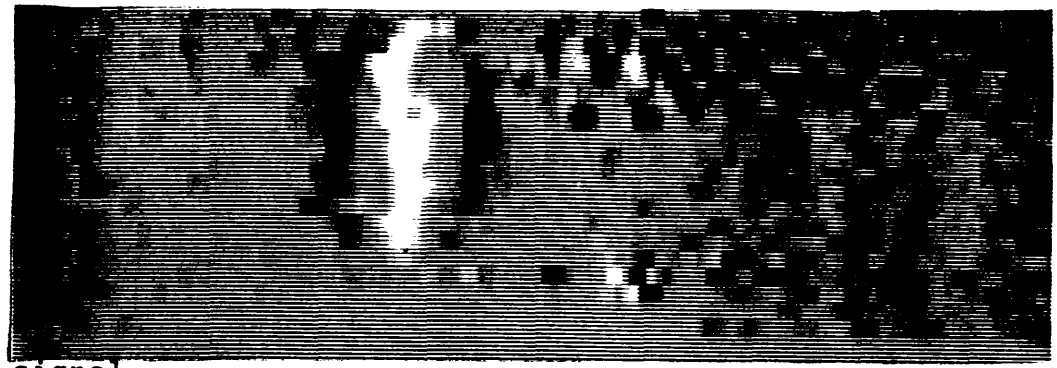
5. **Stremler, F.G.**, "Introduction to Communication Systems", Section 7.6, Addison-Wesley, 1977.

6. **Kittler, J., and Young, P.C.**, "A new approach to feature selection based on the Karhunen-Loeve expansion", Pattern Recognition, Vol. 5, pp. 335-352, 1973.

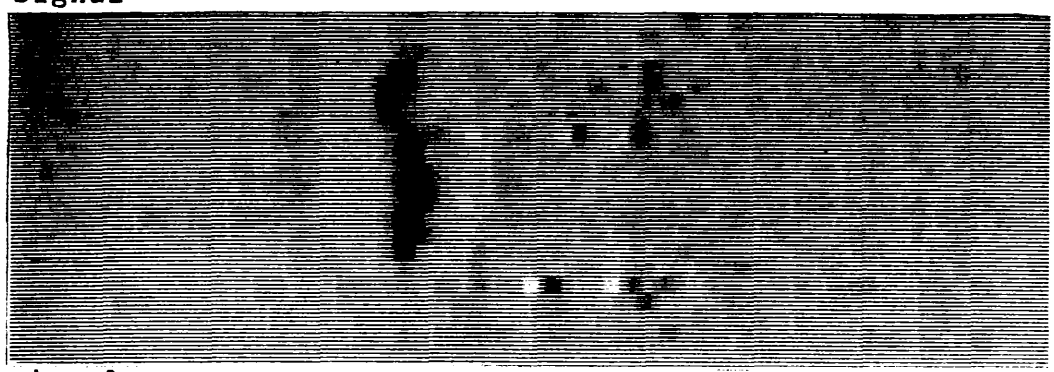
7. **Sammon, J.W.**, "A Nonlinear Mapping for Data Structure Analysis", IEEE Transactions on Computers, Vol. C-18, No. 5, pp401-409, May 1969.

8. **Fukunaga, K. and Mantock, J.M.** ,"Nonparametric Discriminant Analysis", IEEE PAMI-5, No.6, pp671-678,Nov.,1983.

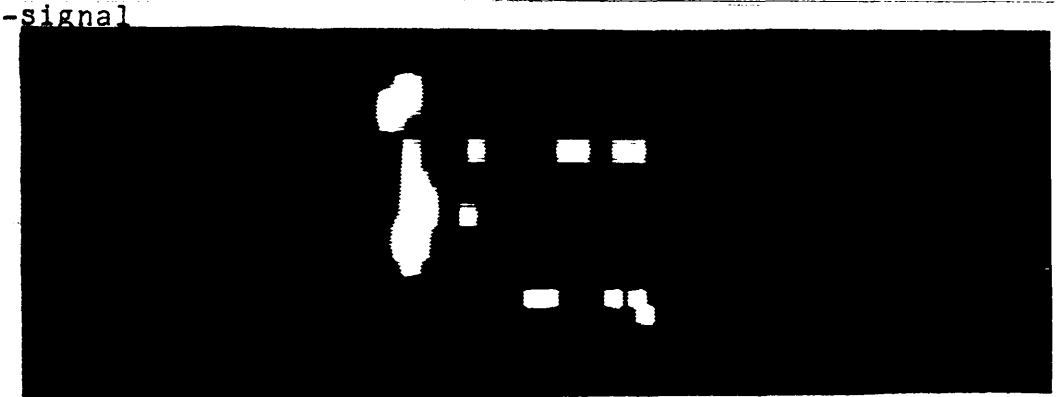
9. **Doctor, P.G. and Harrington, T.P.**, "Analysis of Eddy Current Data Using Pattern Recognition Methods", IEEE 5th Conf, on Patt. Rec., Miami, 1980.



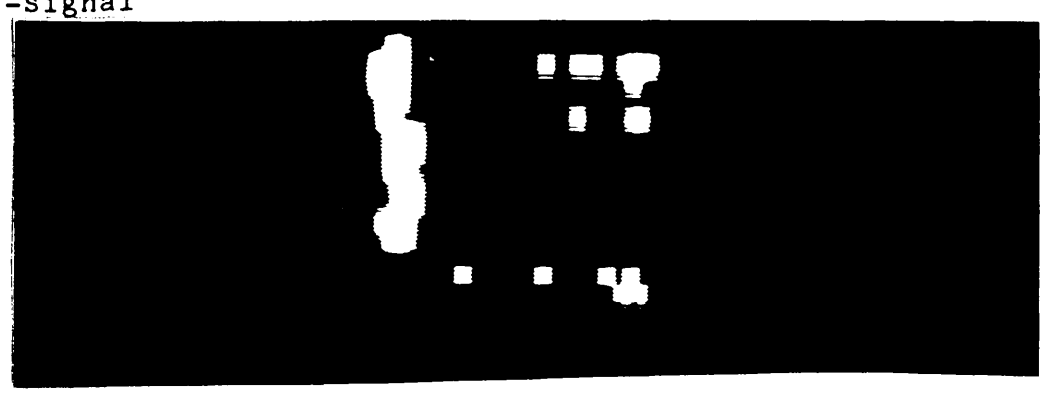
Y-signal



X-signal



Y-signal



X-signal

Intergranular cracking

FIG. 4.1 shown on an image processing system, with brightness proportional to signal voltage over the range -127 to +128 adc units. The bottom pair are thresholded according to the criteria: if $|\text{signal}| > |\text{signal}_{max}|/2$ then pixel=ON, else pixel=OFF.

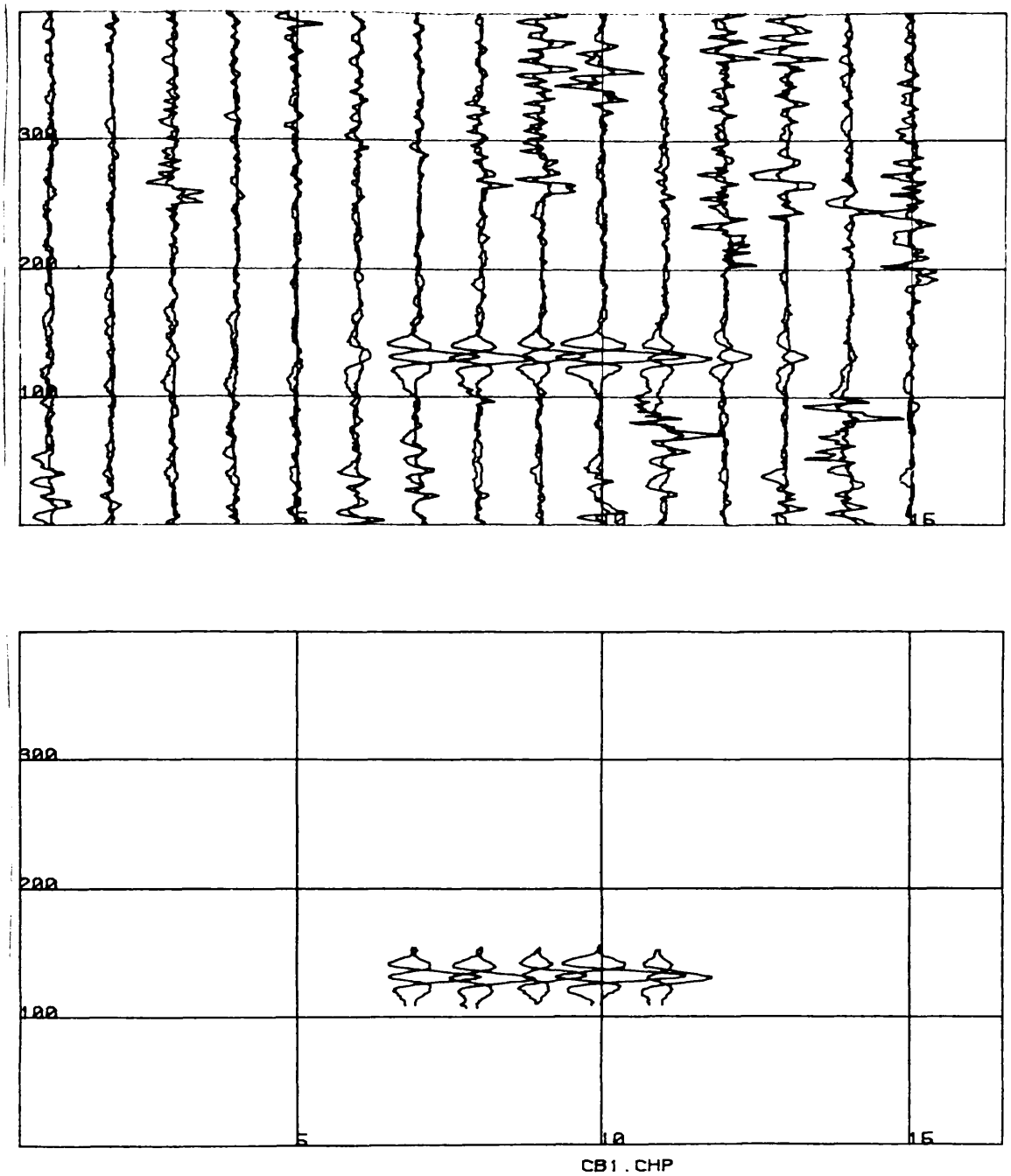
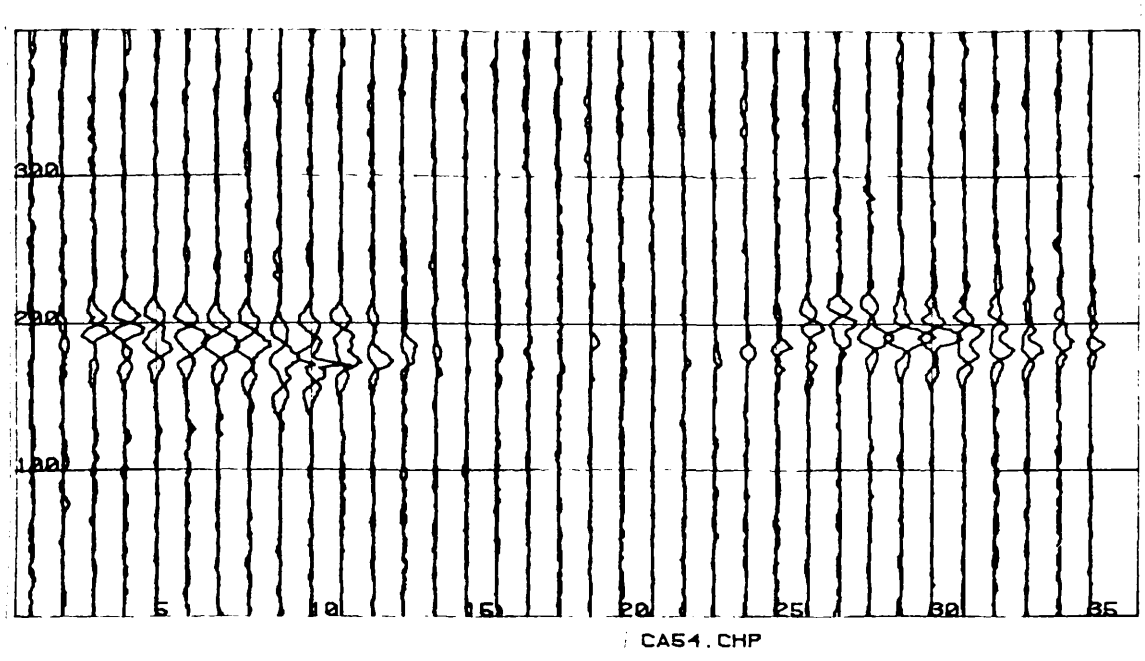



FIG. 4.2 The effect of the segmentation algorithm of section 4.3, upon an axial crack, in a tube which also suffered from substantial surface corrosion. The defect is not detected at its ends, in scans 6,12 &13.



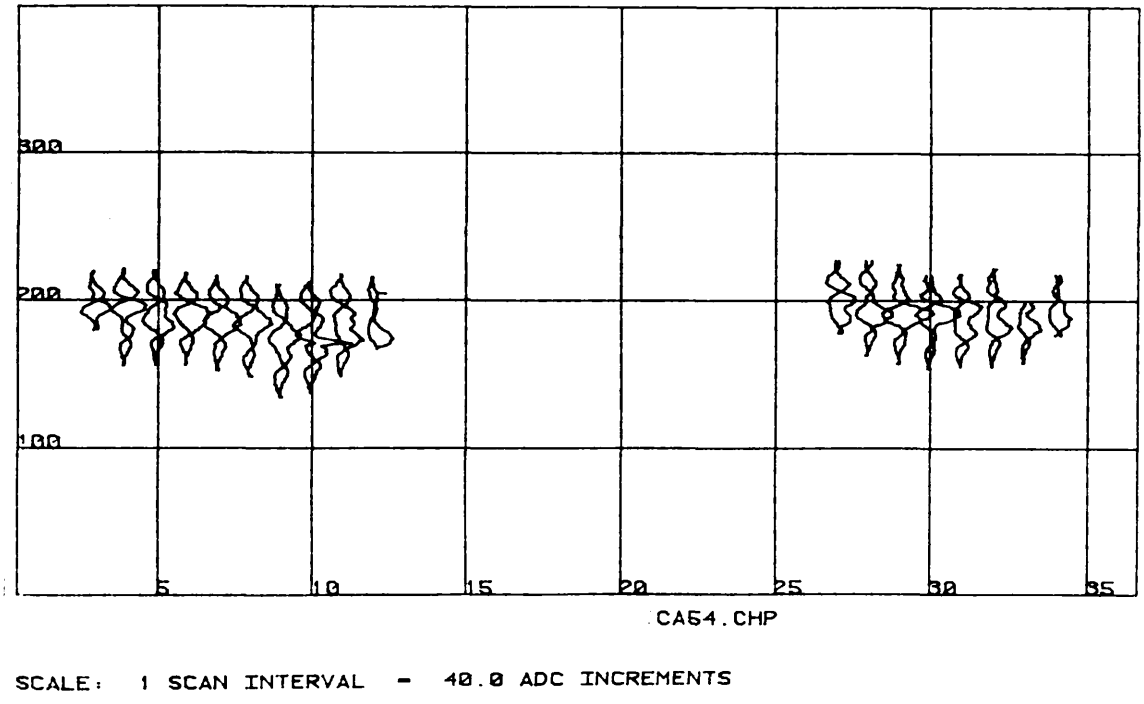
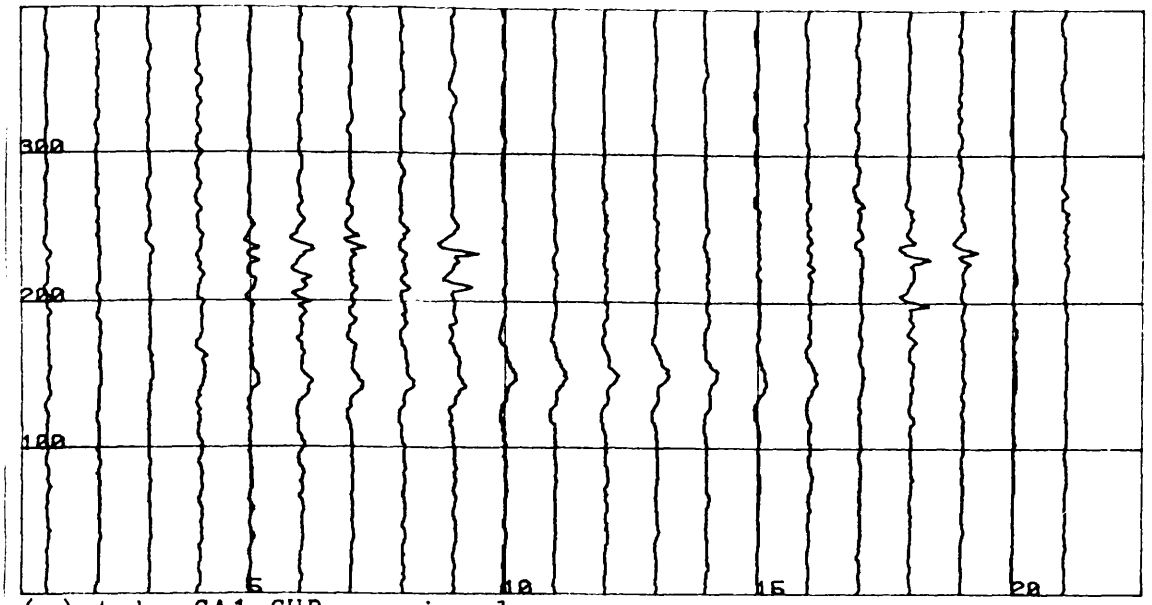
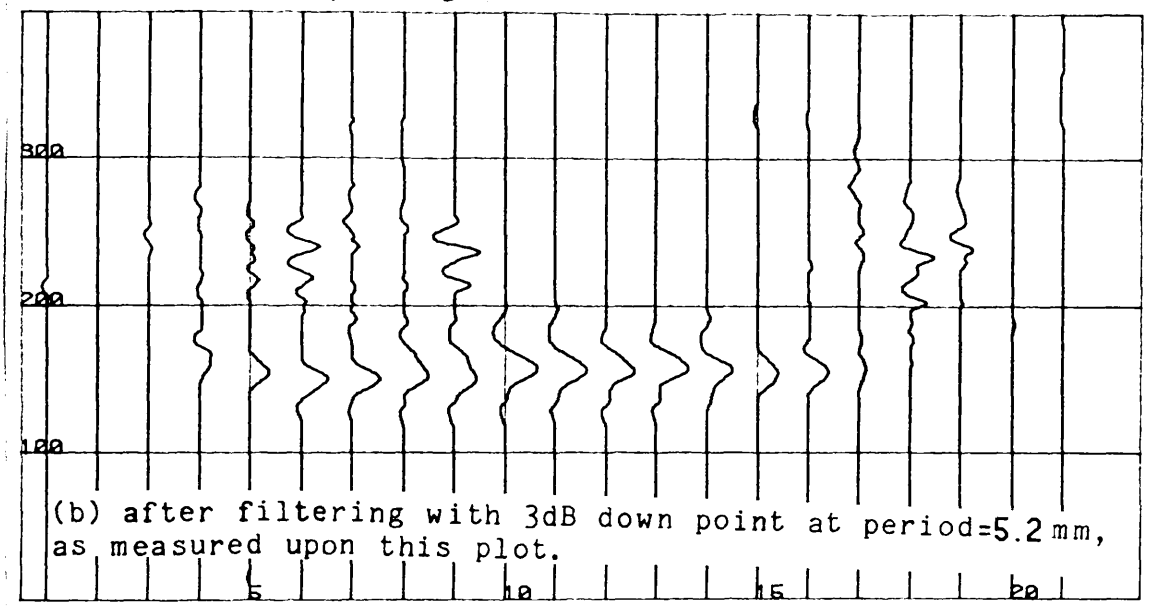



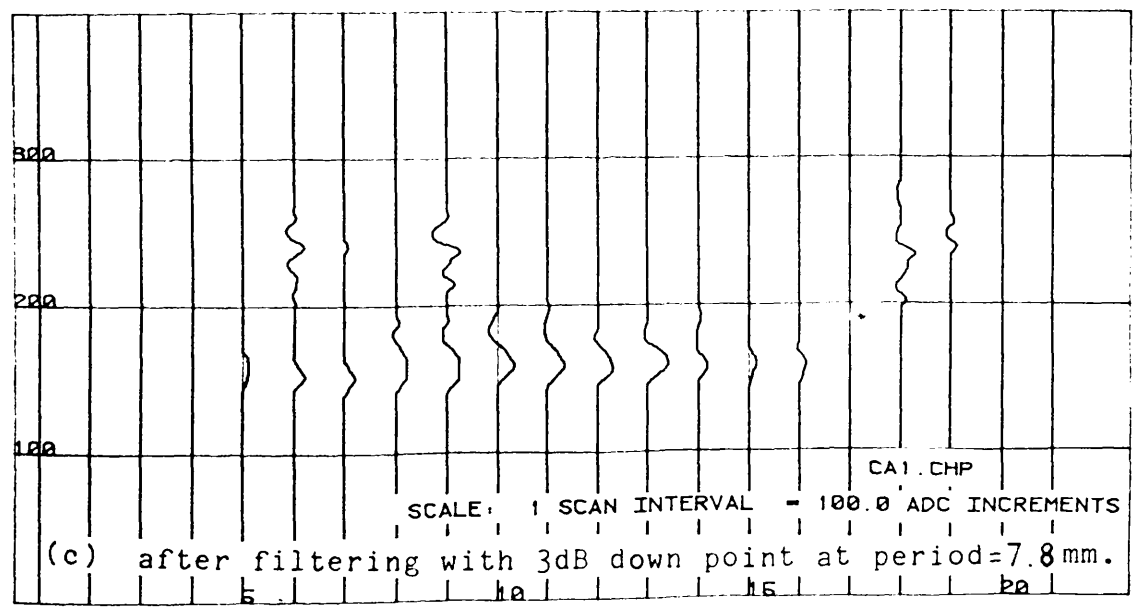

FIG. 4.3 The effect of the segmentation algorithm upon an example of low amplitude intergranular cracking.




(a) tube CA1.CHP, y-signal.

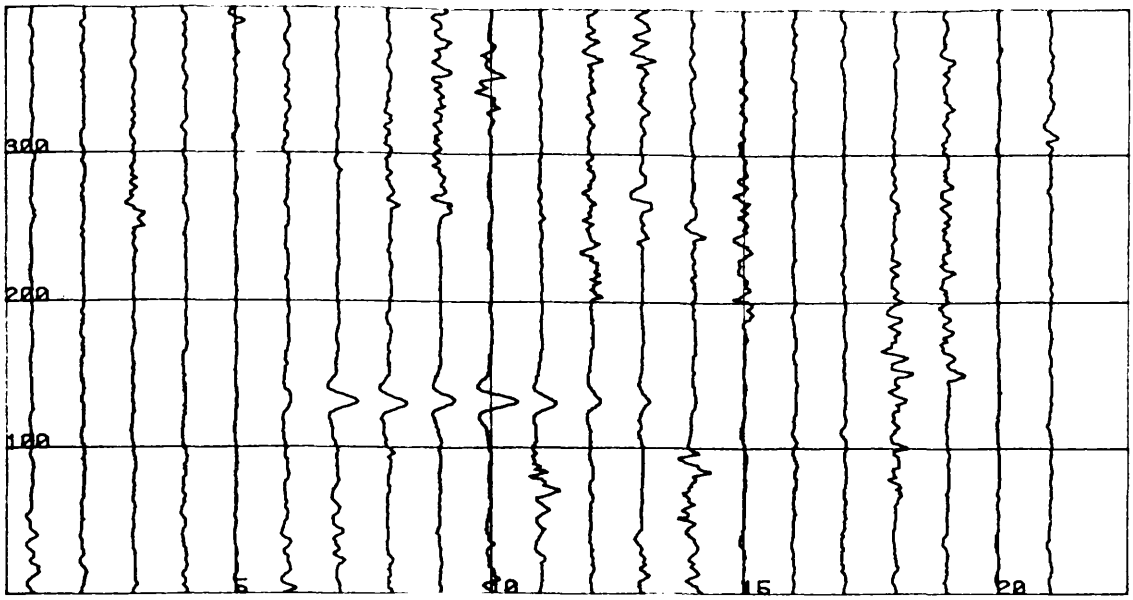



(b) after filtering with 3dB down point at period=5.2 mm, as measured upon this plot.

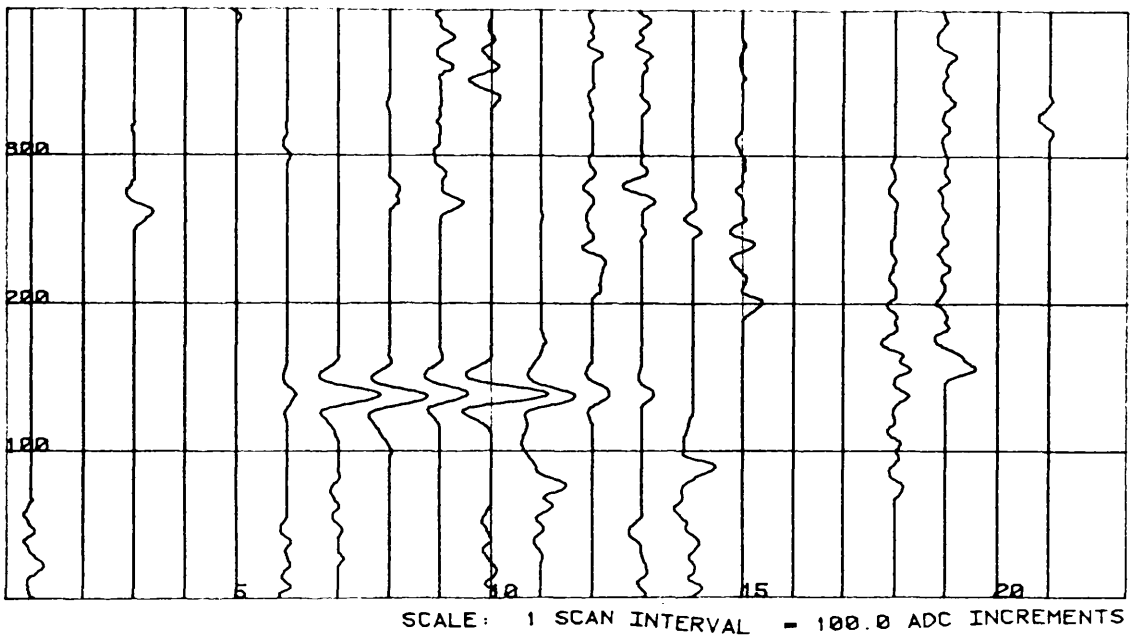



(c) after filtering with 3dB down point at period=7.8 mm.

FIG. 4.4 The effect of Bessel filtering upon an intergranular cracking signal.




(a) tube CB1.CHP, y-signal.




(b) after filtering with 3dB down point at period=5.2mm, as measured upon this plot.

FIG. 4.5 The effect of 2nd order Bessel filtering upon an axial crack signal, with adjacent surface corrosion.

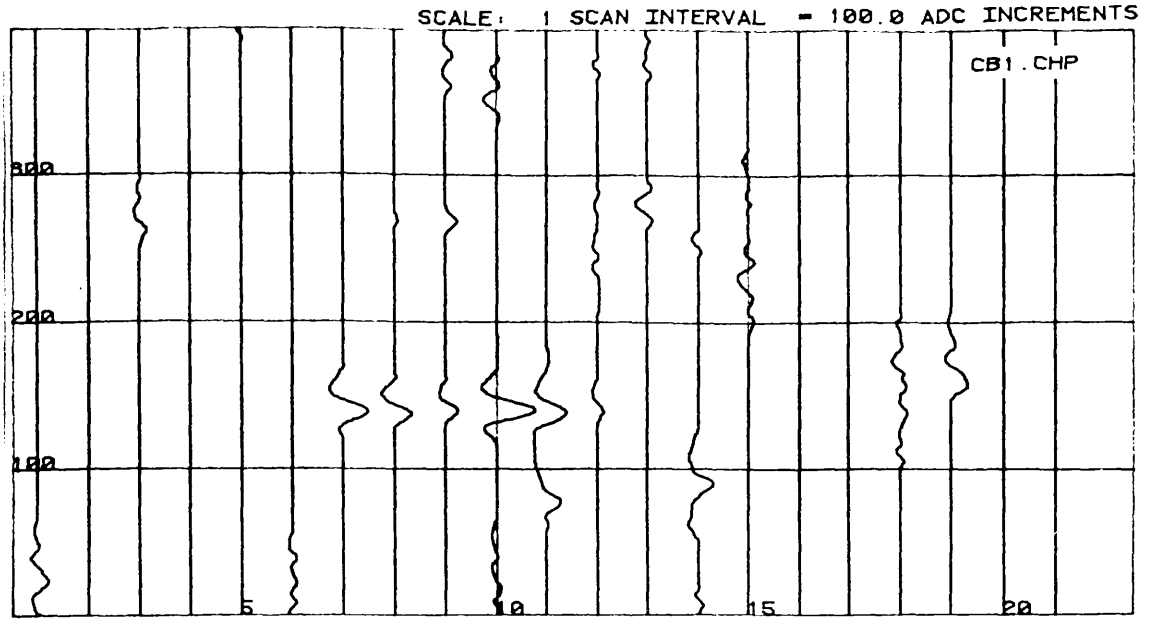



FIG. 4.5 (c) AXIAL CRACK, 2ND ORDER BESSEL FILTER, 3 DB DOWN PT. AT PERIOD OF 7.8 MM.

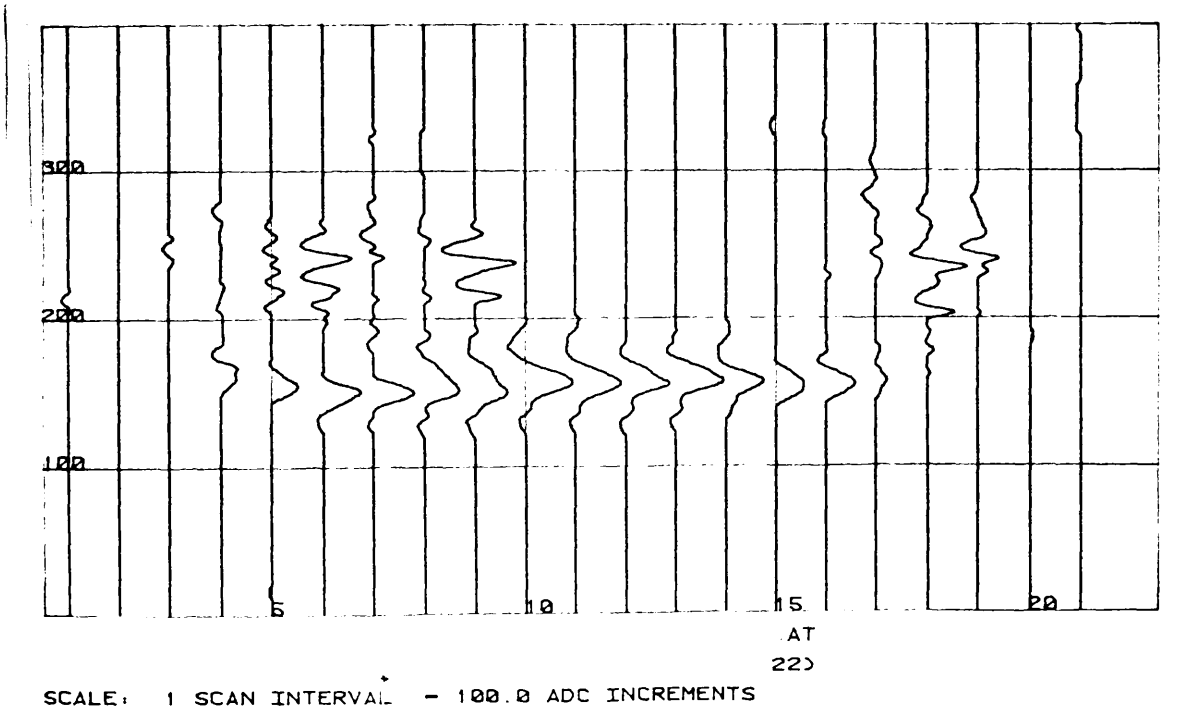



FIG. 4.6 IG AND GCO, 4TH ORDER BESSEL , 3 DB DOWN PT. AT PERIOD OF 5.2 MM.

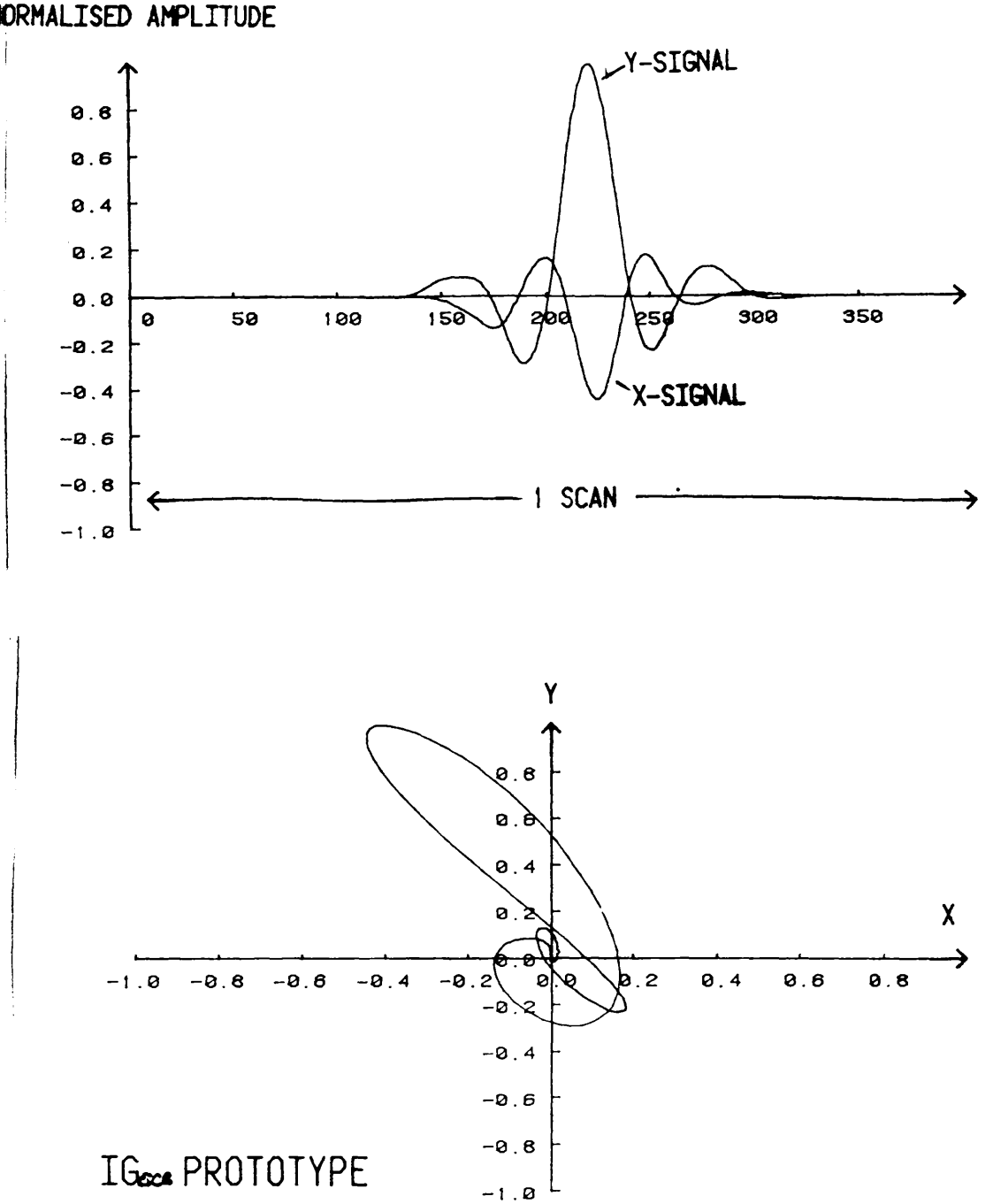



FIG. 4.7 The axial crack prototype scan, used for the cross correlation trials. It was originally scan 8 from tube file CB1.CHP and subsequently filtered, windowed and normalised to unity height largest peak. The lower figure shows the impedance plane signal locus, whilst the upper one shows both the x & y time series for the complete scan.

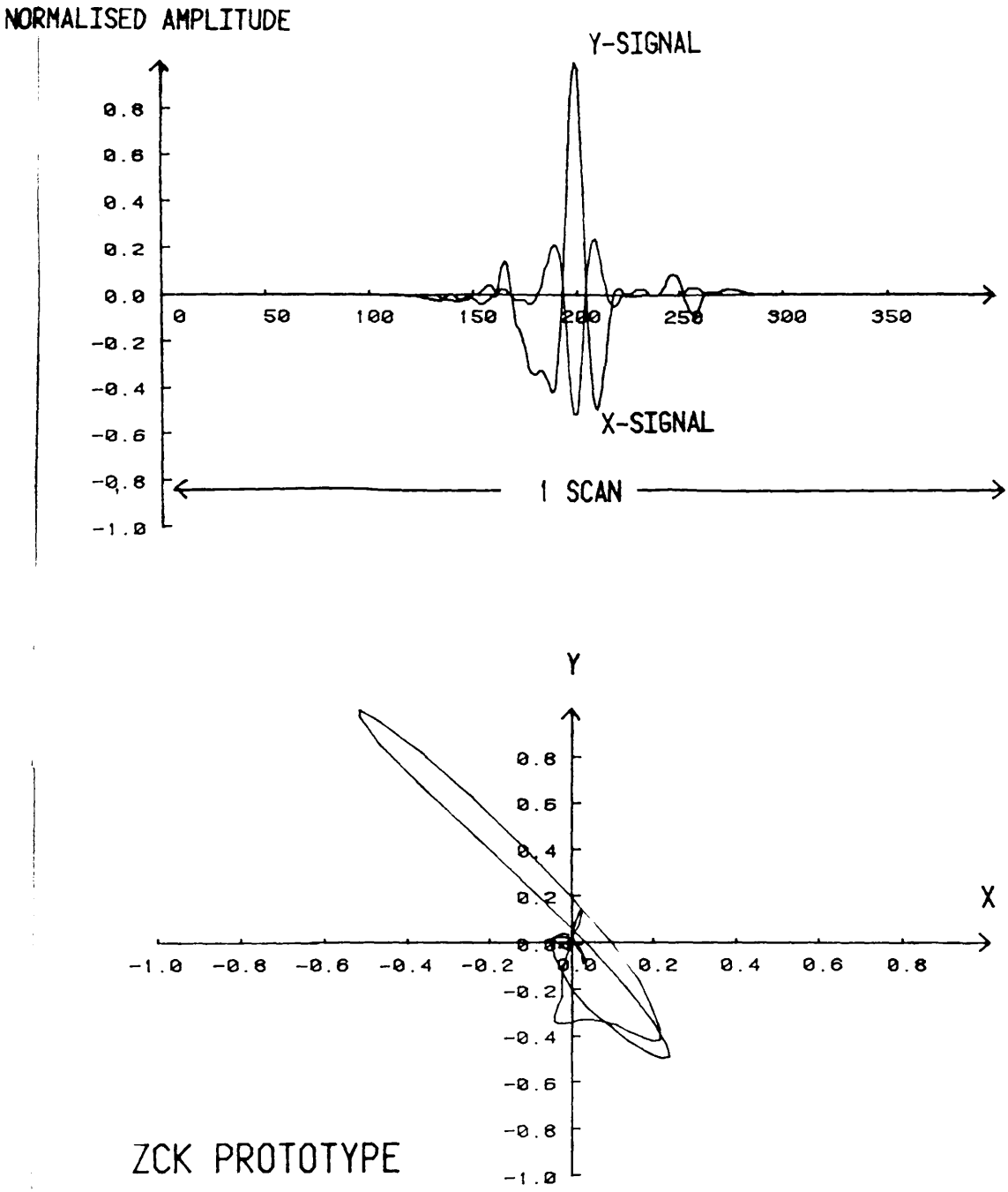



FIG. 4.8 .The intergranular cracking prototype developed for cross correlation purposes, for comparison with fig. 4.7

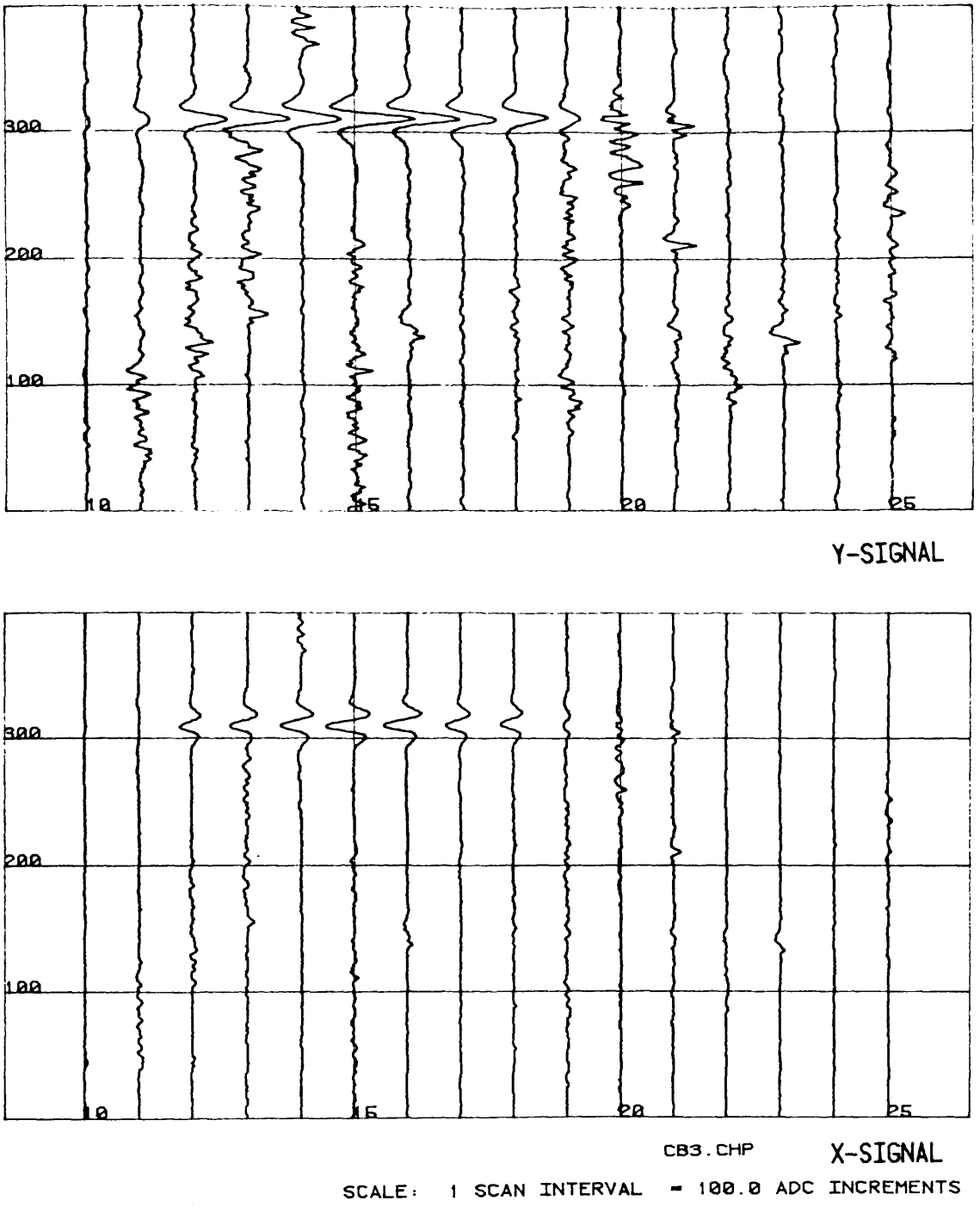
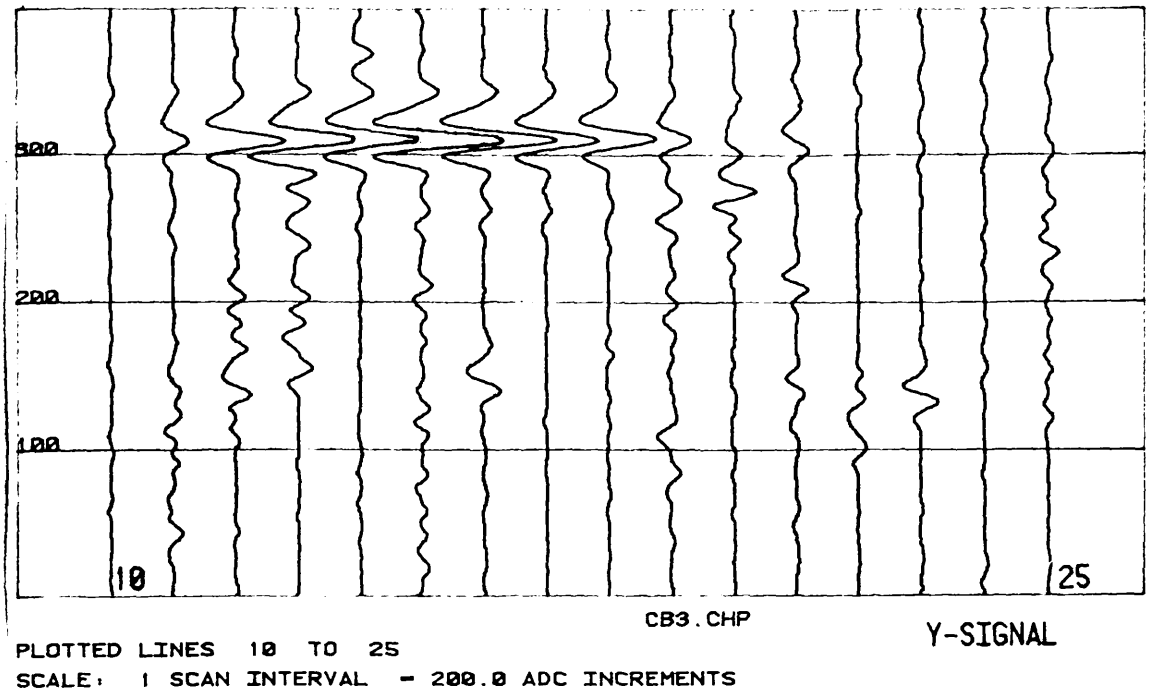



FIG. 4.9. The quadrature(top plot) and inphase(bottom) signals from a tube containing a well defined axial crack and extensive surface corrosion.



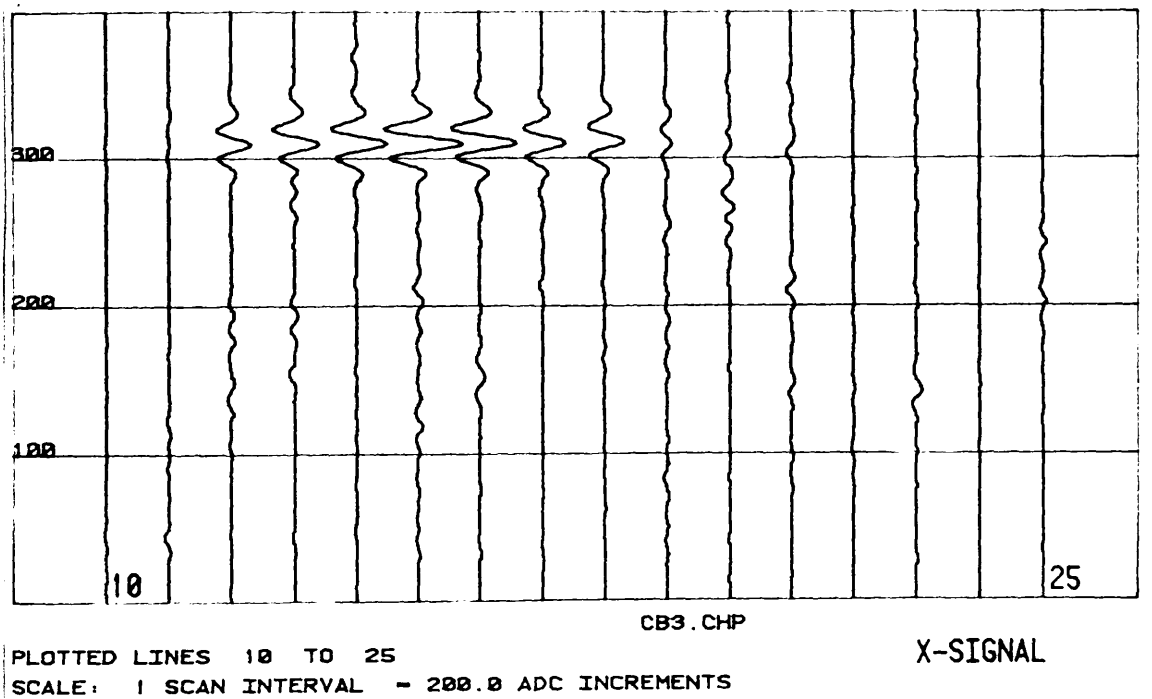
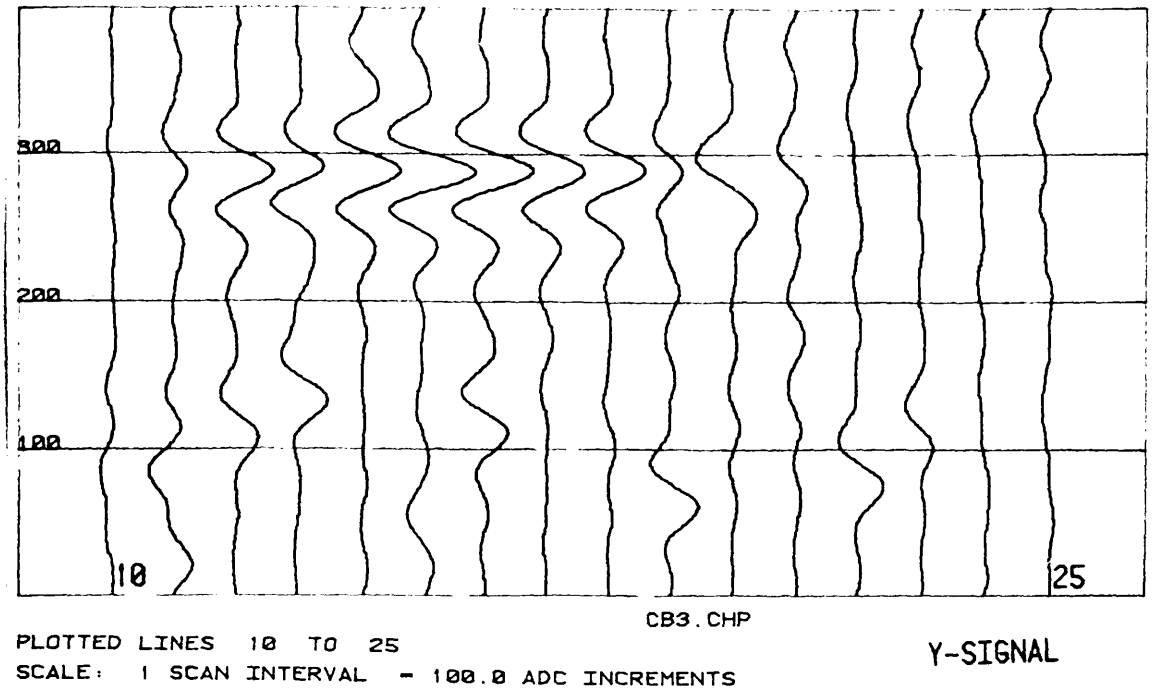




FIG. 4.10 The result of the cross correlation of an axail crack with the prototype axial crack scan of fig. 4.8, for comparison with figure 4.11.



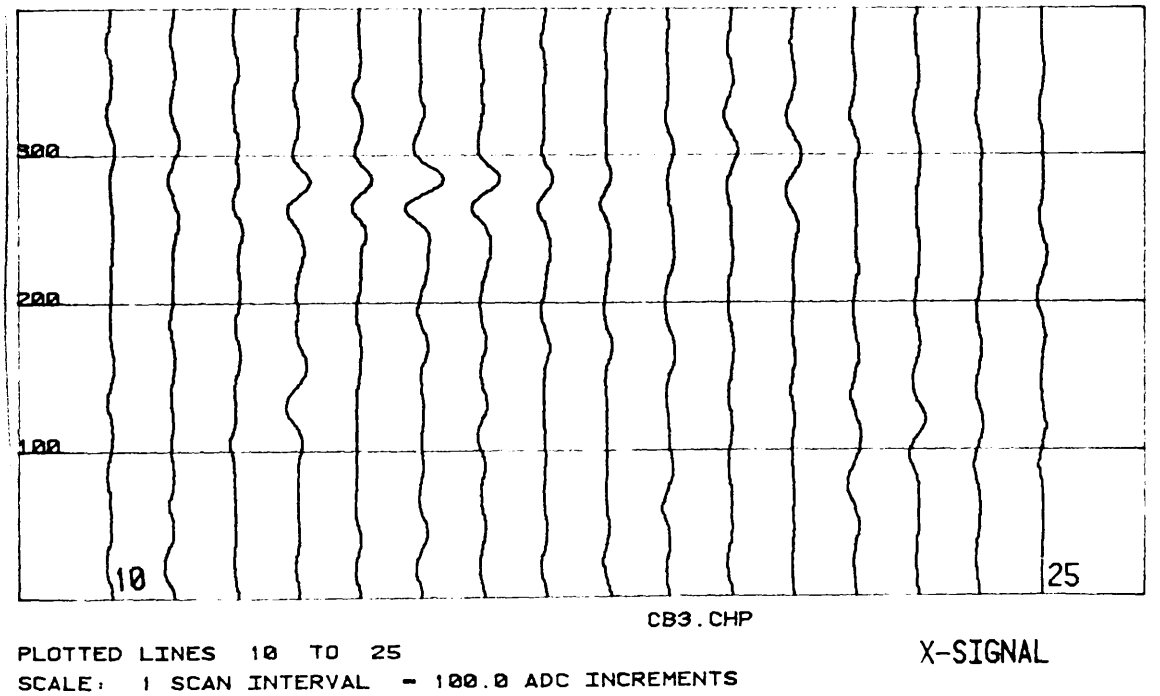
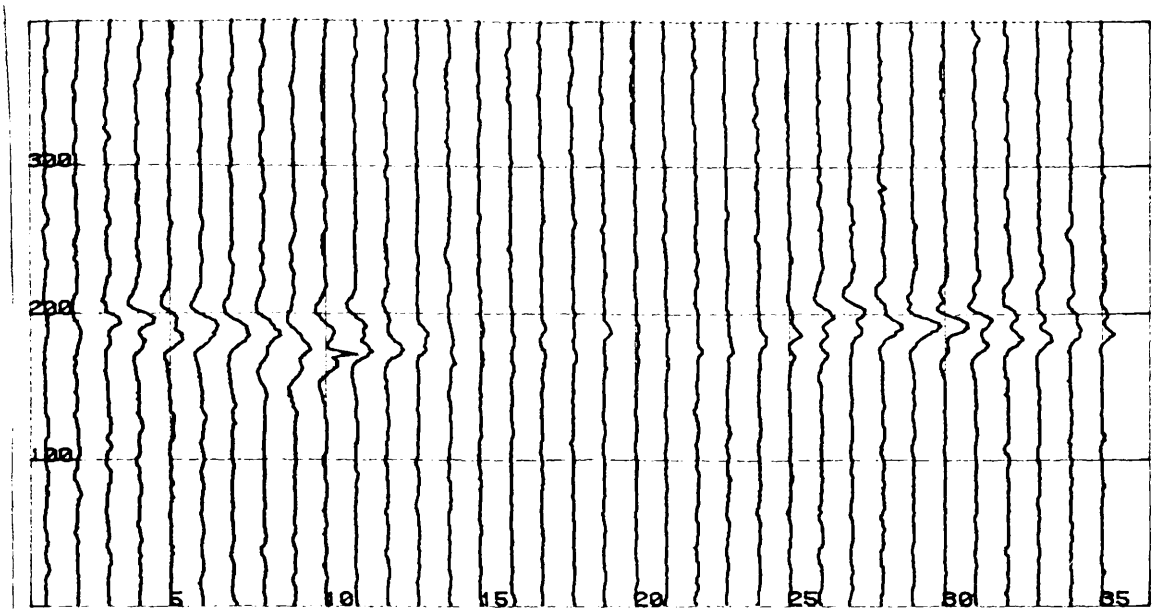
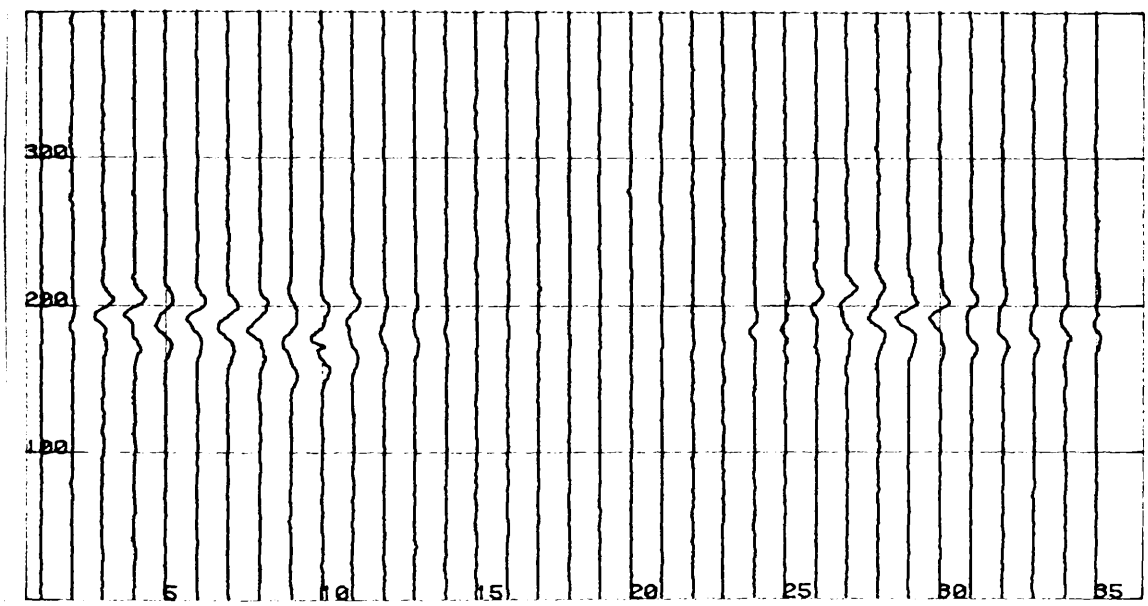




FIG. 4.11 Cross correlation of the axial crack of fig. 4.10 with the prototype intergranular cracking of fig. 4.7 (Note the scale differences between this figure and that of 4.10 .)




Y-SIGNAL

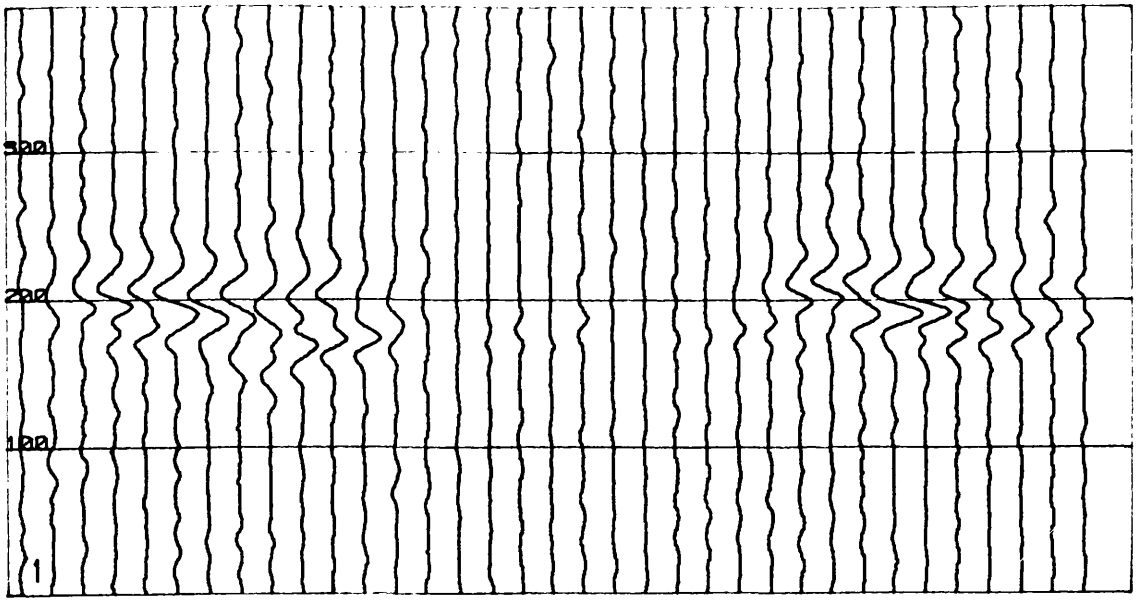



CA54.CHP

SCALE: 1 SCAN INTERVAL = 50.0 ADC INCREMENTS

X-SIGNAL

FIG. 4.12 The quadrature (top) and inphase (bottom) signal plots (y and x respectively) for a tube containing two areas of intergranular cracking. This particular example was chosen because of its low overall amplitude, and the resemblance of scans 6,7 and 29 to broadened axail cracks.



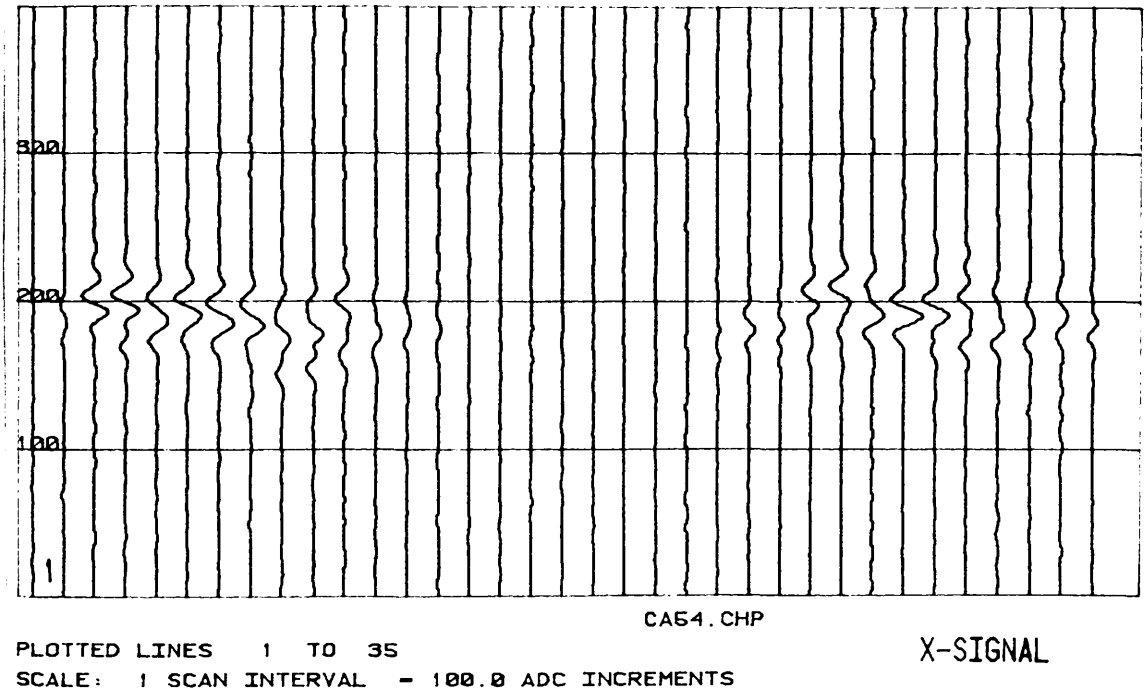




FIG. 4.13 Intergranular cracking, cross correlated with the prototype axial notch of FIG. 4.8. The y-signal is weak and ill-defined in comparison with that of FIG. 4.14.

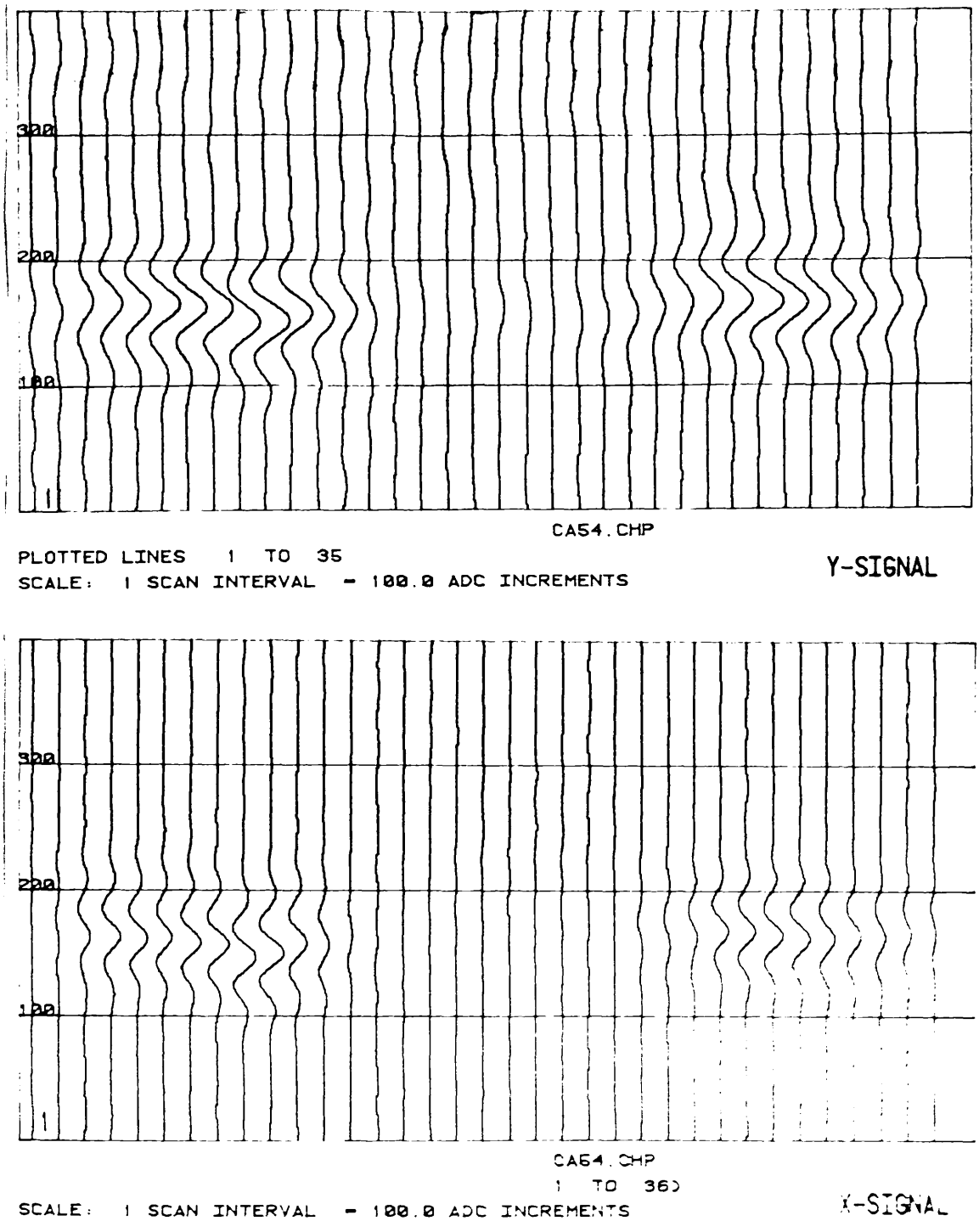






FIG. 4.14 The tube shown in FIG. 4.12, after cross correlation with the prototype intergranular cracking of FIG. 4.7. The scale diffrence between the x & y channels is predominantly due to the larger y signal in both the tube and the prototype; however, the there is also obviously better correlation.

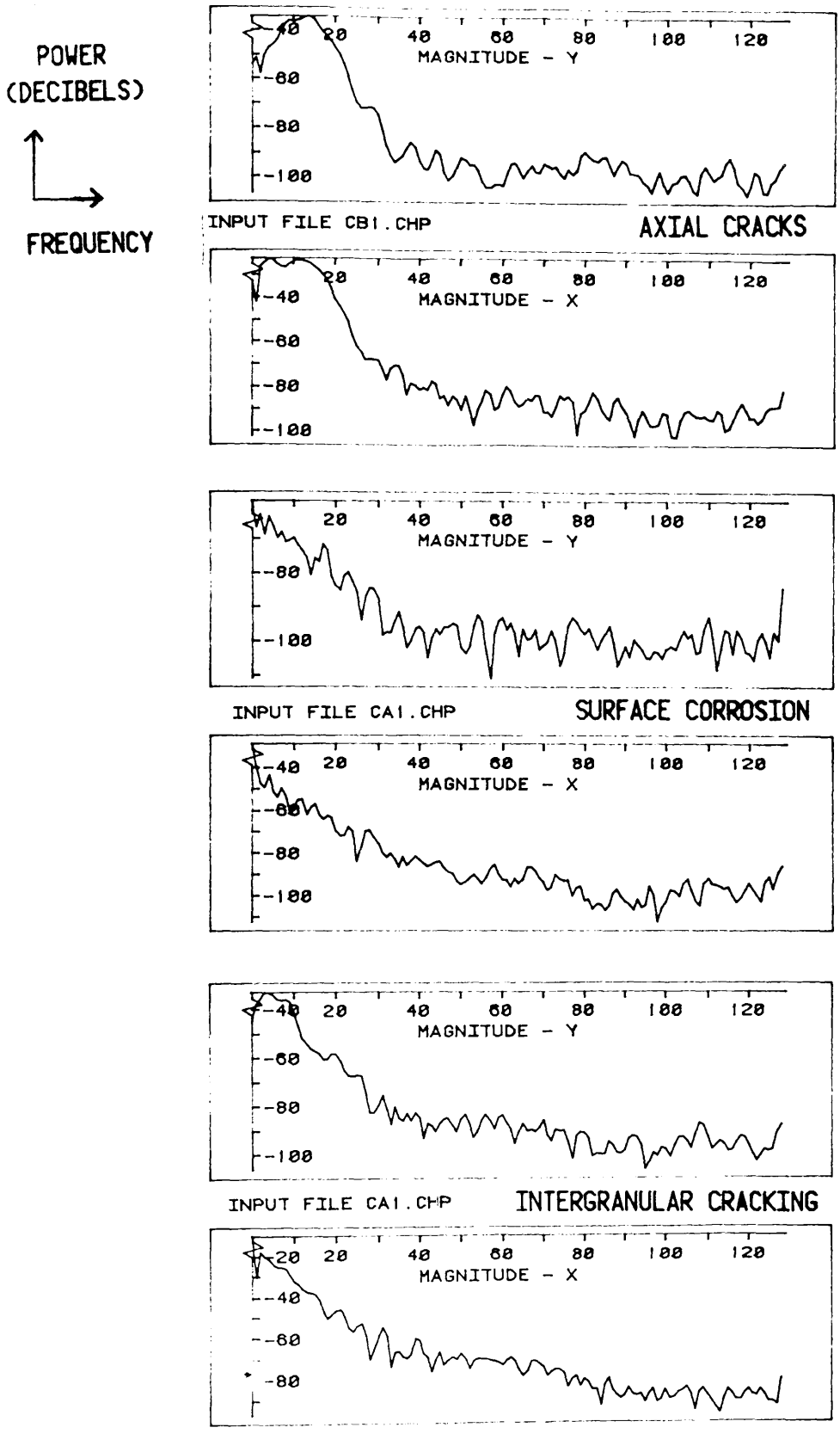



FIG. 4.15 Ensemble-averaged power spectra for the three defect classes. Vertical scale is in dB. The horizontal is in terms of the FFT transform size: 0 to 128 linearly represents d.c. to 2.58 cycles/mm, measured w.r.t. the tube plots in previous figures.

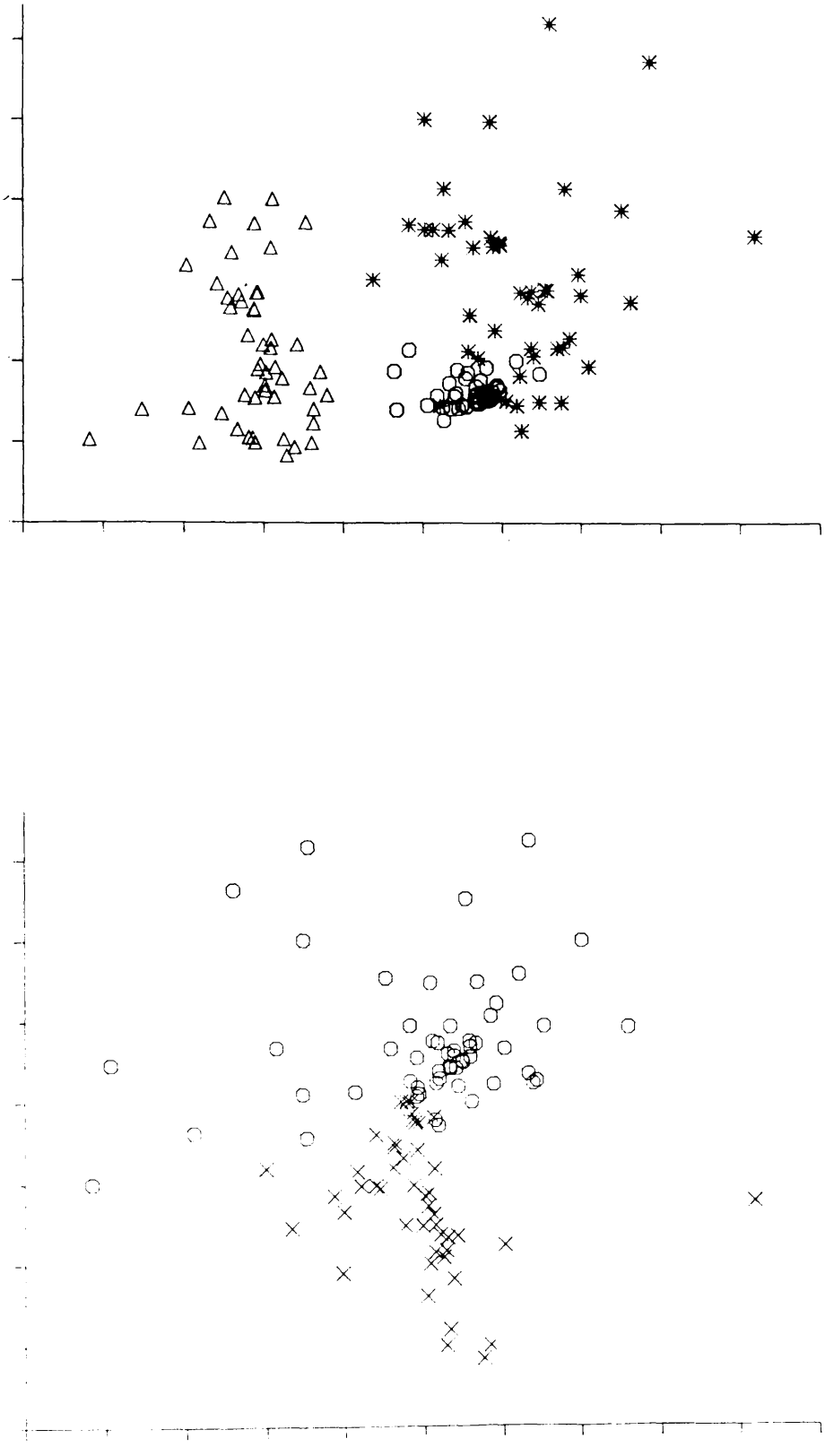


FIG. 4.16 A NON LINEAR MAPPINGS X - ZCK O - IGscc Δ - GCO

SEE SECTION 4.6

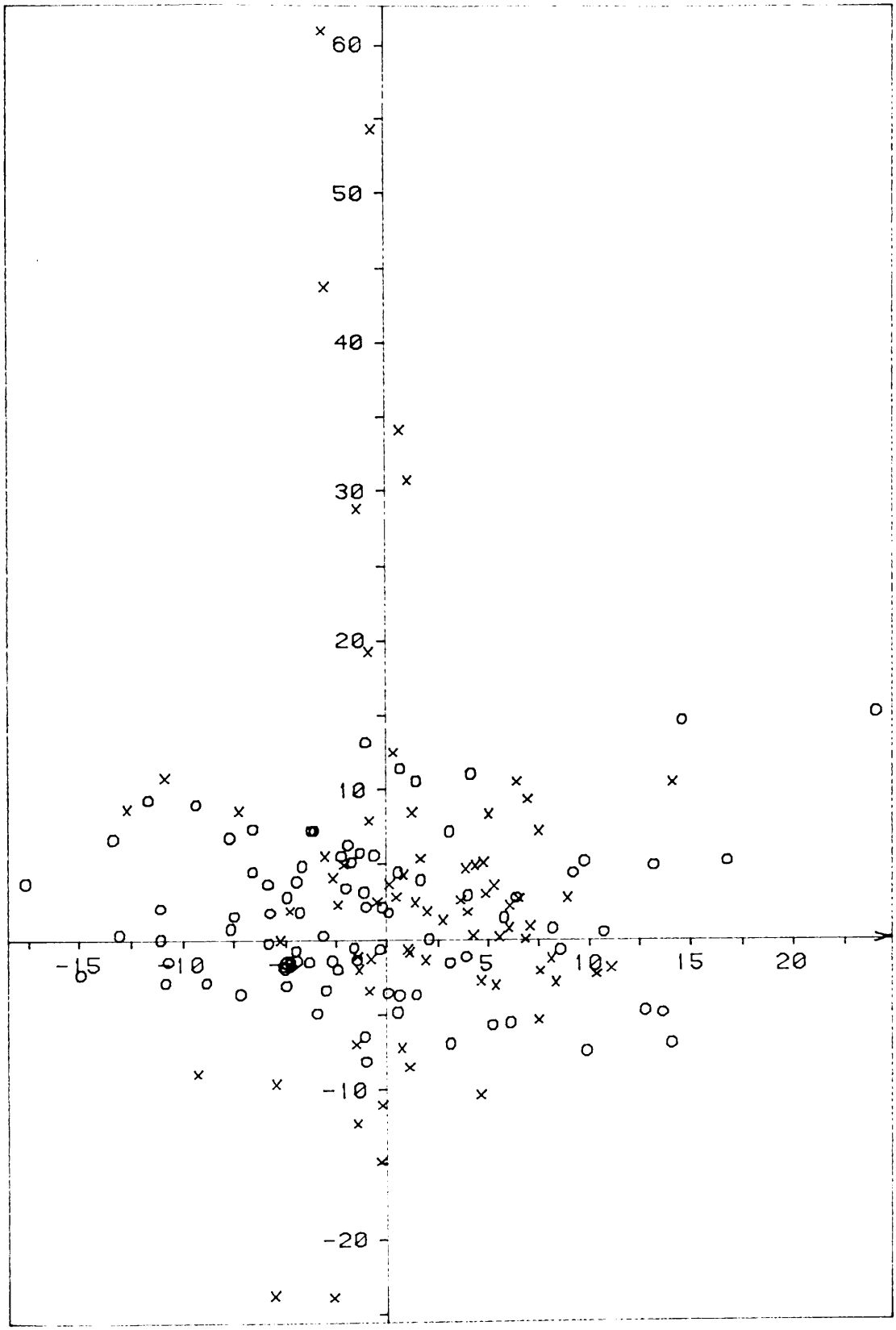
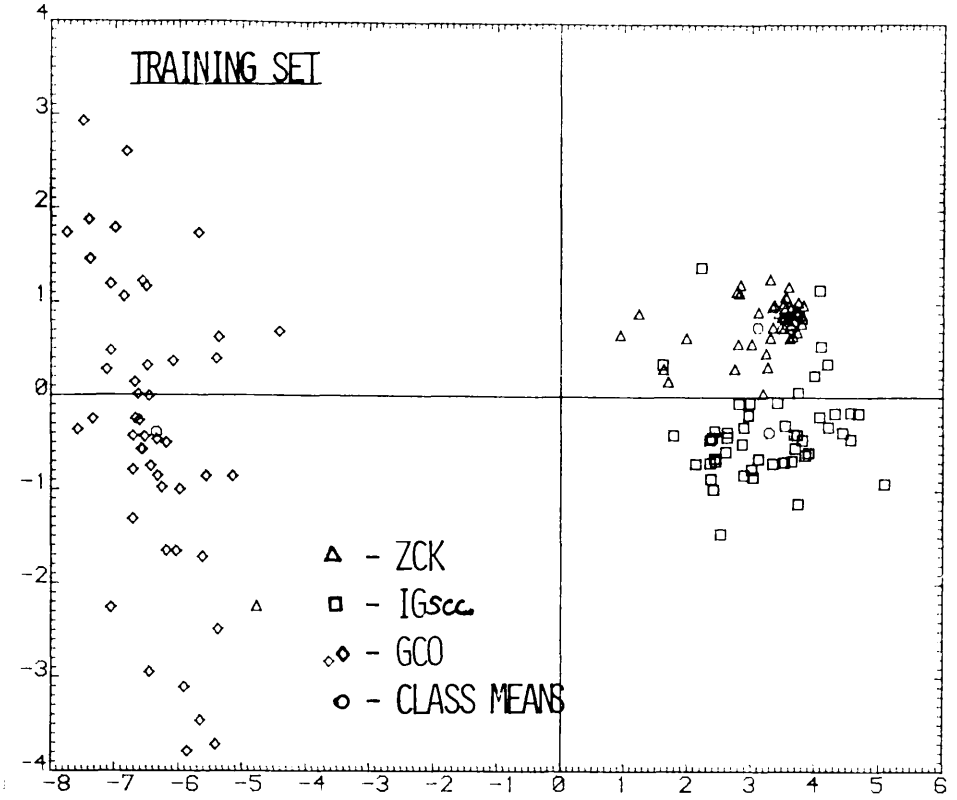



FIG. 4.17 THE K-Y TRANSFORMATION OF 14 DIMENSIONAL ZCK AND IGκSAMPLES (CP9 DATA SET)

X-ZCK O-IGR



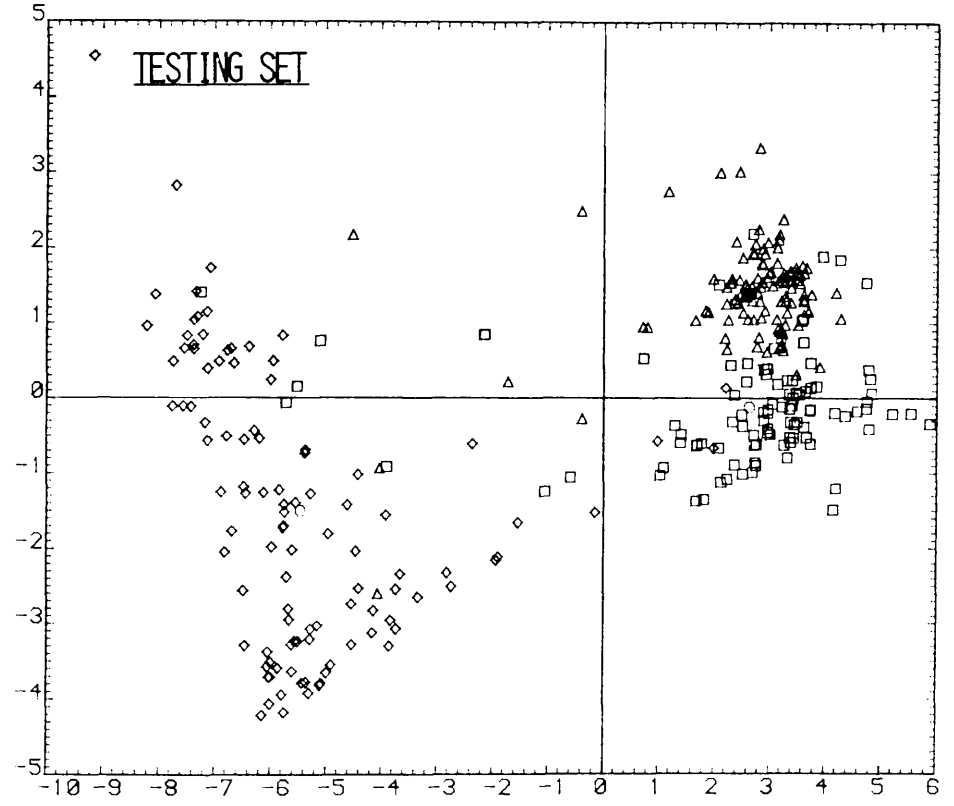




FIG. 4.18 PROJECTION OF TRAINING AND TESTING SETS BY NON PARAMETRIC DISCRIMINANT ANALYSIS METHOD OF FUKUNAGA AND MANTOCK. (SEE SECTION 3.43)

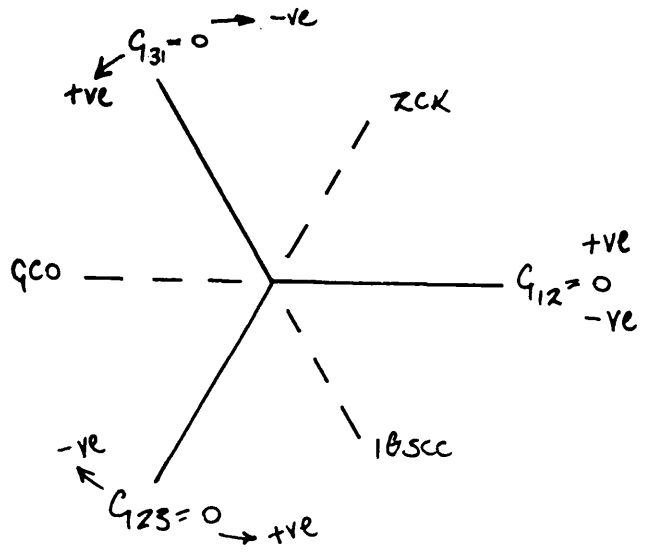



Fig. 4.19 The use of three two class linear classifiers to dichotomise the featurespace according to the three classes. Class membership is determined by a logical AND performed upon the output of each function i.e. whether it is greater than or equal to zero for an as yet unclassified pattern X, or not. For example $X \in GCO$ iff $(G_{23}(X) < \emptyset)$ AND $(G_{31}(X) > \emptyset)$, and the value of $G_{12}(X)$ is unimportant.

**Table 4.1 Fisher Discriminant Ratios for Axial Cracks (ZCK) against Integrgranular Cracking (IGR)**

| feature no. | description | Fisher Discriminant Ratio |
|---|---|---|
| 1. | Average impedance plane phase i.e. slope of best straight line through main lobe. | Ø.Ø166 |
| 2. | Intercept of best straight line with x-axis. | Ø.2Ø5 |
| 3. | Fatness of lobe i.e. (mean square perpendicular distance from best straight line), normalised by the maximum excursion and number of points involved. | Ø.777 |
| 4. | x peak/y peak in respective spectral magnitudes. | Ø.728 |
| 5. | fpk, frequency of peak in y spectral magnitude. | 1.74 |
| 6. | Difference in probe between x and y spectra at fpk. | Ø.479 |
| 7, 8, and 9 | are the spectral areas for the y-signal, up to fpk, from fpk to 25% of fsampling, and above respectively. | 1.22Ø<br>Ø.888<br>Ø.444 |
| 1Ø. | Average radius (1st moment about signal origin). | 1.423 |
| 11. | Maximum radius. | 1.45 |
| 12. | Radius2 average over segment. | 1.271 |
| 13. | Segmented width. | Ø.876 |
| 14. | measure of segment symmetry about middle of segment i.e. moment of y sign about left hand segment end divided by moment of y signal about right hand segment end, where moment means signal amplitude at a point times that point's distance in terms of number of sample points from the "fulcrum". Coded as $f_{14} = \left[\sum_{i=1}^{n} i\,y(i)/n\right] \Big/ \left[\sum_{i=1}^{n} (n-i+1)\,\frac{y(n-i+1)}{n}\right]$ where i is sample number with respect to edge of segment, n is number of samples in segment and y(i) is the voltage measured at the ith sample. | Ø.Ø521 |

**Table 4.2(a)**: multiclass linear classifier results:

| | | assigned class | | | |
|---|---|---|---|---|---|
| | | ZCK | IGSCC | GCO | Misc. Rate |
| true class | ZCK | 88 | 29 | 2 | 26.1% |
| | IGSCC | 0 | 94 | 6 | 6.0% |
| | GCO | 3 | 23 | 79 | 24.8% |

overall misclassification rate = 24.8%

**Table 4.2(b)**: sequential 2 class linear classfier, digital filtering applied to scans as per section 4.3, and fig 4.5(b):

| | | assigned class | | | |
|---|---|---|---|---|---|
| | | ZCK | IGSCC | GCO | Misc. Rate |
| true class | ZCK | 84 | 30 | 5 | 29.4% |
| | IGSCC | 5 | 90 | 5 | 10.0% |
| | GCO | 14 | 6 | 95 | 9.5% |

overall 17.0%

**Table 4.2(c)**: Results of 2 stage 2 class linear classification with all of the new features included, except for defect width, (which, since it was calculated from the output of the segmentation algorithm, did not produce sensible results for the the GCO class):

| | | assigned class | | | |
|---|---|---|---|---|---|
| | | ZCK | IGSCC | GCO | Misc. Rate |
| true class | ZCK | 98 | 16 | 5 | 17.6% |
| | IGSCC | 2 | 86 | 2 | 4.5% |
| | GCO | 2 | 2 | 101 | 3.8% |

overall 6.3%

Note: The number of patterns in the IGSCC class has decreased from 100 to 90 because of the inclusion of 10 of the scans from the lowest amplitude IGSCC patch in tube CA54 from the testing set to the training set (see section 4.7).

**Table 4.2(d)**: Results for the 2 class linear classifier when using the segmentation algorithm, and the full set of 14 features (and therefore no GCO class), the best results were achieved:

| | | assigned class | | |
|---|---|---|---|---|
| | | ZCK | IGSCC | Misc. Rate |
| true class | ZCK | 70 | 3 | 4.1% |
| | IGSCC | 6 | 81 | 6.9% |

overall 5.6%

Note: a few scans were not passed by the segmentation algorithm, and consequently the row totals are reduced with respect to the other confusion matrices above.

**Table 4.3 Nearest mean Classifier Results**

| | | assigned class | | | |
|---|---|---|---|---|---|
| | | ZCK | IGSCC | GCO | Misc.Rate |
| true class | ZCK | 110 | 6 | 3 | 7.6% |
| | IGSCC | 21 | 68 | 11 | 32% |
| | GCO | 1 | 2 | 102 | 2.8% |
| | | | | overall = | 16.6% |

**K-Nearest Neighbour Results from 3 Class, 9 Dimensional Data**

for k=1

| | | known class | | |
|---|---|---|---|---|
| | | ZCK | IGC | GCO |
| assigned class | ZCK | 110 | 20 | 2 |
| | IGC | 6 | 73 | 5 |
| | GCO | 3 | 7 | 98 |
| % correct | | 92.4 | 73.0 | 93.3 |

correct classification rate over all classes=86.3
overall misclassification rate, ($E_1$)=13.7%

for k=3

| | | known class | | |
|---|---|---|---|---|
| | | ZCK | IGC | GCO |
| assigned class | ZCK | 108 | 15 | 3 |
| | IGC | 8 | 78 | 7 |
| | GCO | 3 | 4 | 101 |
| % correct | | 90.8 | 78.0 | 96.2 |

correct classification rate over all classes=89.3%
overall misclassification rate, ($E_3$)=10.7%

for k=5

| | | known class | | |
|---|---|---|---|---|
| | | ZCK | IGC | GCO |
| assigned class | ZCK | 105 | 15 | 0 |
| | IGC | 11 | 78 | 6 |
| | GCO | 3 | 7 | 99 |
| % correct | | 88.2 | 78.0 | 94.3 |

correct classification rate over all classes=86.8
overall misclassification rate, ($E_5$)=13.2%

Confusion tables for nearest neighbour classifiers.

**Table 4.4** Nearest Neighbour Classifier Results upon the original feature vector set.

TRAINING SET FILE: CP7.STA
CLASS 1 47 MEMBERS
CLASS 2 57 MEMBERS

TESTING SET FILE: CP9.STA
CLASS 1 73 MEMBERS
CLASS 2 87 MEMBERS

$E_1$

| KNOWN CLASS | ASSIGNED CLASS 1 | 2 |
|---|---|---|
| 1 | 71 | 2 |
| 2 | 3 | 84 |

PERCENTAGE CORRECTLY CLASSIFIED:
CLASS 1 97.26%
CLASS 2 96.55%
OVER ALL CLASSES: 96.91%
REJECTION RATE = 0.00%
MISCLASSIFICATION RATE = 3.09%

$E_{13}$

| KNOWN CLASS | ASSIGNED CLASS 1 | 2 |
|---|---|---|
| 1 | 73 | 0 |
| 2 | 6 | 81 |

PERCENTAGE CORRECTLY CLASSIFIED:
CLASS 1 100.00%
CLASS 2 93.10%
OVER ALL CLASSES: 96.55%
REJECTION RATE = 0.00%
MISCLASSIFICATION RATE = 3.45%

$E_{13,7}$

| KNOWN CLASS | ASSIGNED CLASS 1 | 2 |
|---|---|---|
| 1 | 72 | 0 |
| 2 | 4 | 77 |

PERCENTAGE CORRECTLY CLASSIFIED:
CLASS 1 98.63%
CLASS 2 95.06%
OVER ALL CLASSES: 96.57%
REJECTION RATE = 2.73%
MISCLASSIFICATION RATE = 3.70%

**Table 4.5** Improved Nearest Neighbour Classifier Results for ZCK (class 1) vs. IGR (class 2) using all 14 features and the segmentation algorithm.

# 5 Chapter 5: Multifrequency Inspection Work

## 5.1 Introduction

Work done to automate the interpretation of multifrequency loci is given in this chapter. Its organisation follows that of Chapter 3, except that here the experimental use of each of the sections of a statistical pattern recognition system is emphasised. The desirability of using multifrequency eddy current has already been explained in Chapter 2.

## 5.2 The Defect Set

In this section, the composition of the defect set is examined, and a number of points are raised concerning the mechanics of the logging of these defects.

### 5.2.1 The Composition of the Defect Set

Six types of defect in Inconel 600 have been recorded: axial and circumferential notches, uniform thinning, flat bottomed holes and intergranular attack (in both patches and in uniform bands). Later in the chapter, in the sections on classification, these are referred to by the 4 letter mnemonics Znch, Cnch, Thin, Hole, Piga and Uiga respectively. All of these were artificially manufactured by mechanical or chemical means and all were external wall defects. There were no internal bore defects, as a consequence of the collaborating bodies' interests, which lay in demonstrating the ability to

detect all defects of any consequence. Defects occuring in tube bores are regarded as easily detectable since 5% through wall cracks in 1mm thick inconel can be discerned from the background. However those of 20% through wall depth on the outside wall of tubing (even when it is not passing through a support plate) may go undetected. In any case, almost all naturally occuring defects grow from the outer sides of tubes inwards, because they tend to form initially in areas in tension on the outer radii of U-bends, or at the tube roll (see Fig.1.4), so omission of ID defects from the data set is equivalent to removing the most easily detected examples.

### 5.2.2 Manufacture of Defects

The axial notches were produced by electric discharge machining (EDM). The circumferential notches were cut. All notches were machined to depth, and then accurately measured by gauging the height of their imprints in a plastercast with a Talysurf stylus.

The uniform thinning was produced on a lathe by reducing the diameter of a short section of tubing and the thinned section was then gauged for accurate depth.

The intergranular attack (IGA) was manufactured by immersion of a tube in a boiling acid bath for a number of hours after it had been "sensitised" to attack by prolonged exposure to a temperature of 650°c, a temperature twice as high as the maximum that tubing would normally be operated at. The ends of the lengths of tubing were stoppered to confine the attack to the outer wall and most of the outer surface was masked off. Small areas of the outer

surface were left unprotected in order to cause patches or bands of IGA.

Two different acids were used, nitric acid and polythionic acid. The type of acid used had a considerable effect upon the type of IGA produced. The former attacked the grain boundaries more vigorously than the latter, and produced a layer of metal with correspondingly higher resistivity, due to the larger gap between the metal grains. Consequently, although IGA induced by polythionic acid gave a signal amplitude which was related to IGA depth, it was of a lower amplitude for a given depth compared with that for the nitric acid specimens. A report produced to accompany these specimens (**Foster,1985**) concluded that the IGA was representative of the two distinct varieties of IGA which occured in SG's under service conditions. Table 5.1 lists the defect sample set.

The tubing used was straight 15.5 mm bore diameter, Inconel 600 tubing with wall-thickness 1.1 mm, of the same type as is use in some currently operational Westinghouse type reactors.

### 5.2.3 Defect Set Sufficiency

Defect depth sizing during in service inspection of tubing is normally regarded as sufficiently accurate if defects can be reliably assigned to one of the depth zones 0-25%, 25-50%, 50-75% or 75-100% through the tube wall. For each defect class (ZNCH, CNCH, PIGA, UIGA and THIN) there were two examples from each depth regime. The Hole class, although being used to simulate pits, were actually angular registration and calibration marks (all of them were 1.3mm 100% through wall) and so no provision had been made for

producing them in different depths. Thus for this popularly accepted classification scheme, the defect depths available were sufficient at least to test for (if not quantitatively assess) any sizing capability possessed by any feature set, leaving aside any considerations of statistical validity due to the small size of the sample set.

### 5.2.4 Defect Set Size

As there were only two samples of each defect type, in each depth regime (including the differently induced types of IGA), each pair was split so that one went into the testing set and the other into the training set, whilst (as in chapter 4) taking care not to put scans from the same defect into both sets.

Unfortunately, not all of the defects manufactured could be used.

During the relogging of the Inconel defect set onto the PDP11/23, a 45 scan limit had to be imposed on each tube file. This made it necessary to locate the defect in the tape recording using an oscilloscope before logging could begin. The tube samples were typically 15 cm. long, and 45 scans represented only 4.5cm.

It was impossible to log defects not visible against the more slowly varying Pilger noise background. In practice these were always CNCH and IGA defects of particularly low amplitude, because the gently undulating form of their cyclic variation was so similar to that of the Pilger noise.

When the tubes were first logged to tape, a chart-recording Medelec

oscilloscope was used to monitor the scans. It produced a plot very similar to the tubeplots (see fig. 5.37(c)). However it was not possible to use the Medelec plots to deduce after how many scans the logging program should be started to capture the defect because the start of the Medelec plots could not be synchronised with the tape recordings (because the paper drive took a small but random number of scans to accelerate up to speed). The problem might have been overcome by reviewing all of the files pertaining to each tube immediately after the tube was logged to disk had it not been for certain practical constraints: plotting was time consuming and uninterrupted access to the Racal tape recorder would have been required because there was not enough disk space to accomodate all of the sections of all of of the tubes at once. Since two such tape recorders were shared between all of the departments at Babcock's Research Centre, exclusive usage in this manner was not possible, especially since the tape recorders spent a lot of time out on site work.

The single largest deficiency in the defect set was the lack of enough defects because there were no more than two examples of any one depth of a given type. Attempts to overcome this were unsuccessful because of the cost of the production of more defects. e.g. the cost of producing the IGA samples to Babcocks was about £3000 for 32 tubes. Economic factors prevented the production of a sample set which was large enough to satisfy eqn. 4.13 whilst treating each defect as a pattern. Consequently, the same strategy was used as in Chapter 4: each scan through a defect was treated as an independent sample, even though this was a slightly dubious procedure from the point of view of statistical independance of the samples.

It is virtually impossible to manufacture OD cracks of controlled depth in tubing without deforming the tube to the extent where it could not be scanned with an internal probe, and no real cracks were available for logging since there are no Inconel-based steam generators in this country as yet. Consequently production of a real crack sample set was not possible. The effect of a small sample set is discussed further in section 5.8 upon classification results.

### 5.3 Signals Associated with Defects.

Computerised simulation of eddy currents around defects, and the resultant effects upon probe coil impedance, is still at too primitive a stage to allow prediction of what should be the salient features for classification of different types (or depths) of defects (see section 2.9). For the design of an adhoc featureset, prior to testing of the featureset with feature selection and classifiers, it was desirable to try to find out what features are important by inspection of the signal time series and impedance plane loci from defects typical of those to be classified. In the next section, the plots obtained from the defect set are reviewed for precisely this purpose.

There is some variability in the scan length and plot quality in the plots at the end of this chapter, due to the use of three different computers in the course of the work for this chapter. The background to this is explained in the appendix. However, plots of similarly oriented scans but of different resolution are otherwise unaltered (compare, for example figs. 5.23 and 5.43)

Plots 5.1 to 5.21 and 5.27 to 5.36 were done using University of Glasgow's

PDP11/45 with a Tektronics 4662 medium resolution plotter. The very high definition tube plots (5.22-5.26) had to be done on a Benson 1302 plotter connected to the departmental GEC4180 computer. This required quite a lot of effort in rewriting of the plotting programs and the new data file handling routines, since the GEC4180 file transfer utility (FTP) would not allow binary data files - only ascii formatted data files. Later plots (figs. 5.37 onwards) were done using a lower definition screen dump from a DEC VT220 terminal specially converted for the purpose, by the installation of a Selenar SG220 Tektronics 4014 emulator module. The data files were relogged as part of the process of transferring the experimental work from the University's PDP11/45 and GEC4180 computers to Babcock's PDP11/23 and at that point were modified to have a lower resolution of 55 points/scan (see appendix). The first set of Inconel 600 defects data files had about 390 points/scan.

Of particular interest are the shapes and relative sizes of the waveforms and impedance plane loci especially within channel 2 (220kHz), and how these vary with defect depth as the frequency of excitation is altered. Channel 2 was of greatest interest because it was the channel with a phase of 90° between ID and OD effects, and also the one in which the loci appeared to be most distinctly different between the classes.

A large number of typical plots have been included at the end of this chapter. There is a plot from each program for each defect type except where the resultant plot is not helpful e.g. a circumferential scan plot of a banded (circumferentially symmetric) defect obviously does not show a lot other than a d.c. level. The relationship between plots and how for instance the same scan

appears in figs 5.2 and 5.14 for the purpose of comparison is perhaps best explained by Table 5.4. In all plots except for those done with the 8 channel tube surface plotting program (figs. 5.22-5.26), individual channel gains are taken into account. In those plots, it was important to display each channel as recorded, rather than as in all of the other plots, scaled with respect to the largest peak due to the defect in any channel. So the channels are shown as recorded, with the vertical scale representing a.d.c. bits, and consequently the actual signals gains are not necessarily equal across the channels.

The horizontal scale in each plot is the number of samples, there being approximately 390 samples per rotation. Due to the improvement in scanning speed stability compared with the circograph used in chapter 4, it was found to be unneccessary to pad lines out to achieve uniform lengths.

Excitation frequencies were 100 kHz, 220 kHz, 300 kHz and 600 kHz, for channels 1, 2, 3 and 4 respectively. The corresponding standard skin depths in Inconel 600 were 1.6, 1.0, 0.8 and 0.6 mm The through wall depths quoted were gauged from the outer wall inwards. The tube wall was 1.1 mm.

**Helical Scan Plots**

Figures 5.1-5.6 were plotted with program SCN4 (a mnemonic for "4 frequency SCaN plot") which plots all eight channels for one scan around the entire circumference of a tube. In these plots the x and y signals represent the horizontal and vertical components of the impedance plane trajectory.

Each scan is a fairly severe example in terms of depth of penetration and consequent signal amplitude in order to prevent probe wobble noise obscuring the salient signal features. The length of the defect in a scan time series

appears to be a discriminatory feature.

The defects shown are as follows:

**Figure 5.1** intergranular attack, patch type (PIGA), 60% through wall, scan number 100, tube 346.

**Figure 5.2** circumferential notch (CNCH), 80% through wall, scan 52, tube 15c

**Figure 5.3** axial notch (ZNCH), 78% through wall, scan 56, tube 5F

**Figure 5.4** flat bottomed hole (HOLE), 100% through wall, scan 10, tube 342

**Figure 5.5** clean tubing (GOOD), scan 10, tube 342

**Figure 5.6** uniform thinning (THIN), 48% through wall, scan 65, tube 18F

Only one banded defect is shown, thinning (fig. 5.6), since this type of defect is not really effectively displayed by a scan around the band, but the other type, uniform (or banded) IGA, is very similar to uniform thinning, differing only in the degree of d.c. shift when displayed in this form.

**Axial Scan Plots**

These plots (figures 5.7-5.12) were plotted using program AXSCN4, which is similar to SCN4 except that the time series now represents a path along and not around a tube. The helix pitch was 1 mm/revolution, and the tubes were 15 cm long.

**Figure 5.7** IGA (patch type)

**Figure 5.8** CNCH

**Figure 5.9** ZNCH

**Figure 5.10** HOLE

**Figure 5.11** GOOD

**Figure 5.12** THIN

All axial scans were arranged to pass through the centre of the defect. Had they been to one side of the defect, the signal would merely have been reduced in size, except in the case of the hole type defect where the signal would have changed in shape from one having two peaks, to one having only one (see fig 2.7).

**Plots of Multifrequency Single Axial Scans in the Impedance Plane**

The scales in these plots are calibrated in terms of a.d.c. units, 128 units representing 1 volt. In some of the figures, the scales extend to 1000 units for maximum deflection, because the attenuation introduced by the range settings on the Racal tape recorder and the Zetec gain switches have been taken into account. The existence of manual prescaling of the signals to use all of the tape recorder's dynamic range meant that an 8 bit representation was adequate. For the single scan, 4 channel plots, the signal origin is indicated by the cross at the centre of each quadrant. The quadrants contain the 100, 220, 300 and 600 kHz loci respectively as indicated in fig. 5.12. The same convention is adhered to in successive figures.

**Figure 5.12** An axial notch (ZNCH,) with and without d.c. component (corresponding to fig. 5.3)

**Figure 5.13** Patch-type (rather than banded-type) intergranular attack IGA ($HNO_3$ induced), the 2nd of which corresponds to fig. 5.1

**Figure 5.14** Circumnferential notches CNCH, (see fig.5.2)

**Figure 5.15** ZNCH, (see fig. 5.3)

**Figure 5.16** A hole (see fig. 5.4)

The offset of each of the loci from the origin visible in fig. 5.12 were

considered as features and included in the impedance plane featureset in section 5.6, but they were found not to be strongly significant (section 5.6).

In each of the remaining plots, (5.13-5.16), this offset has been removed. They show 3 or 4 defects of the same type, of increasing depth, in order to illustrate that the similarity in the shape of the loci, but differences in the overall sizes and angles.

### 5.4 Scan Spectra

The spectra in figs. 5.27-.33 were generated by an FFT routine, acting upon a zero padded 256 sample array, which contained the defect signal. The program operated by taking the FFT's of the x & y signals separately, and then graphing their individual complex amplitude magnitudes and combined phase differences. The time series were pre-windowed by a 4th order Blackman-Harris (raised Cosine) window (**Harris**,1978), and post-smoothed by a 5-point moving average window.

The vertical scales in the plots are linear ones, representing spectral energy or degrees as appropriate. The horizontal scale represents only the first 12 ordinates in the frequency domain plus the d.c. component. The upper 116 points in the 128 point transform results were always of negligible magnitude, and the accompanying violently varying phase above the $12^{th}$ ordinate did not appear to be at all related to the type or depth of defect involved.

## 5.5 Characteristics and Effects of Pilger Noise

In the plots of individual helical scans (figs. 5.1 to 5.21) and of complete tubes (figs. 5.21 to 5.26) a cyclic variation in the background can be seen. This is due to a combination of the Pilger effect and lift-off, the latter of which was due to some non-alignment of the axis of probe rotation with that of the tube, an effect that is difficult to avoid. This slow rippling of the background at a frequency very close to 1 cycle/scan makes segmentation of defect signals more difficult than it was for the 316 tubing in Chapter 4 where it did not occur because a differential probe was used.

An account of the results of various segmentation methods follows in section 5.6, but first the effects of Pilger noise must be defined because they must be cancelled by any successful preprocessing or segmentation stage.

In the impedance plane plots, only a narrow segment from a scan has been plotted in order to show only the defect and not the background. In some of these, the locus does not end at its starting point (see for example, fig 5.20). Also, many loci from defects which would be expected to have approximately identical outward and return trajectories, enclose a readily visible area. Both of these effects were seen to happen regardless of the speed of rotation of the probe, indicating that the effect were not due to high speed deformation of the magnetic field associated with the probe, as was first suspected. Had this been true, this would have accounted for the asymmetric loci which were being experienced, even for symmetrical V-profile spark notches but the probe tip speeds at which this would normally be expected are 20 times higher (Morgan,1985). It was suggested that removal of the Pilger noise would help segmentation by simplifying the

segmentation operation to one which tested the signal to determine if it had gone outside a ring around the probe balance point. Classification might be helped too, since the background lends a spurious thickness to the figures, distorting their shapes and perhaps confusing the difference between classes.

The rippling background of figs 5.22-5.26 tended to be more obvious at the highest excitation frequency since it was an internal diameter effect where the probe sensitivity would be biased toward inner surface phenomena due to the skin effect. In tubes that had been recorded with no attenuation in order to study a shallow defect, e.g. fig. 5.22, the wavy cyclic background was most severe. Since it was normal practice to reduce the display gain by a factor of ten when recording the Y2 channel for 100% through wall notches and other similarly strong signal producers, the background for such defects appears constant due to the loss of the Pilger signal between quantisation levels. Pilger noise is referred to very little by other authors, possibly because it is a more severe problem for rotating absolute probes than the more common differential bobbin probes. **Van Drunen and Sharp** (1983) reported a similar type of rippling waveform due to localised oxidation in Zirconium alloy Candu tubing. Pilger noise is a less serious problem in the UK than in the USA and Canada because the specification against which heat exchanger and steam generator tubing is tested places a limit on the amount of Pilger noise allowed.

## 5.6 Segmentation

The aims of the work in this section were twofold: to segment the defect signals from the background, and also to cancel out the additive background

signal, thus leaving only the signal due to the defect itself for presentation to the next stage, the feature extractor.

### 5.6.1 Identification of Pilger Noise

From an inspection of a large number of impedence plane plots it was concluded that the Pilger noise loci always lay along the same line in the impedance plane, assuming that the phase rotators of the Zetec were all set to zero (which they were in all cases). The angles at which the pilger noise loci lie are shown in fig. 5.38(a), along with the angles relative to an OD defect at each of the four inspection frequencies.

The angular relationship between the Pilger noise in channels 2 and 4 is (by inspection of the plots of the tubing set) always the same, even although the Zetec was balanced for each tube immediately prior to logging. This signified repeatability of the bridge balancing operation and that the conductivity and permeability of the tubes did not vary, even for those that had been subject to prolonged heating while IGA was induced in them.

Occasionally, as in fig. 5.38(b), the Pilger locus can be fatter than it is long, with the minor axis apparently lying along the characteristic Pilger direction. This fattening of the Pilger locus seemed to happen whenever probe eccentricity was particularly bad, but it was not possible to be certain of this since there was no independant means available of measuring probe/tube axis eccentricity, or profile variations in the internal and external tube diameter.

The importance of considering cases of unusually fat Pilger loci lies in the necessity to ensure that any scheme to remove Pilger noise and/or segment

out defects will continue to work even under exceptional circumstances. The final method, apart from being tested upon normal loci (see fig. 5.38(a) and 5.41), was also tested with the locus shown in fig. 5.38(b), and worked acceptably providing that the mix had been developed upon normal Pilger noise (i.e. noise with a strongly linear shape).

### 5.6.2 Rotation of Loci with Frequency

All defects in the tube set were surface breaking at the outer tube wall. Since Pilger noise is essentially an inner wall phenomena, there is a phase difference between Pilger noise and OD defect loci, and the amount of the phase separation increases with frequency, as subsurface OD currents lag increasingly with respect to their ID counterparts.

As expected, it was observed that the OD defect loci rotated with frequency (i.e. with increasing phase lag of the OD currents). However it was also observed that the Pilger loci rotated. This is not as predicted for a purely surface based phenomenon such as Pilger noise, since the surface currents always have the same phase angle with respect to the driving sinusoid. Thus the combination of the probe bridge and Zetec demodulation circuitry does not preclude the possiblity of a frequency dependence of the orientation of the OD/ID axes themselves. The automatic mixing algorithm was designed such that this had no impact.

During the first 5mm. of the tube, the background level differed markedly from that for the tube away from the end, due to distortion of the normally circular eddy currents by the proximity of the nonconducting air

metal interface.Any d.c. bias in the signal was negated before the use of a segmentation criterion, by subtracting the value of the first sample in any scan from all of the other samples in that scan. This procedure prevented tube end effects from erroneously triggering the segmentation mechanism. Where defects can be guaranteed not to fall across the transition from one scan to another, as here, this is an acceptable strategy. Obviously, this constraint could not be tolerated outside a laboratory. It also served to counteract an intermittent problem with a leaking capacitor in the Zetec bridge balancing circuitry, which would cause a considerable drift in background level during the time taken to log a 15cm. length of tube (see for example fig. 5.10)

### 5.6.3 Segmentation Experiments

All of the following methods assume the availability of clean tubing from which to capture an impedance plane locus due to Pilger noise alone. The methods tested by experiment were

(i) **Ringing the Signal Origin with a circle** which just enclosed the Pilger noise from the prototype scan and then treating any signal excursion outside of it as a defect. This method failed to detect defects in tubes where the defect peak amplitude was less than that of the Pilger noise.

(ii) **"gift wrapping" the Pilger noise locus, by forming its convex hull in the impedence plane. Akl & Toussaint's** (1978) convex hull algorithm produces a list of all of the external points in a polygon (in this case the one described by the Pilger noise), in clockwise order. This is analogous to wrapping all of the

points in the locus with a sheet of paper in 2 dimensions, and discarding all of the points which do not come in to contact with the paper. **Akl & Toussaint's** algorithm operates by

**(i)** finding the 4 points with maximum and minimum x and y values,

**(ii)** for the 12 o'clock to 3 o'clock quadrant, sorting the point set with x greater than the x value of the point with max y, into ascending order by value of x ,

**(iii)** find the most convex point with respect to the 12 o'clock and 3 o'clock positions (i.e. the one with max(x+y) and then

**(iv)** discard the points as within according to the sign of the vector cross product of their position vectors with respect to adjacent points, of which the least clockwise of the external points is the first.

The method was picked because it is guaranteed to work for all sets of points in 2 dimensions, unlike some others which they cite. In this application, all of the points in the resultant polygon were dilated with respect to the centre of gravity of the convex hull, to allow a safety factor before segmentation was triggered. Detection of any excursion outside the hull was also done by reference to the sign of the vector cross product of each points' position vectors to the two nearest points in the convex hull, which must be taken in clockwise order in the multiplication. When a point is within the hull, the sign of the vector cross product will always be positive.

A convex hull was thought to be superior to an elliptical threshold because it described the allowable locus of impedances more accurately. This method worked and was effective in segmenting defects with associated signals smaller in magnitude than the Pilger noise, but it is more complex and

slower than the equally effective method (vii) which uses a sloping line fitted to the Pilger noise.

It does however yield the possibility of an automatically calculable segmentation boundary of arbitrary rather than merely just convex shape around a number of unwanted signal loci by modifying the algorithm to treat a point set in regions, by advancing along the convex hull, treating only points within a radius of limited size. The exact value of this radius would effectively control the tighness of the fit of the hull to the detailed shape of the signals.

(iii) **high pass filtering, followed by simple amplitude thresholding.** In theory, it ought to have been possible to separate non-banded defects from Pilger noise with a high-order, high pass filter, with a cut-off frequency of 1 cycle/scan. However the filter was divergent after several scans. This may have been a result of storing the previous filter outputs as byte length variables which was unavoidable (even with overlaying) due to memory constraints.

(iv) **simulation of a differential probe.** It was found that the signal from a differential probe could be synthesised (see fig. 5.34) by taking the moving difference between the present sample and a previous one which was equivalent in lag to the "rise time" of the probe. The comparable undifferentiated scans are plotted in figs 5.13,14 & 16. In order to use this technique for segmentation it had to be followed by code to detect the variation of either the differentiated x or y signals away from zero, and then "fill in" the gap between the extremities of any defect within one scan, since it is only the extremities which register with a differential probe. Part of the algorithm

developed in chapter 4 (the part which linked the positive and negative going swings at the starts and ends of defects) was used to do this. The differentiation approach required some rather involved code to re-orient scans in order to work for both banded and non-banded defects. It worked as well as the mixing approach (method(vi)), which is as expected, since both methods eliminate Pilger noise.

(v)**subtraction of wobble** It initially appeared that the slow advance of the cyclic Pilger noise might be sufficiently regular to allow it to be tracked and predicted and therefore subtracted from each scan, leaving any defect on a zero level datum: the Pilger loci was seen not to change in the presence of defects, i.e. the defect signal could be treated as a "bulge", merely being added to one segment of the scan time series. It was determined from inspection of several tube plots that the period of Pilger noise is 52 scans, i.e. the peak advances around the bore of the tube in 5.2 cm. One typical Pilger scan was picked as a prototype, and computationally advanced at the rate of $1/52^{nd}$ revolution per scan, and subtracted from each successive scan in turn. The results are shown in fig. 5.39. Close inspection of consecutive scans in several tubes showed however that Pilger noise is less simple than it initially appears (fig. 5.40) and cannot be described as a cyclic function which can be predicted by rotating the present scan in a recirculating buffer by a small fraction of one cycle. Also, the shape of the peak can change quite markedly from one scan to the next. Simple approaches like tracking and extrapolating the position of the next peak are therefore precluded. With the variations in wave shape and the nonconstant increments in peak position from one scan to the next, the effort involved in evolving an adaptive curve fitting strategy (i.e. one where both the shape of subtracted Pilger noise scan to be subtracted

and the degree of its advancement are varied) would have been prohibitive, especially given that a simpler method was developed, as described next.

If the Pilger noise loci had remained constant in size and it had been found possible to track the advance of the signal (as mentioned above), it would have been possible to store a scan of purely Pilger noise and remove it from each scan in turn. Considerable advantages ought to have resulted from this method in that defect loci to be classified would have had the Pilger noise element removed from them, thus elimanating a source of spurious variation between loci from the same class of defect, or even from different scans through the same defect.

(vi) **by Computationally Developed Mix** The subroutines which did this (ANNUL and MIX) operated as follows. ANNUL was run on a nominated scan containing only Pilger noise, in order to calculate the rotation, and horizontal and vertical magnification which required to be applied to channel 4, in order to match (and hopefully cancel) the Pilger noise loci in channels $X_2$ & $Y_2$. It did this by

**(a)** finding the centre of gravity of Pilger noise in both channels 2(220kHz) and 4(600kHz), individually, by calculation of the x & y. Then

**(b)** rotating both loci until the position vectors to their respective centres of gravity point along the +ve x-axis.

**(c)** calculating the standard deviations along the major (x) and minor (y) axes for both ch2 & 4.

**(d)** calculating scaling factors to match the standard deviations between ch2

and ch4.

Subsequently, subroutine MIX used the two angles found in step(ii), and the scaling factors from step(iv), to generate the same mix as was used by ANNUL to elimanate the Pilger locus.

Mixing normally distorts the characteristic shapes expected from the various different types of defects which exist. In practice, distortion of the defect locus was not troublesome. Certainly, the mixed loci were of a different shape compared to their unmixed parent loci, but the effect was not objectionable. If anything the mixed defect signals were closer to the idealised comma shapes expected for cracks than the unmixed original loci (see fig. 5.41, mixed locus). This is because of the removal of the Pilger noise which tends to fatten the locus. The previous method would also have removed this, had it worked. Other authors have published loci of defects, which because of mixing took on a much different shape and phase angle, making interpretation of the loci difficult because all of the characteristic shapes and angles had changed.

Initially, the mixing annulment subroutine used a line fitting subroutine to find the gradient of the locus, for calculation of the rotation required to map the locus on to the +ve x-axis. However, this method failed whenever the Pilger loci was not linear in form, because the line was always fitted to the major axis, which might not necessarily lie along the characteristic Pilger direction. Furthermore, the intercept under such conditions might not be zero. This caused the $ch_2$ and $ch_4$ loci not to superimpose well after rotation and scaling. Furthermore, there were problems in determining which quadrant the loci predominated because the linear regression algorithm being

used was only capable of returning the slope, making it awkward to calculate $\phi$, the rotation to be administered, when using the standard function $\tan^{-1}$ available in Fortran. Subsequently, the centre of gravity was used to quantify the characteristic angle of the locus, and surprisingly, proved quite adequate.

(vii) linear threshold fitted to the Pilger locus, using the unmixed channel 2: A line was fitted to the Pilger noise locus using a linear regresssion algorithm. The threshold boundary was calculated from the equation of this line, by adding on an amount to the intercept equal to 2 standard deviations of the Pilger noise about the fitted line. The method worked acceptably on the samples available. It was more sensitive to small defects in the presence of large Pilger noise, than for instance, the null point ringing method in (i), because of the differing sensitivities in the OD and ID directions. It was quicker in operation than the convex hull method because the segmentation operation consisted of testing each point to determine if $y \geq mx+c$, rather than a large number of vector cross product operations. However, the scheme would have to be amended to include another parallel threshold plus two more, perpendicular to the first, to be capable of detecting ID defects and other phenomena.

### 5.6.4 Choice of Prototype Scan

It was found to be essential that the scan picked for a calculation of the mix or for the training of a linear threshold as in method (vii) contain strong Pilger noise, which will be guaranteed to have a strongly linear loci. This minimises any error in the determination of the Pilger directions in ch2 & 4. Provided

that this was so, a mix calculated upon one tube (scan 4, tub14c.raw tube data file) was found to be applicable to other tubes, even although the bridge in the eddy current equipment was nulled separately on each tube using the automatic balancing facility. Having found such a robust mix, investigations with more complicated optimisations by functional minimisation were thought unnecessary.

## 5.7 Feature Extraction

Within the time available after the multifrequency defect set was logged, it was only possible to test an ad hoc featureset (basically because of the problems encountered with the systematic feature extraction methods discussed in detail in chapter 3). The design of the ad hoc featureset is now justified.

The loci in all four channels were very similar except in size and in angle and since this was true regardless of the type of defect, the features were drawn essentially from just one channel: channel 2, the channel in which the defect signals appeared to be most different. The frequency was picked by the sponsors for greatest sensitivity to IGA. The information obtained from the other channels is encoded in additional features which describe the relationship between the loci in these other three channels and channel 2.

### 5.7.1 Ad Hoc Impedance Plane Features

The impedance plane featureset basically consisted of

1. the overall angle ( see feature 1 in table 5.1)

2. the size of the locus,
3. the area enclosed
4. the concentration of the area towards the tip,
5. towards the origin
6. towards the middle

Features were also included to test for any significance in the angle and magnitude of the d.c. bias in the locus signal origin (features 7 & 8).

To quantify the shape of locus, the outward trajectory from the null point to the locus extrema was described by the 6 coefficients of a $5^{th}$ order polynomial in x fitted to it (features 9-15), after the locus extrema had been aligned with the x- axis to ease the operation of the curve fitting algorithm.

It might have been preferable to have used a line which was the average of the outward and inward signal paths, but this would have taken too long to program since it is actaully a rather difficult task to perform due to the mismatch present in sample spacings along the two paths. Furthermore, the two paths did not differ greatly in comparison to the overall size of the locus (it was not certain until the mixing algorithm described in sections 2.11 and 5.4 had been developed that the difference in the paths was due solely to Pilger noise).

The $16^{th}$ feature, a boolean one, indicated if the outward locus crossed the inward one: a characteristic seen in all of the thinning loci. Another feature based upon the $16^{th}$ one was at one point considered, to detect whether the locus was described in a clockwise or anticlockwise manner. It worked by determining which path, inbound or outbound, was uppermost after the locus

had been aligned with the positive going x-axis. However, this did not seem to be important: a set of loci were tested with the measure and they were all described clockwise. This could not have bben determined by eye becasue the plotter went too fast and inbound and outbound paths were too similar for the direction of rotation to be visible in most cases).

A $5^{th}$ order polynomial was chosen so that the s-shaped loci characteristic of thinning could be adequately described. In fact such a curve is capable of being adequately modelled with a $3^{rd}$ order curve, but some extra latitude was included in the maximum order so that the curve at the outer end of the locus would be described too. The loci from the other classes of defect were all comma shaped, which ought to be capable of description with a $2^{nd}$ order curve. It such cases, the higher order terms ought to be zero, and indeed they were when the curve fitting routine was tested upon the outward trajectory of simple synthesised figures (Joukowski aerofoils).

### 5.7.2 Ad Hoc Spectral Features

The spectral featureset is given in table 5.2. It follows the impedance plane featureset in that most of the features are drawn from channel 2, with the remainder attempting to describe the bulk differences between the spectra from the other channels and that from channel 2.

Some difficulty was experienced in interpreting the spectra in plots 5.27-.32, in order to devise features. The differences in the relative scales of the x and y spectral magnitude plots within each channel, which were unavoidable if the shape of the distribution was to be easily visible in all of the plots,

made featureset design particularly difficult. In order to present the information more graphically, and simplify the assimilation of so much information, another form of displaying the spectra was sought. A new program (SPCROT) was written which displayed each channel pair's spectrum as a "corkscrew" view, i.e. end on from the low frequency end of the spectrum (fig. 5.33).

The magnitude and phase of the d.c. ordinate seemed to be related to whether a defect was an axial notch or not. In the SPCROT plots this was almost invariably the point of maximum magnitude. Feature 5 was designed to indicate a peak above the $5^{th}$ ordinate, feature 6 was thought to be a significant difference between ZNCH and HOLE classes, and feature 7 was included to give a measure of in which spectra the spectral energy was concentrated.

The spectral phase seemed to vary only slightly in the lowest parts of the spectrum, i.e. in those parts where the power was concentrated. Only when spectral power was negligible, did appreciable phase shifts occur.

### 5.7.3 Synthesis of Axial Scans

The feature extractor used to generate both the impedance plane and spectral featuresets had the ability to operate upon an axial scan for those classes of defects where it was appropriate. This was done by a preprocessing stage within the feature extractor program which dumped one point from each circumferential scan into a buffer which at the end of which the resultant synthesised input signal array was in effect a synthesis of an unrotated

axial scan. The same feature was added to the axial scan plotting program to produce the synthesised bobbin probe plots in figs. 5.17-.21. Feature extraction in this pseudo-bobbin probe mode was very slow, due to the necessity to read through the entire file for every pattern added to the data set, so only every 10th possible axial scan through each defect was actually used.

### 5.7.4 Practical Aspects

The feature extractor was tested for correct operation by using the features associated with the curve fitting algorithm to reconstruct the outward going trajectory of a test locus from a real defect. The part of the feature extractor which produced the spectral features was tested in a similar manner, except that the x and y time series were reconstructed separately, using the inverse transform. The reconstruction was perfect when the windowing and smoothing was removed and the full set of complex spectral amplitudes was used as the input series. Features not capable of being tested by reconstruction of a defect locus were not tested as rigorously but were tested for correct operation on simple shapes where the expected value of the feature could be checked by hand calculation.

The feature extractor was run twice for each class of defects, once to produce the training set of feature vectors, and once for the testing set. On each occasion it was done in batch mode using a data driver file because this made concatenation of the feature files easier when classes were being grouped together for separability tests (e.g. for banded vs. non-banded defects as is done in the next section), and also because of the time taken for each run was long enough for there to be a reasonable probability of a machine crash (due

to the age of the computer) and the results for that run lost. At the time at which the feature extractors runs were done, the PDP11/45 was being decommisioned and the maintenance contract on the PDP11/45 had expired and was not being renewed, and machine crashes had become frequent. Typically the feature extractor took 3 hours overnight (i.e. probably with no other users logged on) to process one class.

The process of determining the axial and circumferential extent of the defects within the tube data files was performed by hand, because the segmentation techniques desribed in the previous section had not yet been developed. Hand segmentation was one of the most time consuming jobs in the project, especially because it had to be performed again when the tubes were re-logged on Babcock's PDP11/23 system. The hand segmentation task resulted in the data driver files for the feature extractor given in table 5.5. These scan number references are specific to the tube data files which were logged on the PDP 11/45.

### 5.7.5 Feature Selection and Transformation

In most applications of pattern recognition, the next operation after the extraction of a big featureset is to select a subset using some measure of the class separation along each axis, culminating in selecting the best subset of features (see section 3.4 on feature selection). Most pattern recognition text books explain why considering features individually is incorrect but it is frequently done out of necessity anyway because searching algorithms such as the Branch and Bound method were only developed relatively recently, and are difficult to code, whilst exhaustive testing of all of the possible feature

subsets takes an impossibly long time. Consequently, it is not uncommon to see the Mahalanobis distance calculated for each axis individually being used as the criterion for the elimanation of features.

This is not a desirable way to proceed, so some development work was performed (see section 3.4) to produce a Branch and Bound– based feature selection method with a criterion function powerful enough to deal with the awkwardly distributed data set available.

Consequently, tests were run upon all of the features individually, for interfeature correlation, Fisher discriminant powers and rankings within Karhunen Loeve transformation matrices. The results of these were used to reduce the dimensionality of the feature vectors to an acceptable size (20) for classification purposes on the limited-size sample set. The features used ("*" in Tables 5.1 & 5.2) were those which were not intercorrelated by 95% or more, with the highest individual Fisher discriminant powers. These were in most cases about 0.5.

Figure 5.35 shows a plot of the 2-space ranked uppermost by the Kittler-Young transformation, which was typical of several done. The scans in each defect class fragmented into clusters which corresponded to the defects from which they had been taken (which differed in depth). Features 1 and 2 ( $\phi$ at Rmax and Rmax) were initially thought to be responsible because they are primarily related to defect depth and volume respectively. It was reasoned that if these two features were related only to depth and volume, then for certain combinations of classes, the class separation would be along other axes, and that if a feature axis transformation was tested upon the data, then these two features would not be close to the directions giving greatest separation. As a

result, a Kittler-Young Karhunen-Loeve transformation (see fig 5.35) was produced for each class taken against the remaining classes. Surprisingly, even when the subspaces corresponding to the highest ranked eigenvalues were plotted out, the classes were still fragmented according to the defects from which the scans had come. A nonlinear mapping of a subset of the data using Sammon's method is shown in fig 5.36 for comparison. This further confirms the lack of a coherent class structure suggested by the linear transformations.

### 5.8 Classification Results

Both nearest neighbour and multiclass linear classifers were used on the data, and the results are given in the table 5.3. The linear classifier results are a little pessimistic given the multimodality of the class structures determined in the last section. The nearest neighbour results were all done with a 1-NN rule classifier in order to observe the rule of thumb that k in a k-NN classifier should be about $\sqrt{n}$, where n is the class size. Since n here is really the cluster size, and some clusters had only 4 or 5 scans in them, k=1 was thought to be the largest that k could validly be. This means that the classifier results are perhaps as much as twice as bad as a (ideal) Bayesian classifer's would be. This is further backed up by the fact that the features in general had low individual Fisher discriminant powers, and the Fisher discriminant ratio calculated over all of the features between class pairs was typically about 0.5 and such a ratio would have led one to expect a misclassification rate of about 30% between two classes, because in one dimension, two normally distributed classes (with equal priori probablility of occuring and equal standard deviation) when separated by only 1 standard

deviation (equivalent to a Fisher Discriminant ratio of 0.5) overlap by 30.8% of their cumulative p.d.f.'s (from standard statistical tables of the cumulative normal distribution function)

The results obtained with the classification techniques tested upon the data certainly were not as bad as this, although they were not nearly as good as the results obtained with the signals from the single frequency equipment (which supposedly conveyed less information about the defect because there was only one channel). However, there were fewer defect classes in that instance.

The nearest neighbour classifier was run upon several different groupings of the classes in the defect set (see Table 5.3), in order to determine if the featureset would allow discrimination between axially symmetric phenomena (i.e. good tubing, uniform IGA and thinning), or between non-axisymmetric defects (CNCH's, ZNCH's, HOLE's and patch-type IGA). Using a 1-NN classifier, grouping axially symmetric defects (good tubing, thinning and uniform IGA) against non axially symmetric defects (ZNCH, HOLES, PIGA, CNCH), a misclassification rate of only 9.3% was attained. Separation of the various types of non-axially symmetric defects from each other was less successful (60% correct classification).

The best rate achieved for one stage discrimination between the 7 classes mentioned above (6 defect classes and good tubing) was 61% correct classification. This may appear poor, but it is not all that bad considering the inadequate defect set available, especially taking into account that within most of the classes, more often than not, those scans were correctly classified i.e. the leading diagonal in the confusion matrix is still more heavily

populated than the rest of the matrix put together. The correct classification rate never fell below 54% with the exception of the class containing patches of IGA in table 5.3(a). Considering that there were seven classes, and that six of these had a multimodal class structure, the results are really surprisingly good.

The decision that the classes were multimodal and that each mode represented the scans from one tube was made after careful scrutiny of fig. 5.35, and 6 other similar plots not shown, which bore out essentially the same result. Each one was produced by running the Kittler Young transformation for that class versus the union of all of the others, and then plotting the two dimensions ranked highest by their eigenvalues, in the hope that there would be a clear boundary between that class and all of the others. In order to determine conclusively that each cluster of points within the class in question was indeed from one tube, a line by line check by hand was performed, between the original feature vector file (which contained a scan source identifier for every vector) and a scatter plot with all of the points individually numbered specially for that purpose.

When it became apparent that within each class, the feature vectors scans were grouping according to the tube from which they were drawn (i.e. from one defect of a particular depth), it had to be determined whether or not the only difference between defects of the same type lay in the size of the locus (characterised by the maximum radius $R_{max}$) and in the angle of the major axis with respect to the horizontal ( $\phi$ at $R_{max}$). If this were true, then the clusters representing each defect would separate only along features axes "$R_{max}$" and "$\phi$ at $R_{max}$". This was determined not to be so: a conclusion

arrived at due to the following observations.

All of the features were normalised with respect to $R_{max}$ and for the locus geometrical featureset, the locus was rotated to remove the effect of $\phi$ prior to feature extraction, in order to prevent R or $\phi$ from biasing the other feature values because of the volume or depth of the defect. The scatter plots of which fig. 5.35 is typical, have axes which are made up of a linear weighted sum of the original feature axes. By inspection of the weightings of the original features in the new axes, it was clear that these axes were not predominant in the composition of the transformed feature axes, so the splitting of each of the classes into a cluster per defect was not confined to only these two feature dimensions.

Further details of the data structure were required in order to make any clear statement concerning the degree of separability of the classes, given their apparent multimodality, so **Sammon's** nonlinear mapping was chosen for this purpose because had already been used successfully in Chapter 4. An earlier 160 point version of the data set was used (see fig. 5.36) in order to determine if the clusters within each defect class were really as severely interpenetrating in the full feature space dimensionality as fig. 5.35 would seem to imply. The results are discouraging because the plot indicates that the classes are not separable to any practically useful extent. It would have been desirable to have been able to give a plot of the same 432 point data set as is given in fig. 5.35, but since the time to iterate to an acceptable representation error is related to $(n_{points})^2$, and the time taken to produce fig. 5.35 with 160 points was about an hour, this was not practical.

Given the small size of the data set it would have been preferable to

have used a leave one out approach, whereby the entire data set is used as a training set, with the except of one point, which is then classified according to the training set, and this procedure repeated for each point in turn, while the misclassification rate is calculated as an average of the aggregated results. The method would have allowed a larger feature vector to have been used, but there was insufficient time to run the classifier the required number of times. This would have been approximately 800 times, equal to the total number of defect scans in the final scan set.

### 5.9 Conclusions

At the time that the classification work described in Sections 5.6 and 5.7 was done, no automatic segmentation method existed. Such a method has now been successfully tested upon real loci (section 5.6).

The practical problems of determining the best feature subset for classifying a number of classes of signal have fuelled the development of an automatic feature selection scheme which is intended to be robust enough to function when little is understood about the multidimensional class structure in the featurespace, and so no assumptions can be made concerning the parametric form of the data. Such a measure has been devised and implemented (see section 3.4), and the resulting measure has been tested upon a synthetic data set. Testing on a featureset extracted from real loci can proceed as soon as a feature extractor is available.

The design of the ad hoc feature set has been explained. Some ability to classify defects by type was obvious in this featureset though not as much as

had been hoped. Work has been performed on systematically defined features in the hope of producing a better featurest (see section 3.3). However the resulting methods have not yet been successfully applied to real loci because the features produced cannot yet reproduce the original locus (which is as near as it is possible to get to a validation test for the method).

There are a number of other ways in which the results might be improved and the goal of accurate automatic classification achieved, and these form the bulk of the next and last chapter.

**5.10 References**

1. **Foster, D.J.**, "Intergranular Attack Specimen Preparation", Babcock Internal Report (62)/85/23, May 1985.

2. **Harris, F.J.**,"On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform", Proc. IEEE, Vol. 66, no.1, pp51- 83, Jan. 1978.

3. **Van Drunen, G. and Sharp, F.L.**, "Eddy Current Inspection of Installed Candu Pressure Tubes", Proc. 3rd Conf on Quantitative Reliability in NonDestructive Examination, pp691-700 , Zurich, 1983.

4. **Akl S.G. and Toussaint, G.T.**,"Efficient Convex Hull Algorithms for Pattern Recognition Applications", Proc. 4th Int. Joint Conf. on Patt. Rec., Nov. 1978, Kyoto, Japan.

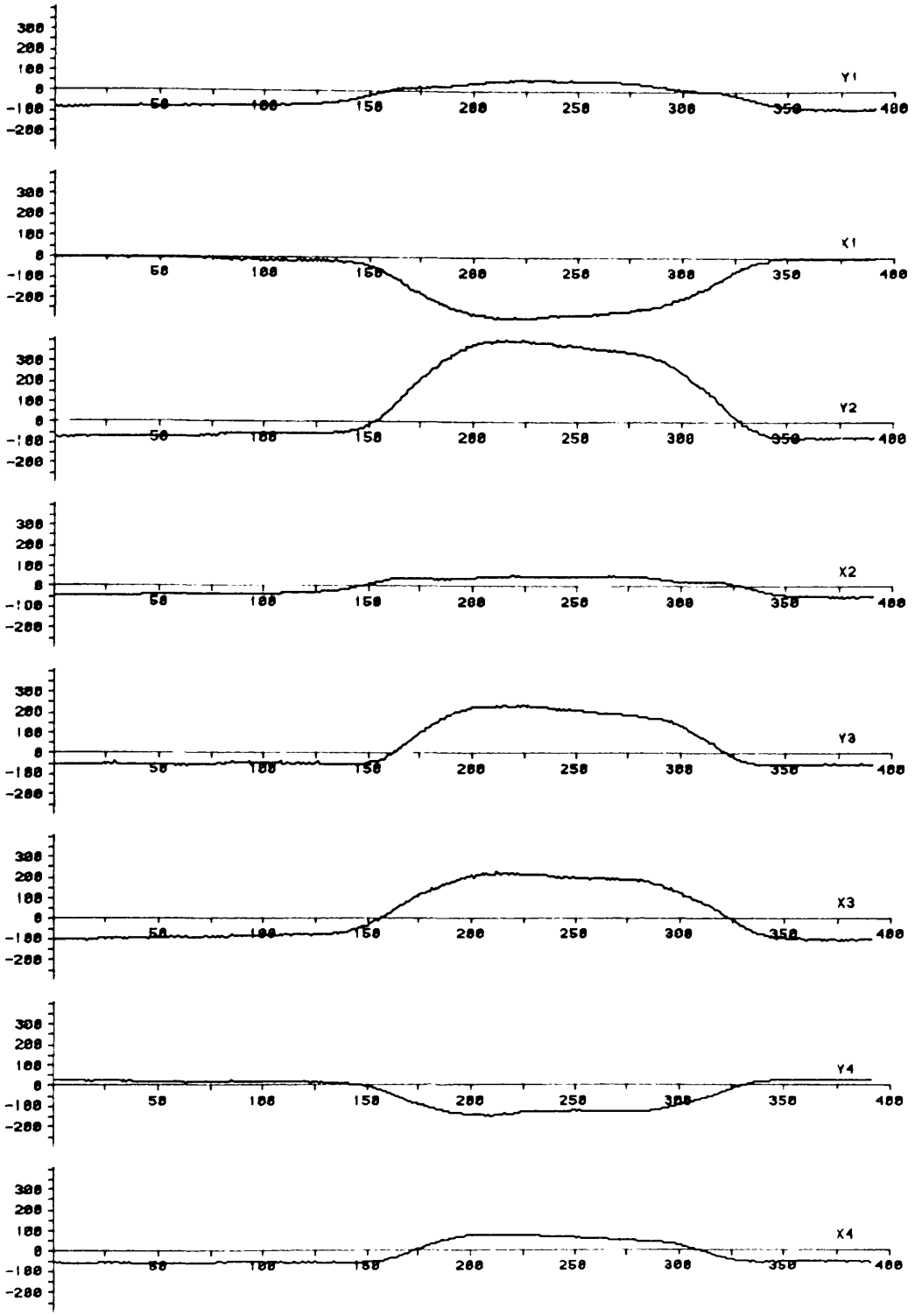



FIG. 5.1

CENTRAL SCAN FROM PATCH OF 60% THRU WALL IGA.
INPUT FILE DP0:TUB346.RAW
SCAN 100 TIME SERIES (VERT SCALE = A.D.C. UNITS)
PLOTTED ON 10-AUG-84 AT 11:22 (+128 = 1 VOLT)

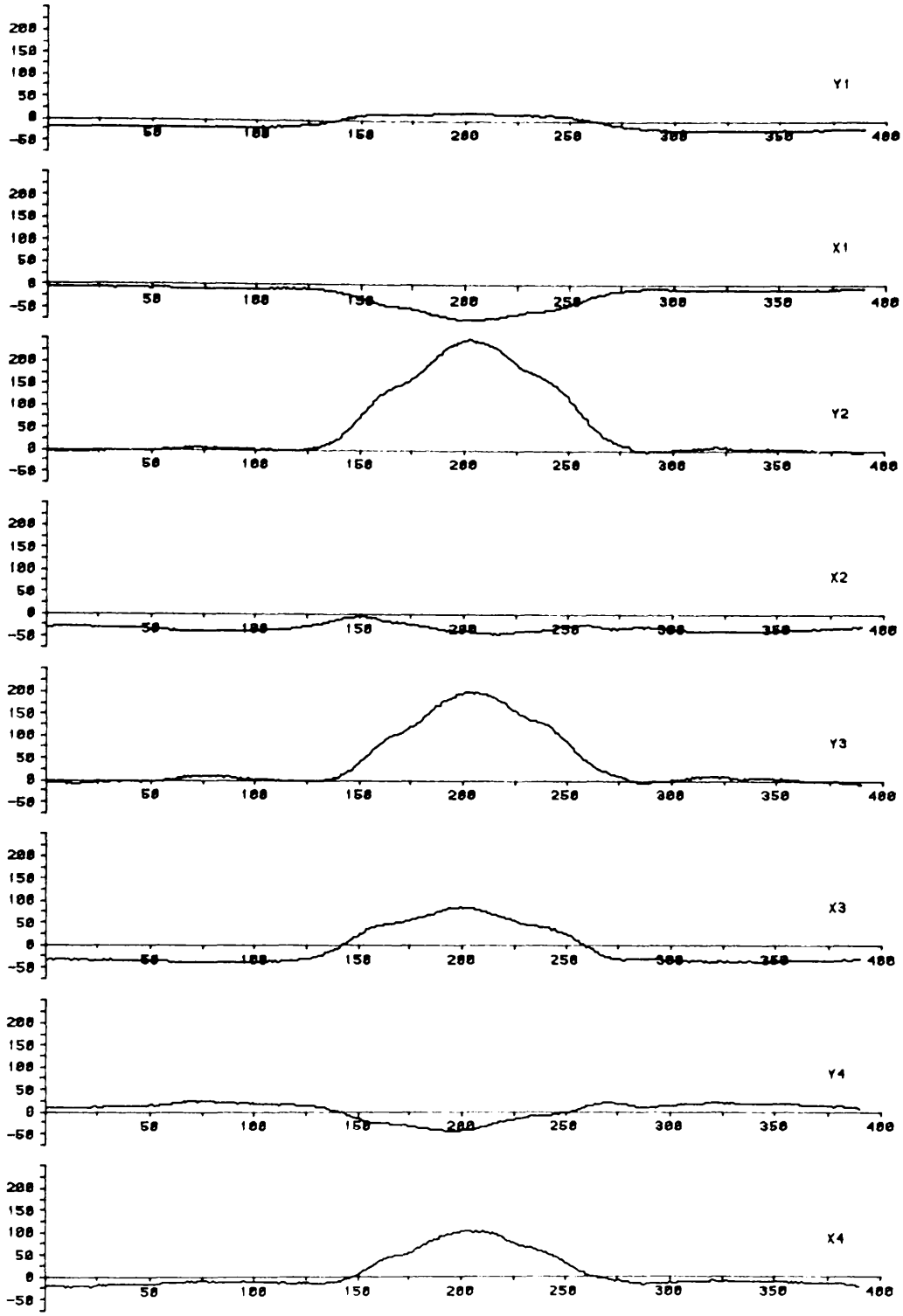



FIG. 5.2

CENTRAL SCAN THROUGH AN 80% THRU WALL CIRC. NOTCH
INPUT FILE DP0:TUB15C.RAW
SCAN 52 TIME SERIES (VERT SCALE = A.D.C UNITS)
PLOTTED ON 10-AUG-84 AT 10:42 (+128 = 1 VOLT)

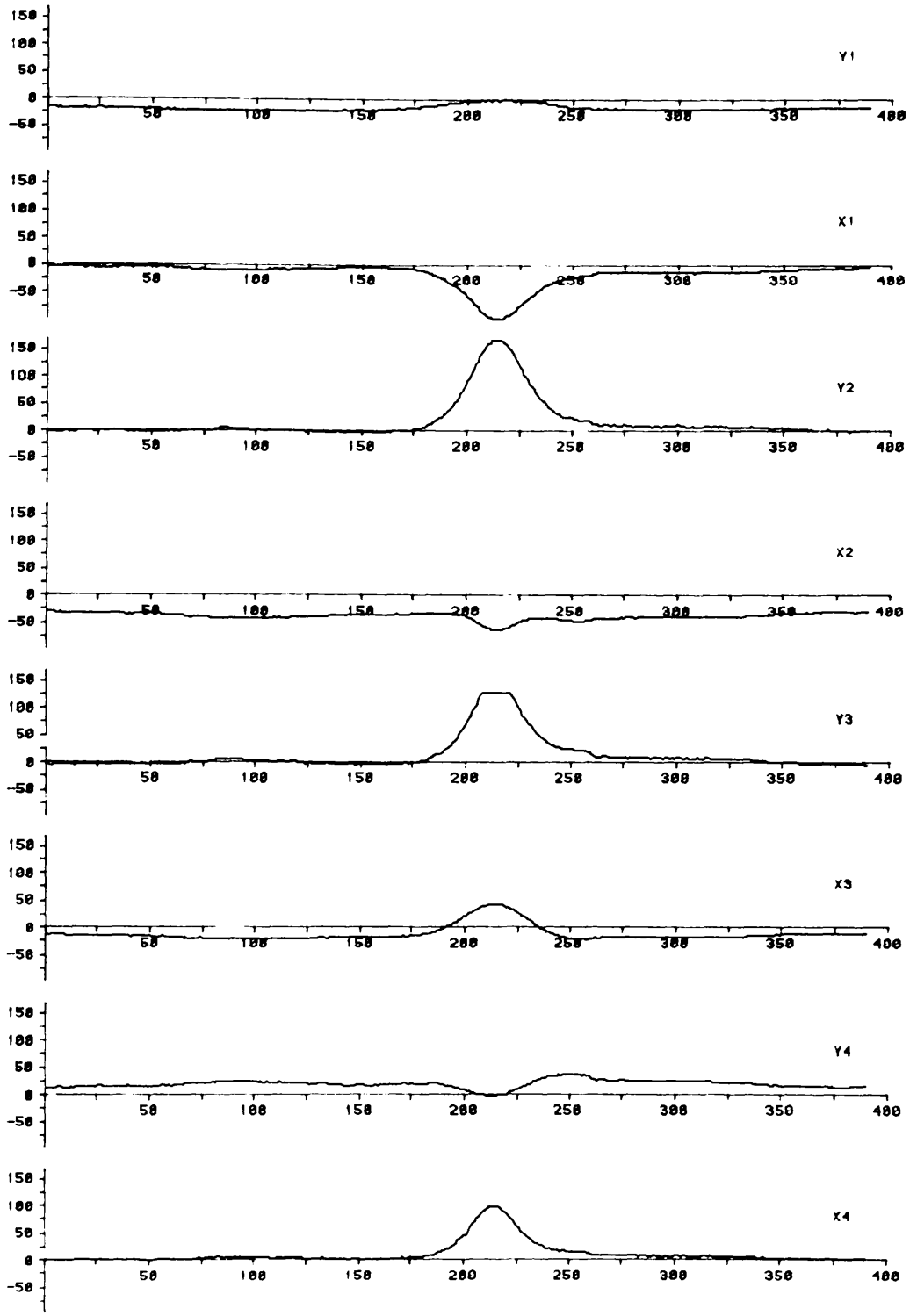



FIG. 5.3

A SCAN CENTRAL TO A 78% THRU WALL AXIAL NOTCH
INPUT FILE DP0:TUBE5F.RAW
SCAN 56 TIME SERIES (VERT SCALE = A.D.C. UNITS)
PLOTTED ON 10-AUG-84 AT 10:22 (+128 = 1 VOLT)

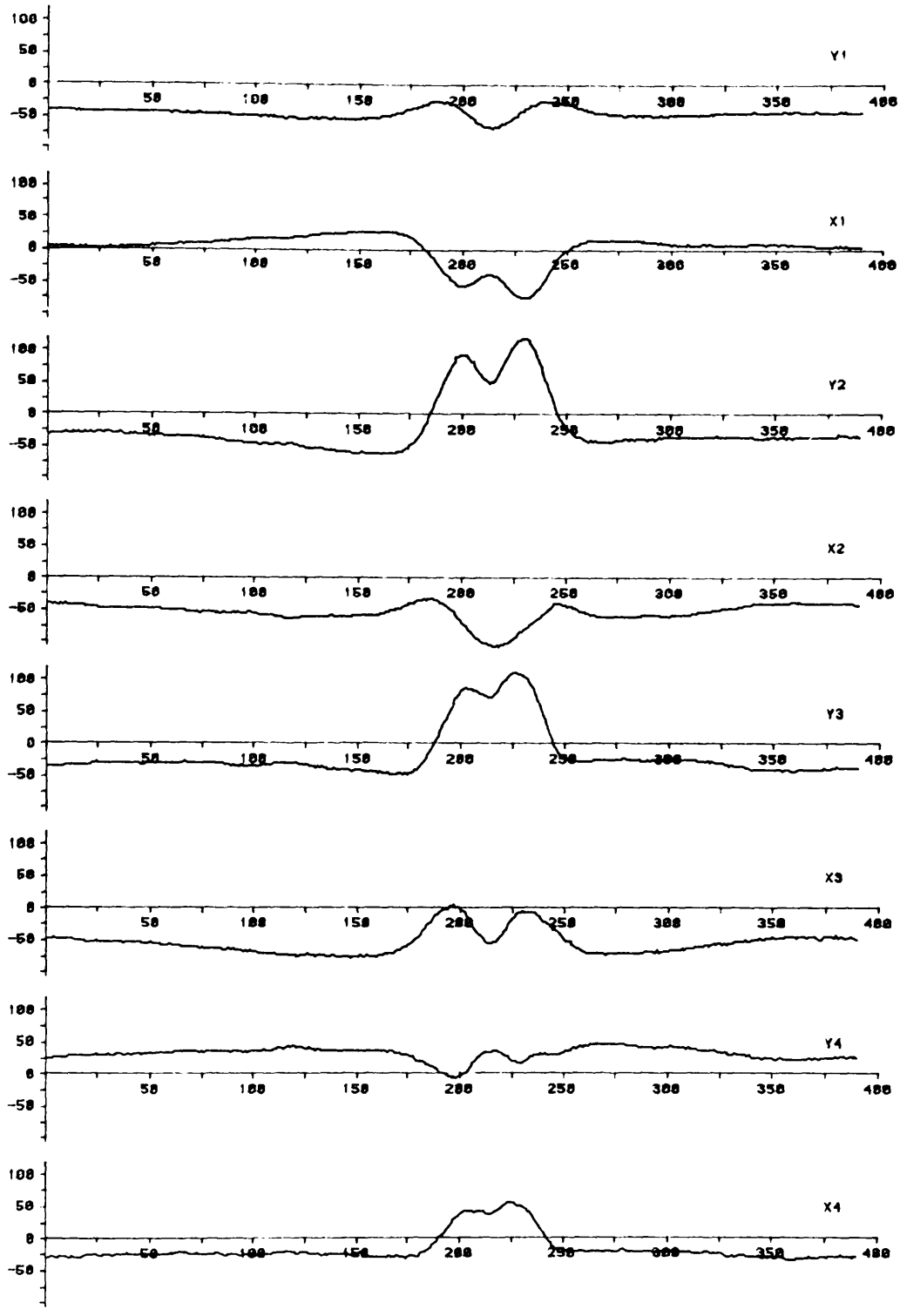



FIG. 5.4

SCAN ACROSS CENTRE OF FLAT BOTTOMED HOLE.
INPUT FILE DP0:TUB342.RAW
SCAN 10 TIME SERIES (VERT SCALE = A.D.C. UNITS)
PLOTTED ON 10-AUG-84 AT 13:26 (+128 = 1 VOLT)

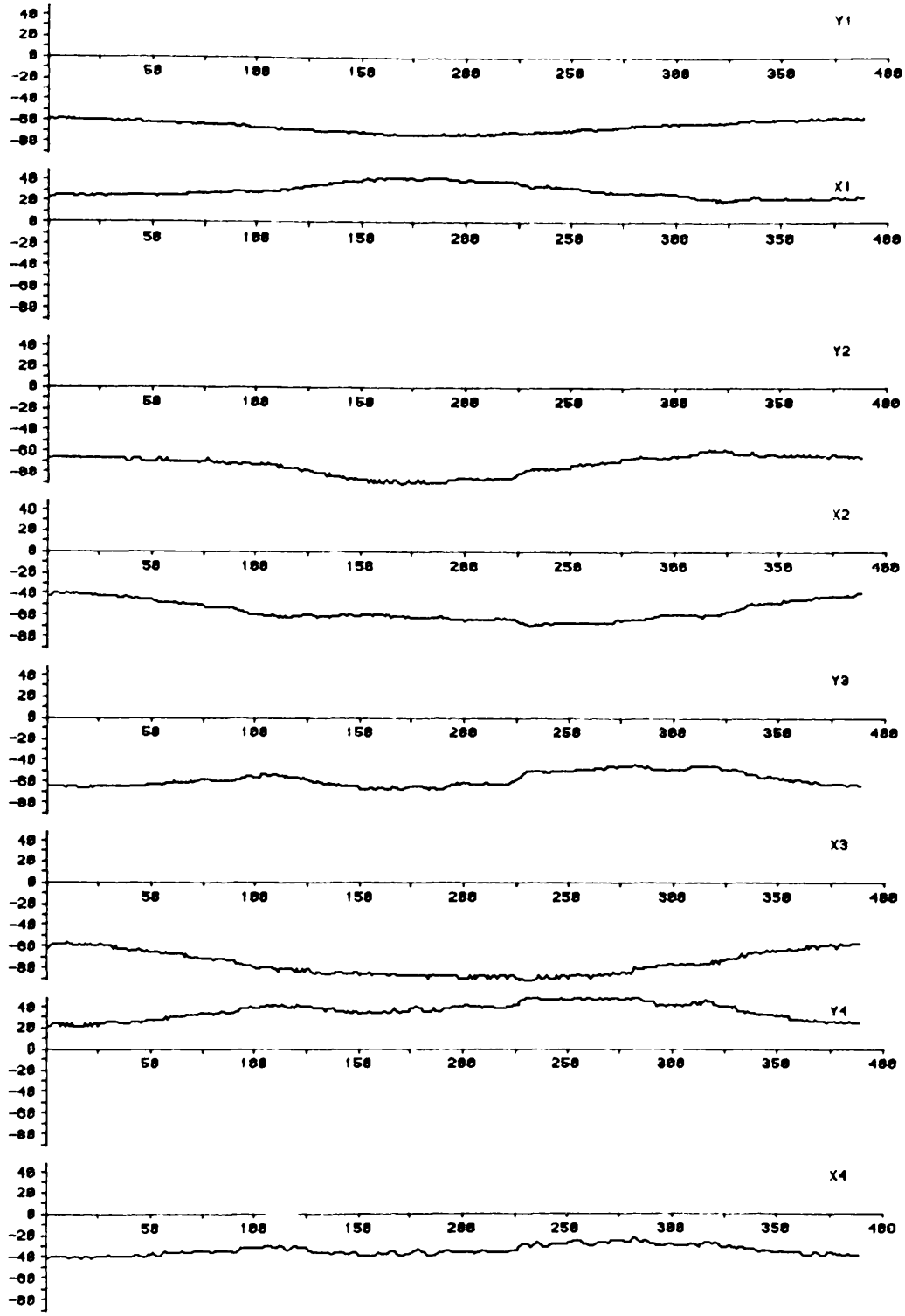



FIG. 5.5

A SCAN FROM CLEAN TUBING.
INPUT FILE DP0:TUB342.RAW
SCAN 40 TIME SERIES (VERT SCALE = A.D.C. UNITS)
PLOTTED ON 10-AUG-84 AT 14:35 (+128 = 1 VOLT)

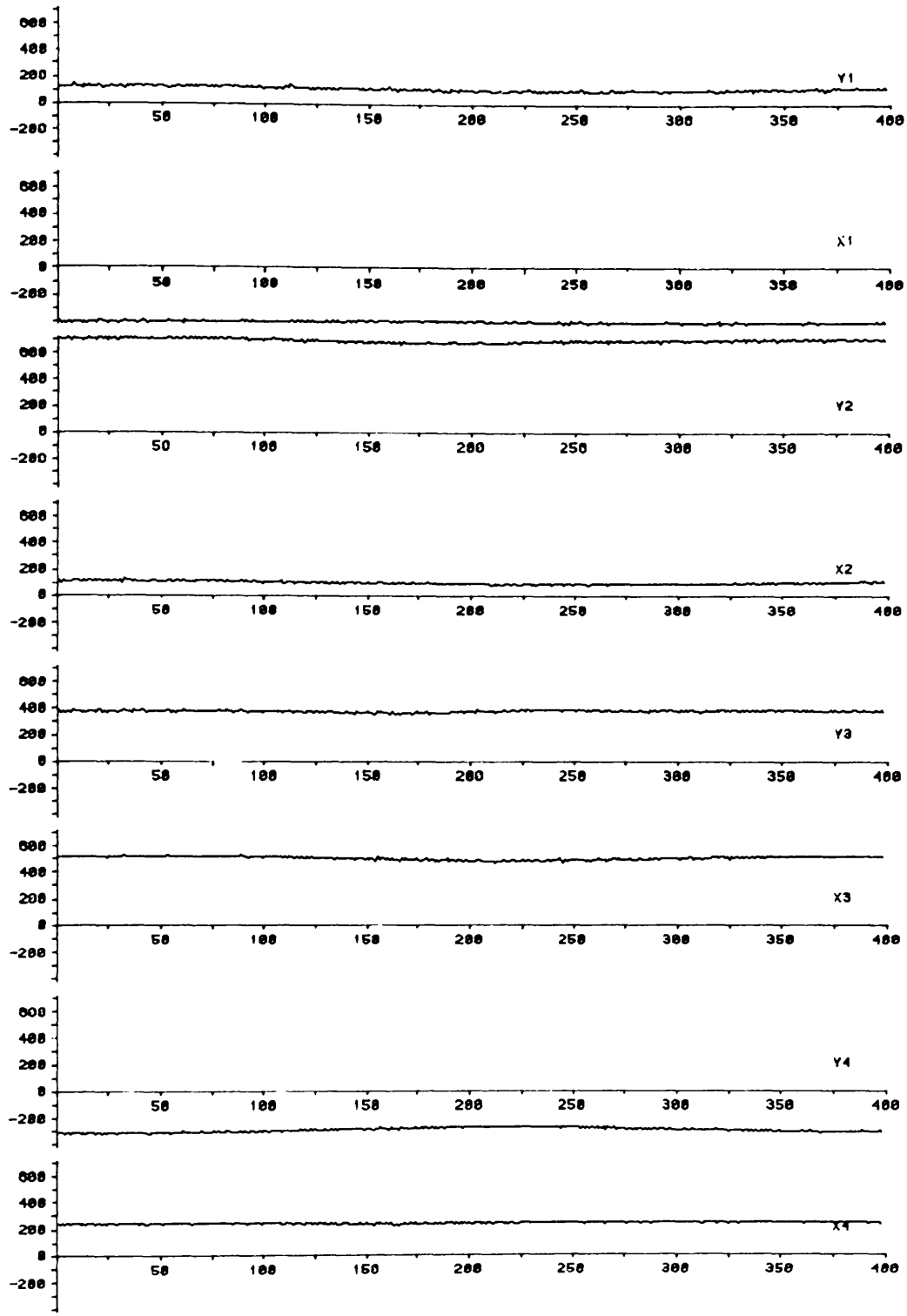



FIG. 5.6

48% THRU WALL UNIFORM THINNING, CENTRE OF BAND
INPUT FILE DP0:TUB18F.RAW
SCAN 65 TIME SERIES (VERT SCALE = A.D.C. UNITS)
PLOTTED ON 12-AUG-84 AT 13:55 (+128 = 1 VOLT)

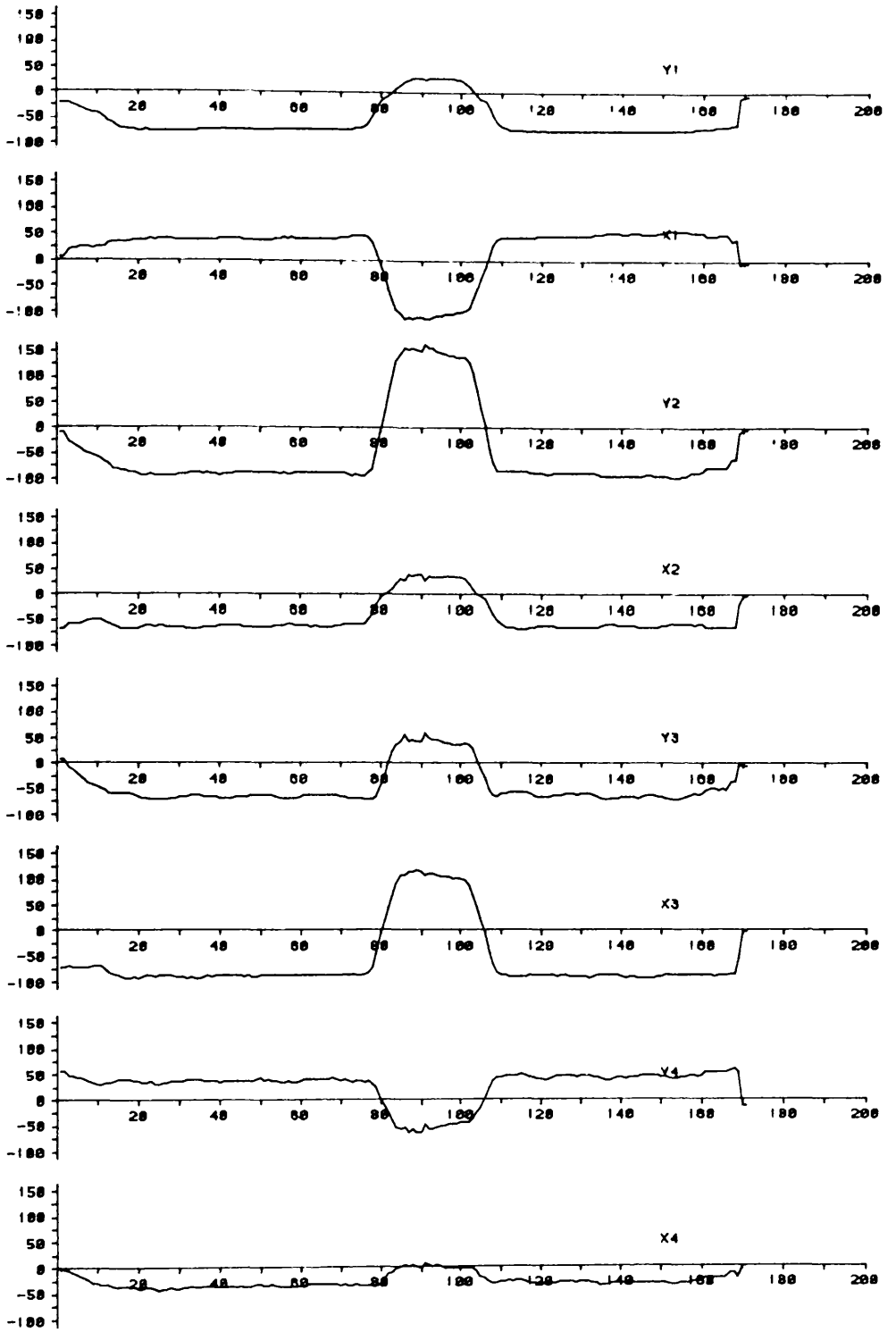



FIG. 5.7

AXIAL SCAN THROUGH PATCH OF 40% THRU WALL IGA
INPUT FILE DP0:TUB342.RAW
SCAN 0 TIME SERIES (VERT SCALE = A.D.C. UNITS)
PLOTTED ON 11-AUG-84 AT 15:15 (+128 = 1 VOLT)

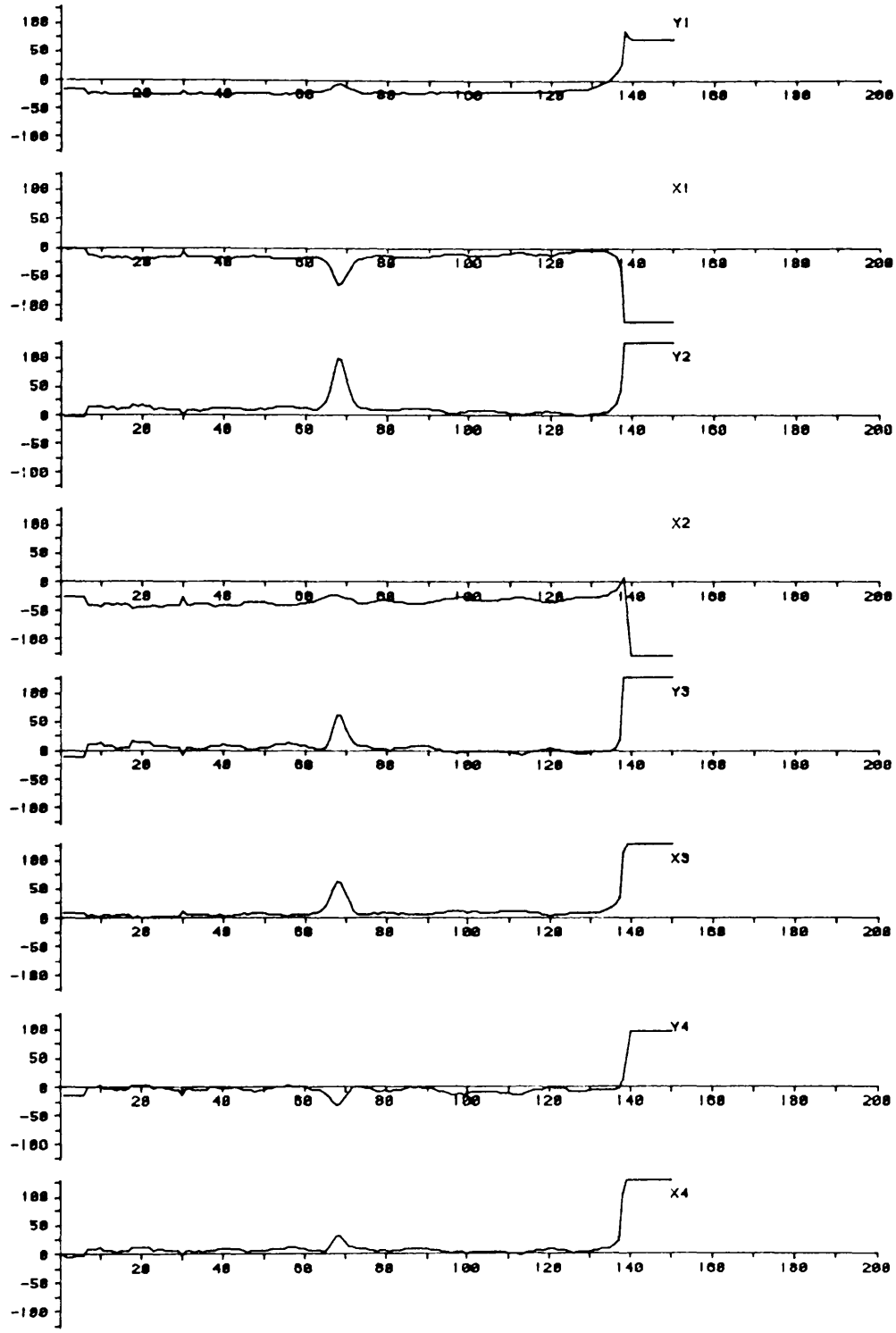



FIG. 5.8

56% THRU WALL CIRC. NOTCH, AXIAL SCAN THRU CENTRE
INPUT FILE DP0:TUB14C.RAW
SCAN 0 TIME SERIES (VERT SCALE = A.D.C. UNITS)
PLOTTED ON 11-AUG-84 AT 16:12 (+128 = 1 VOLT)

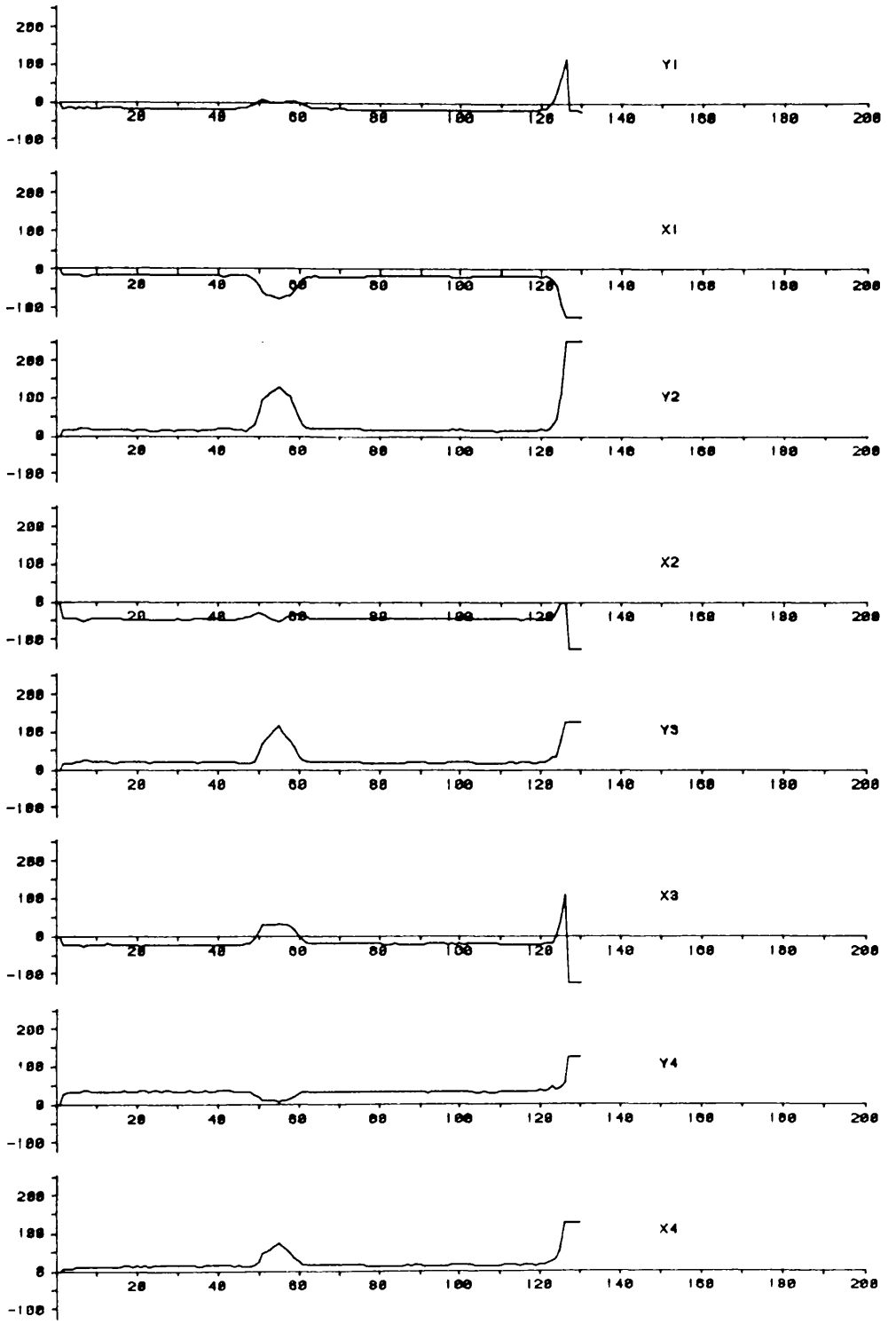



FIG. 5.9

AXIAL SCAN ALONG A 56% THRU WALL AXIAL NOTCH
INPUT FILE DP0:TUBE5F.RAW
SCAN 0 TIME SERIES (VERT SCALE = A.D.C. UNITS)
PLOTTED ON 11-AUG-84 AT 16:48 (+128 = 1 VOLT)

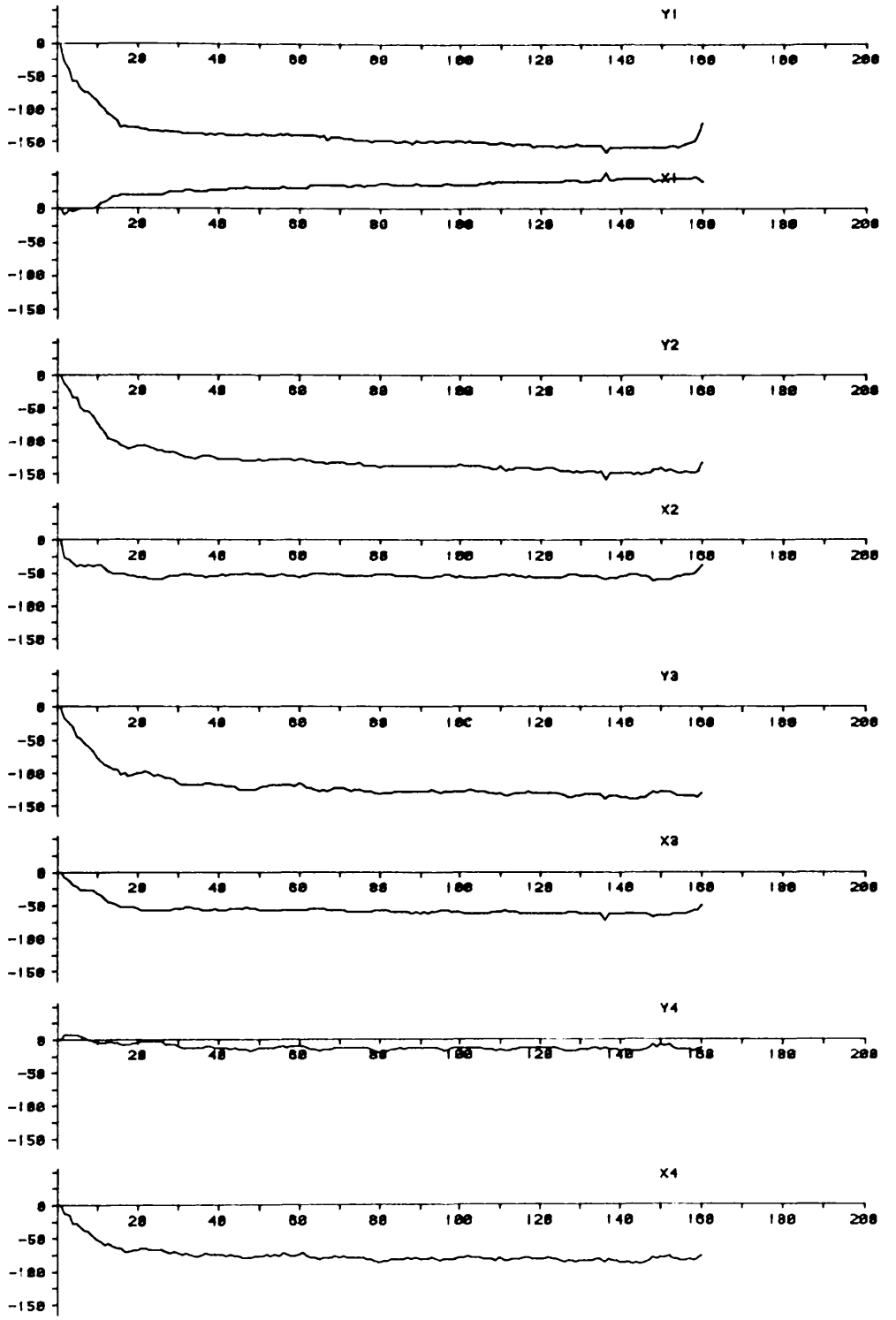



FIG. 5.10

AXIAL SCAN TO SHOW THE PILGER (EXTRUSION) EFFECT
INPUT FILE DP0:TUBE9.RAW
SCAN 0 TIME SERIES (VERT SCALE = A.D.C. UNITS)
PLOTTED ON 11-AUG-84 AT 17:08 (+128 = 1 VOLT)

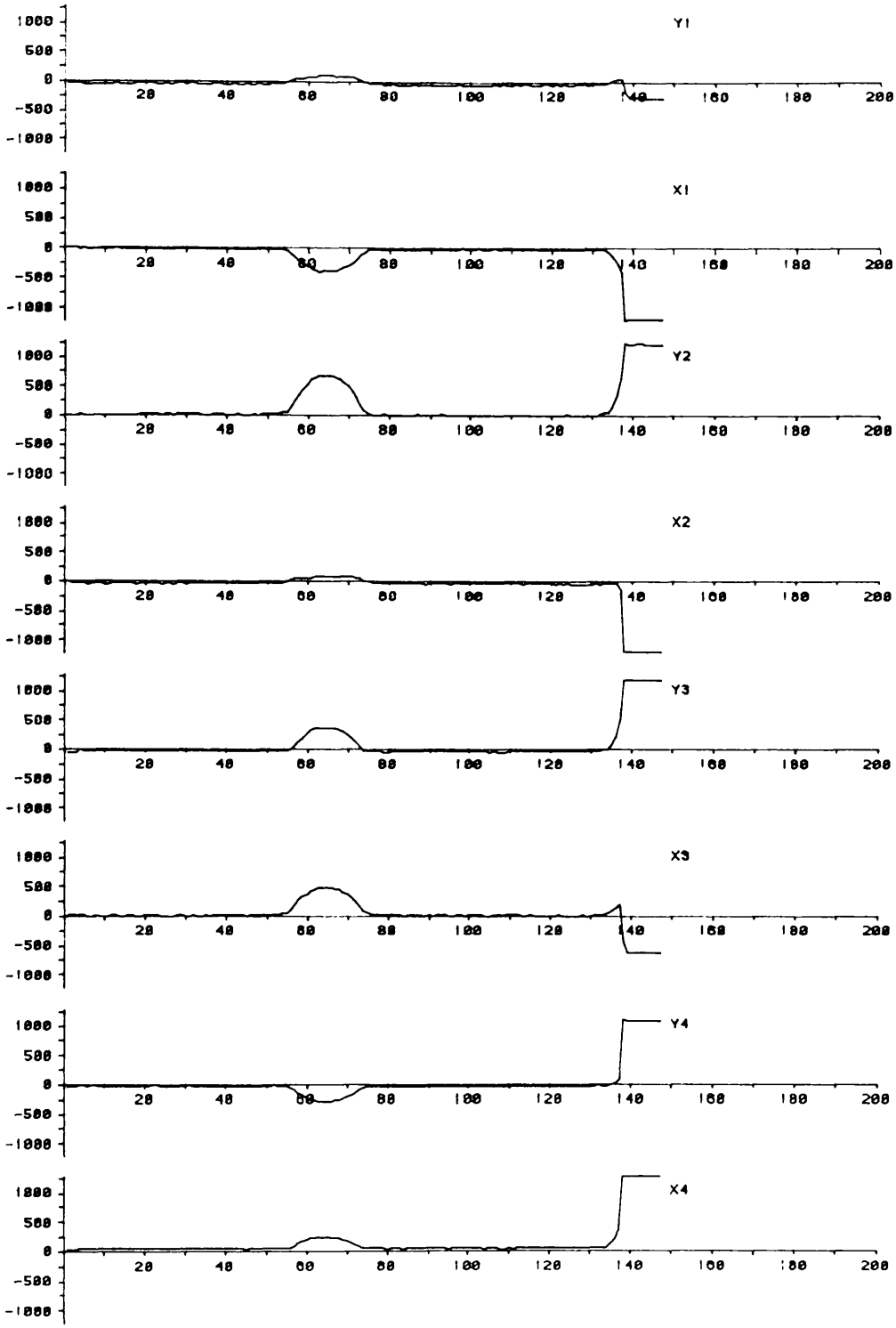
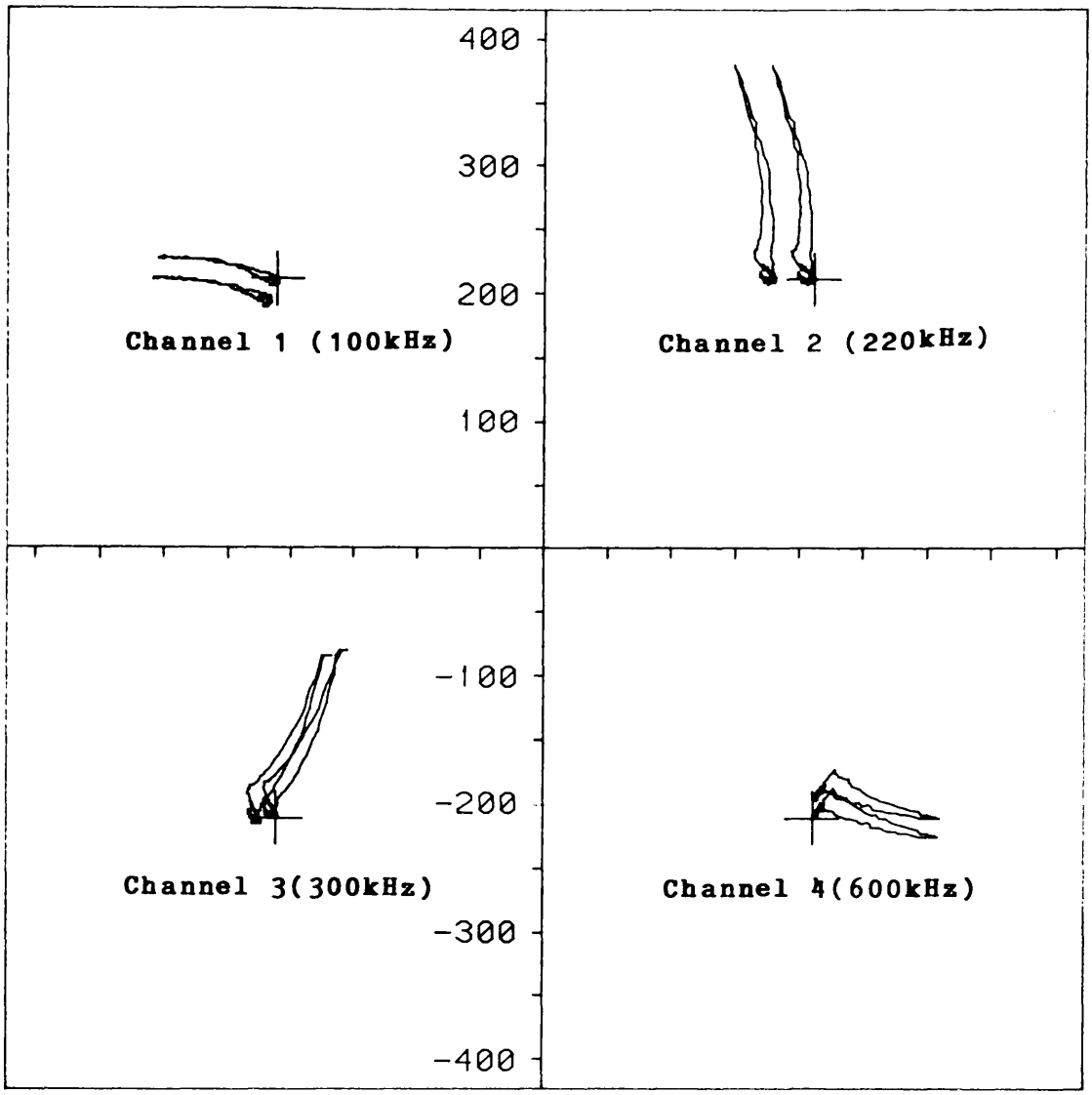


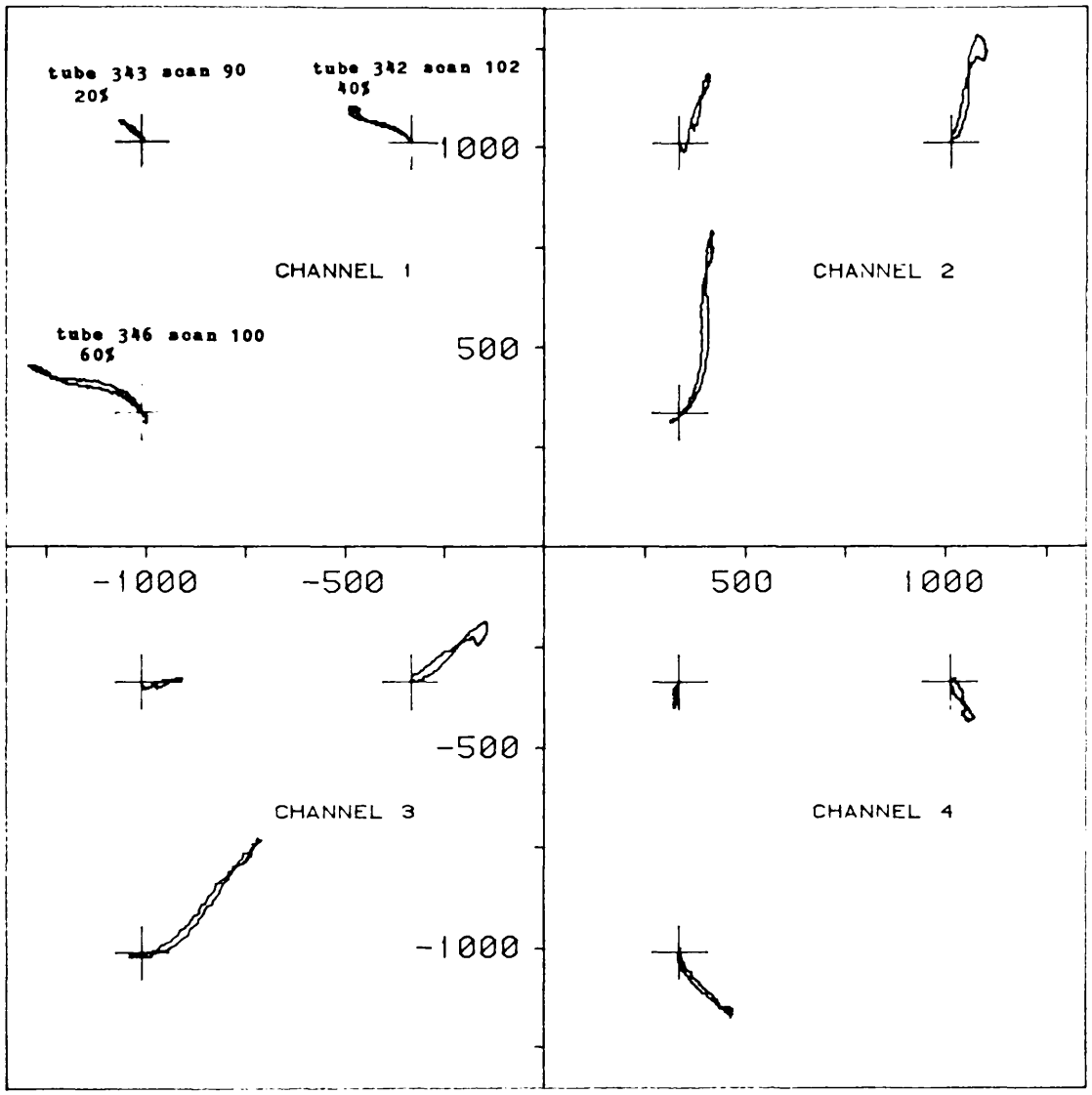

FIG. 5.11

AXIAL SCAN THROUGH BAND OF 48% THRU WALL THINNING
INPUT FILE DP0:TUB18F.RAW
SCAN 0 TIME SERIES (VERT SCALE = A.D.C. UNITS)
PLOTTED ON 11-AUG-84 AT 15:43 (+128 = 1 VOLT)



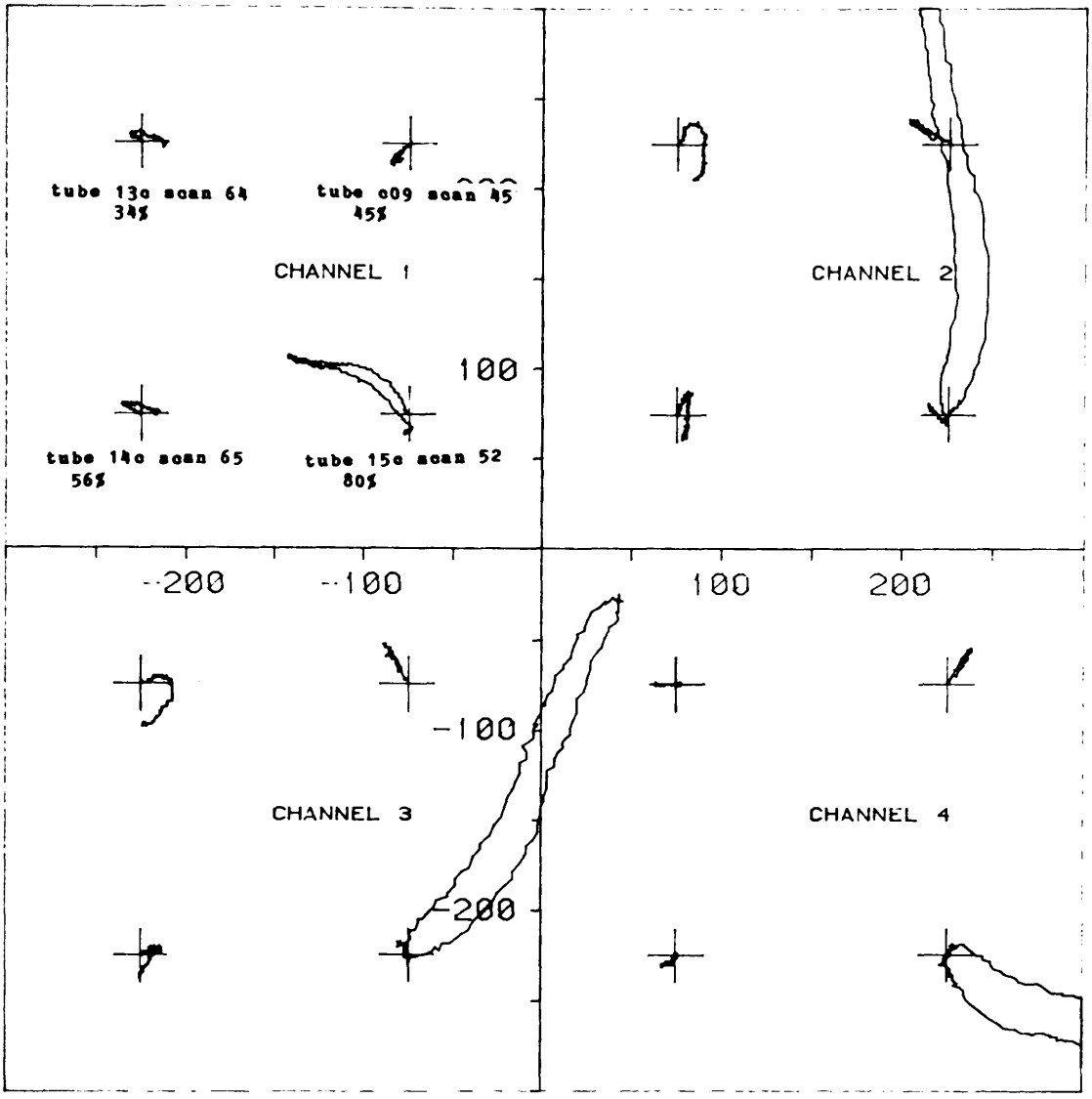

SCOPE4 RUN ON 12-AUG-84 AT 22:09
LINE 56 POINTS 1 TO 390 FILE DP0:TUBE5F.RAW
BLEMISH CLASS ZNCH
78% THRU WALL AXIAL NOTCH, D.C. OFFSET RETAINED

FIG. 5.12 ZNCH, WITH AND WITHOUT D.C. COMPONENT




SCOPE4 RUN ON 13-AUG-84 AT 20:50
BLEMISH CLASS IGA
VARYING DEPTHS OF IGA, D.C. OFFSET REMOVED

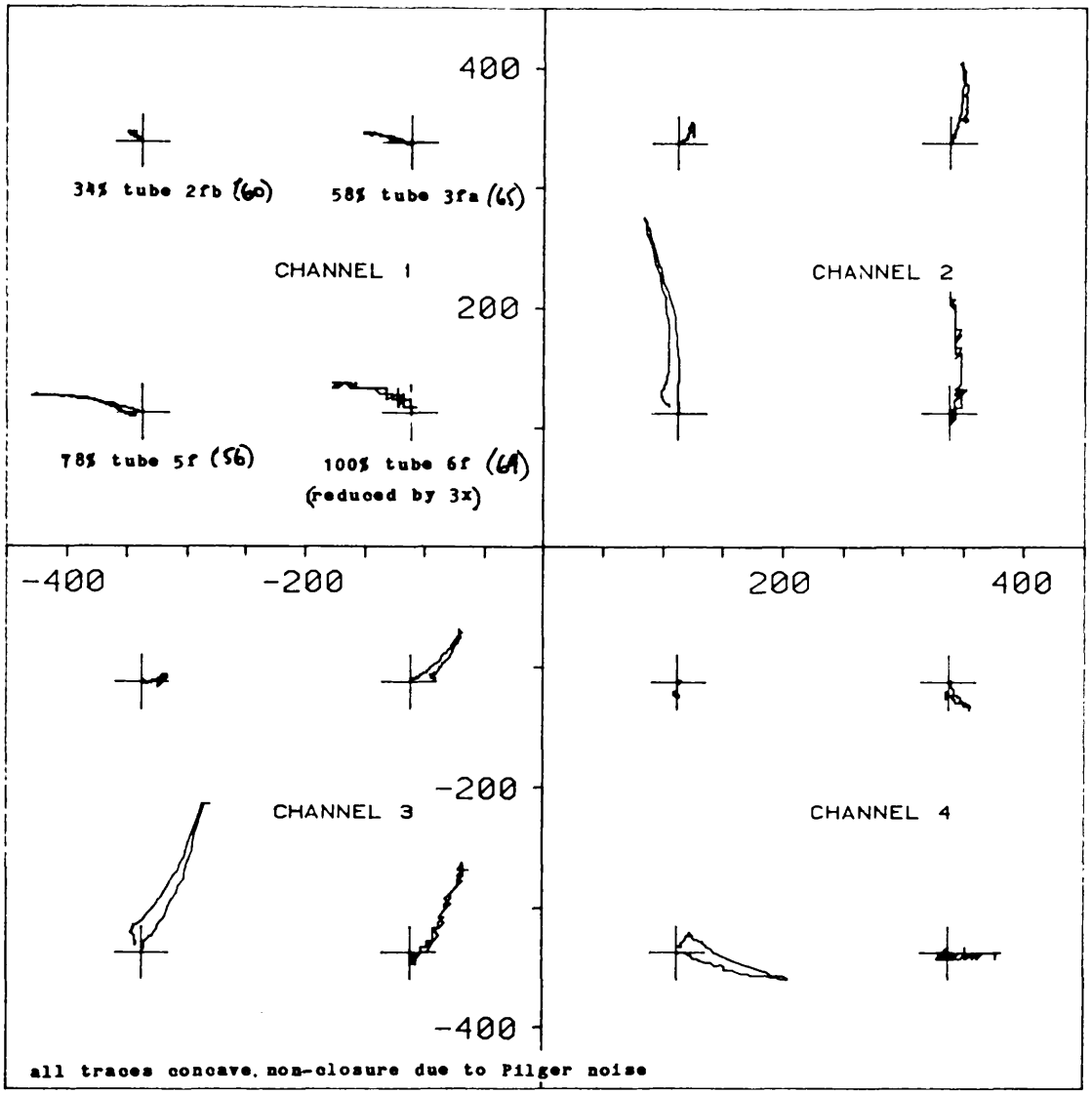
FIG. 5.13 IGA ($HNO_3$ INDUCED), VARIOUS DEPTHS




SCOPE4 RUN ON 13-AUG-84 AT 21:38
BLEMISH CLASS CNCH
VARYING DEPTHS OF CIRCUMFERENTIAL NOTCH

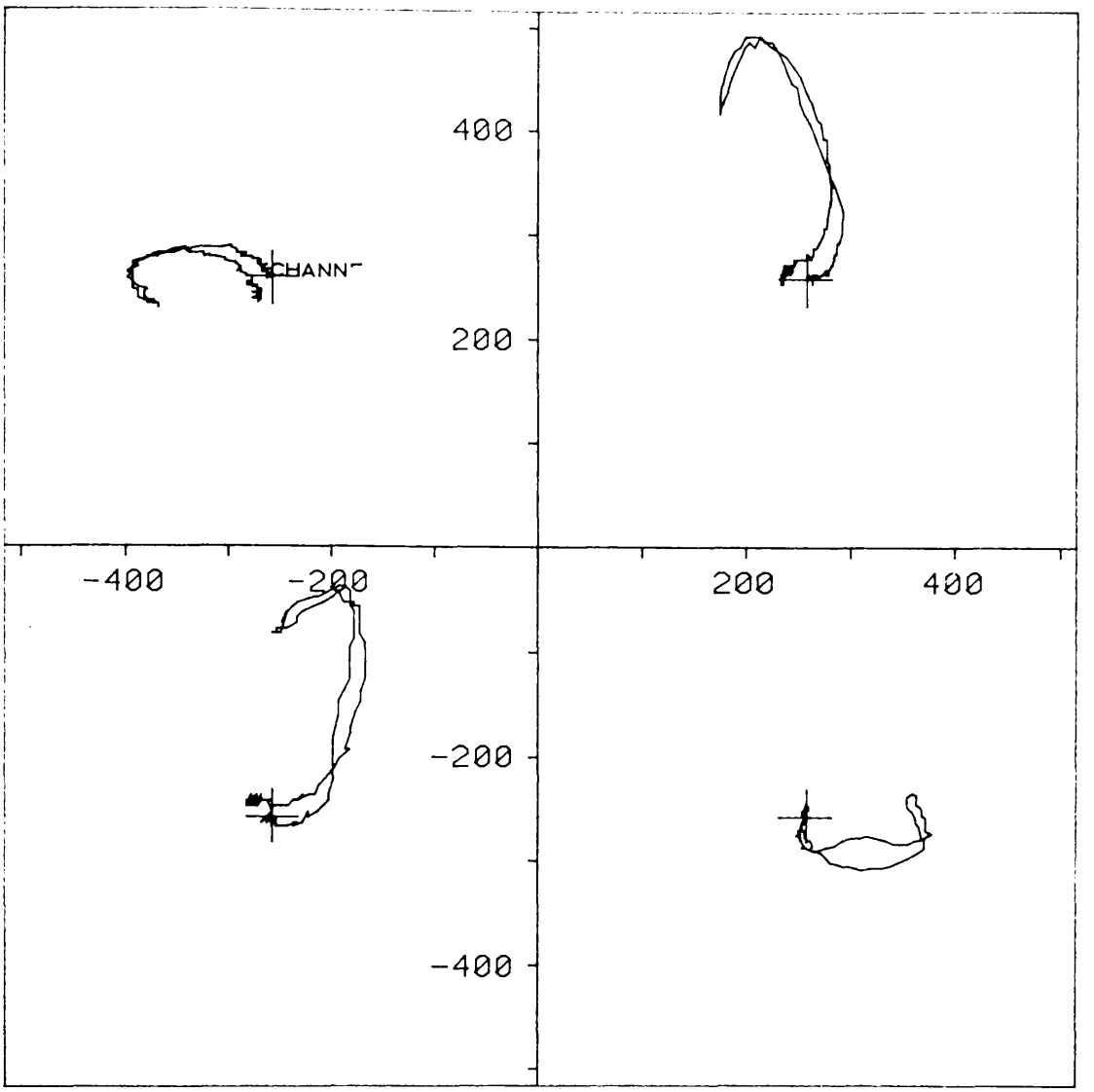
FIG. 5.14 CNCH, VARIOUS DEPTHS

(Note that the loci for all four depths rotate equally with increasing frequency.)



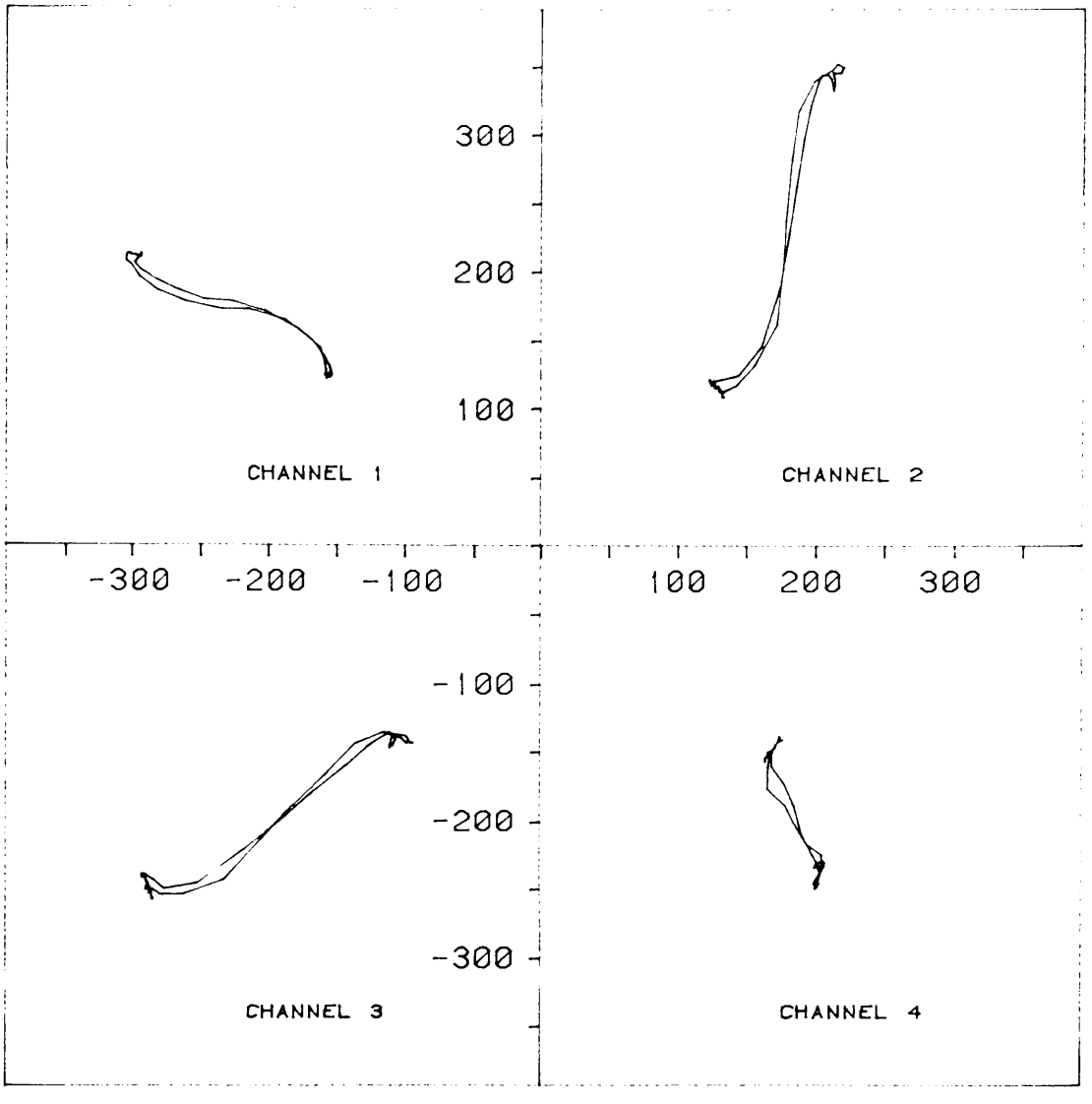

SCOPE4 RUN ON 12-AUG-84 AT 23:19
BLEMISH CLASS ZNCH
VARYING DEPTHS OF AXIAL NOTCH, D.C. REMOVED

FIG. 5.15 ZNCH, VARIOUS DEPTHS



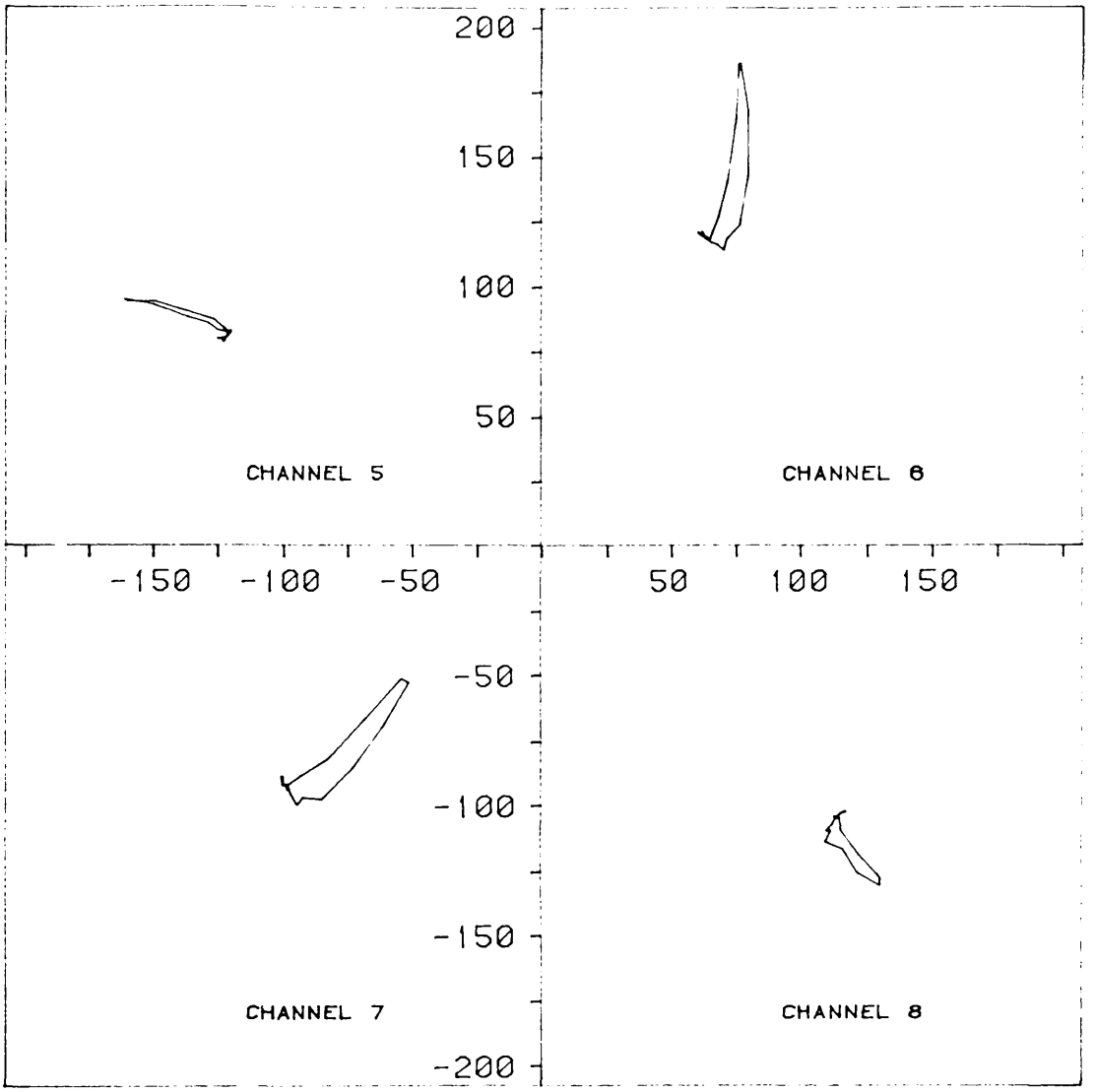

SCOPE4 RUN ON 13-DEC-84 AT 19:44
LINE 10 POINTS100 TO 300  FILE TUB346.RAW
BLEMISH CLASS HOLE
CENTRAL SCAN THROUGH HOLE

FIG. 5.16



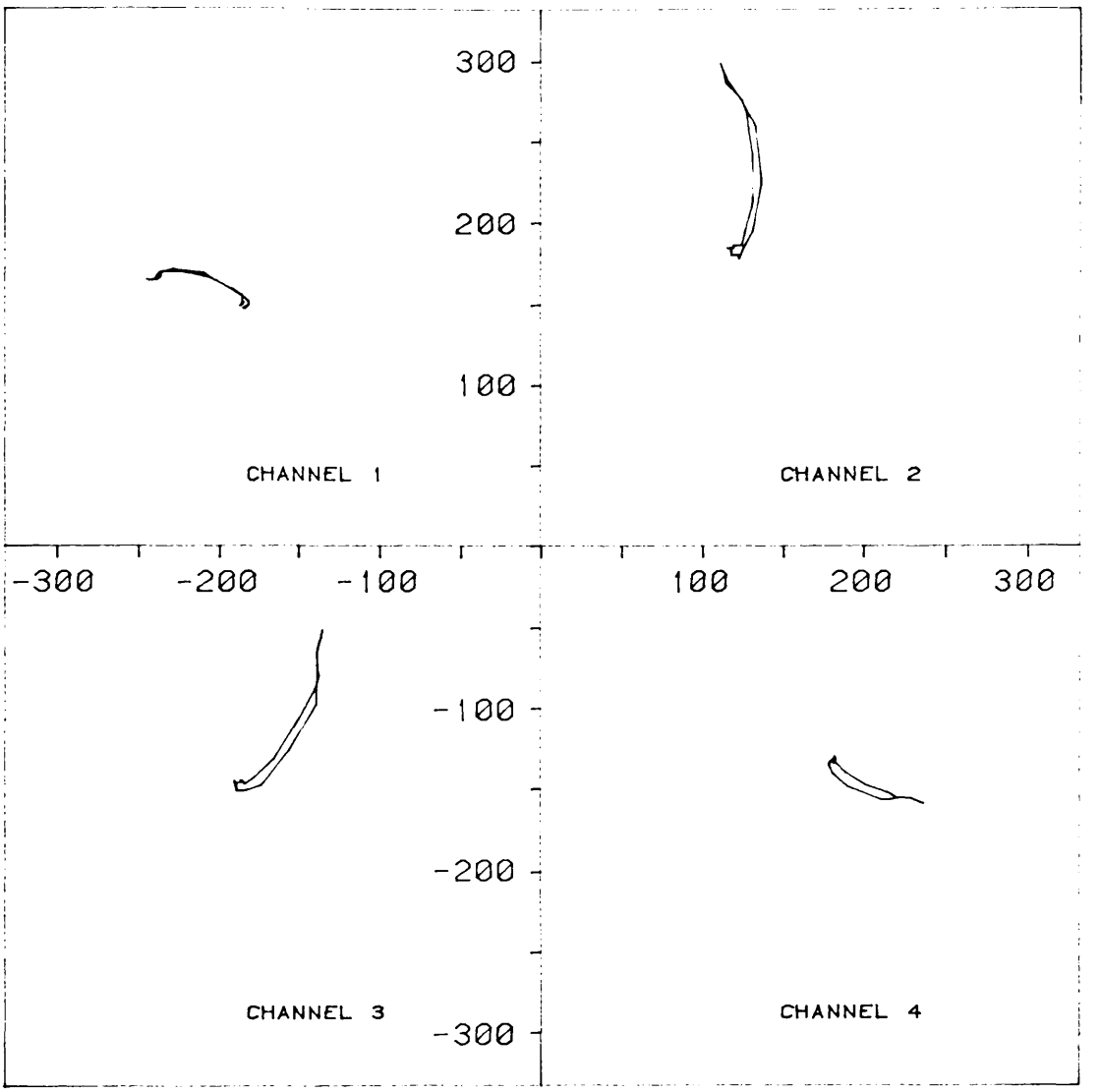

BOBBN4 RUN ON 01-OCT-84 AT 06:51
THETA 200 LINES 60 TO 120 FILE DP0:TUB342.RAW
BLEMISH CLASS PIGA
40% THRU WALL IGA, CORRESPONDING TO FIGURE 5.13

FIG. 5.17



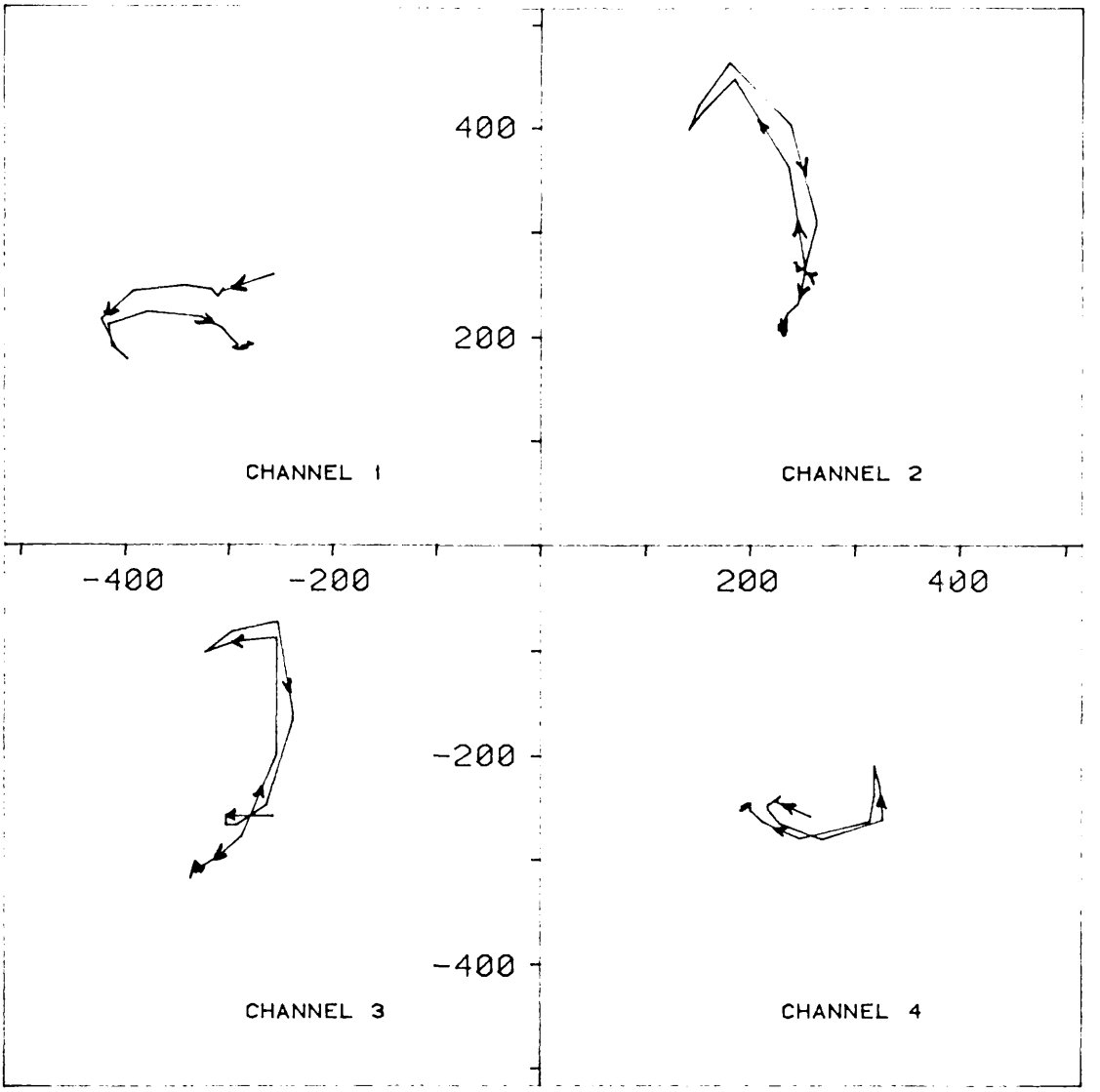

BOBBN4 RUN ON 01-OCT-84 AT 07:04
THETA 200 LINES 60 TO 80 FILE DP0:TUB14C.RAW
BLEMISH CLASS CNCH
56% THRU WALL CIRC. NOTCH, AS IN FIGURE 5.14

FIG. 5.18



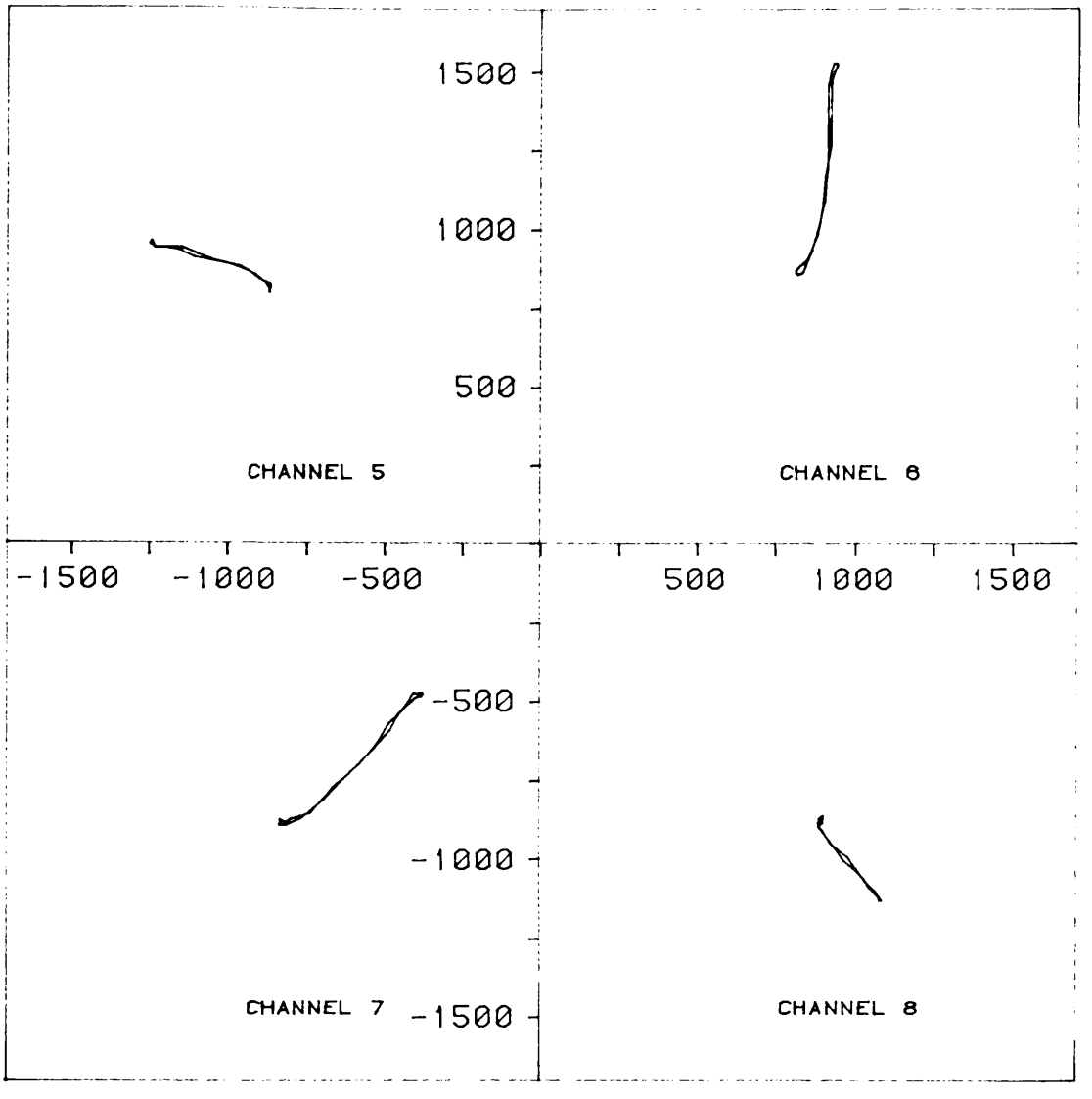

BOBBN4 RUN ON 01-OCT-84 AT 07:16
THETA 225 LINES 40 TO 70 FILE DP0:TUBE5F.RAW
BLEMISH CLASS ZNCH
56% THRU WALL AXIAL NOTCH, AS IN FIGURE 5.15

FIG. 5.19




BOBBN4 RUN ON 13-DEC-84 AT 19:40
THETA 200 LINES 1 TO 40 FILE TUB346.RAW
BLEMISH CLASS HOLE
AXIAL SCAN THROUGH CENTRE OF HOLE
COMPARED WITH FIG. 5.16

FIG. 5.20




BOBBN4 RUN ON 01-OCT-84 AT 07:29
THETA 200 LINES 50 TO 80 FILE DP0:TUB18F.RAW
BLEMISH CLASS THIN
48% THRU WALL THINNING, AS IN FIGURE

FIG. 5.21 (NO EQUIVALENT SCOPE 4 PLOT)

FIG. 5.22 PATCH IGA PILGER NOISE AND A HOLE

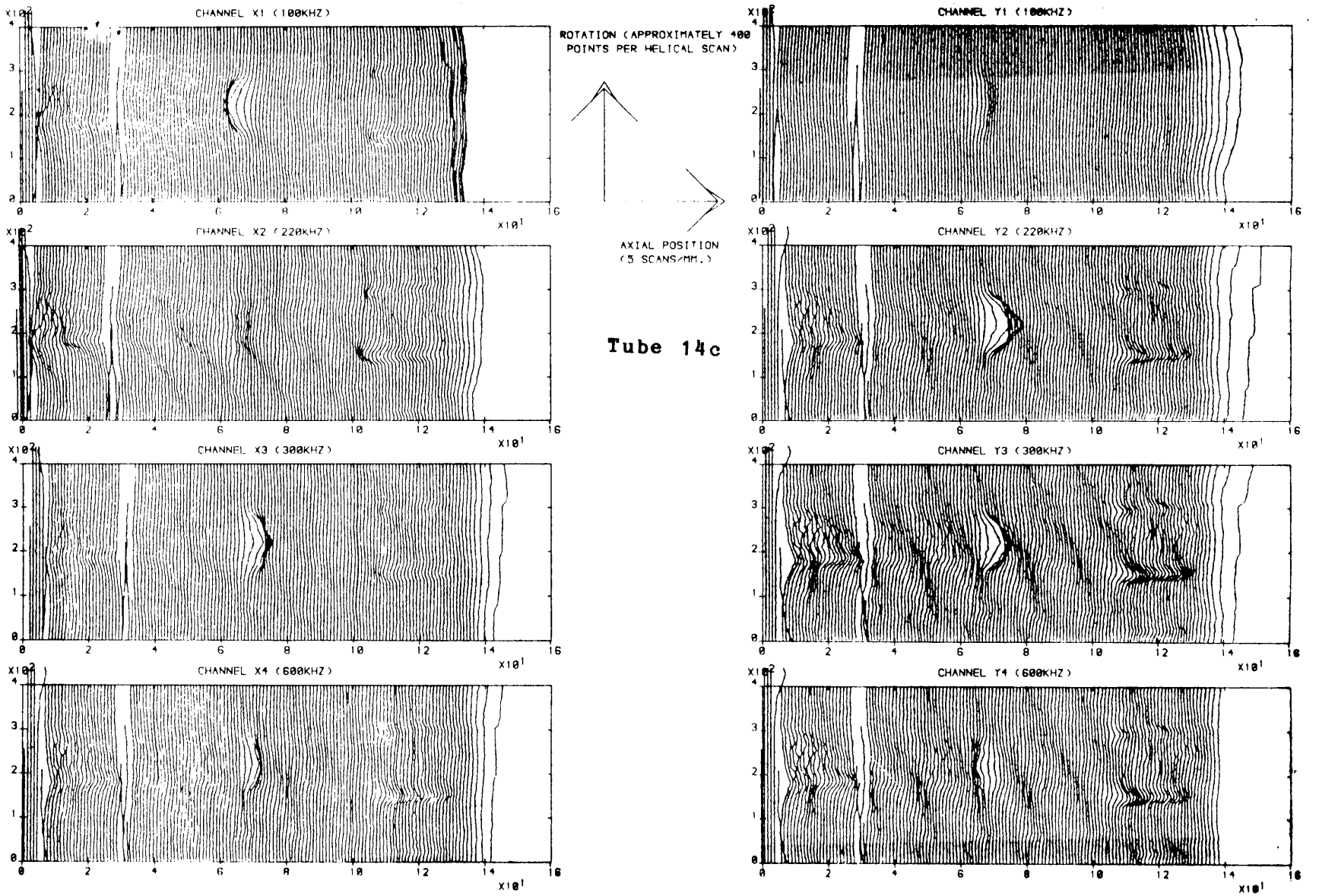



FIG. 5.23 CNCH AND A SCRATCH

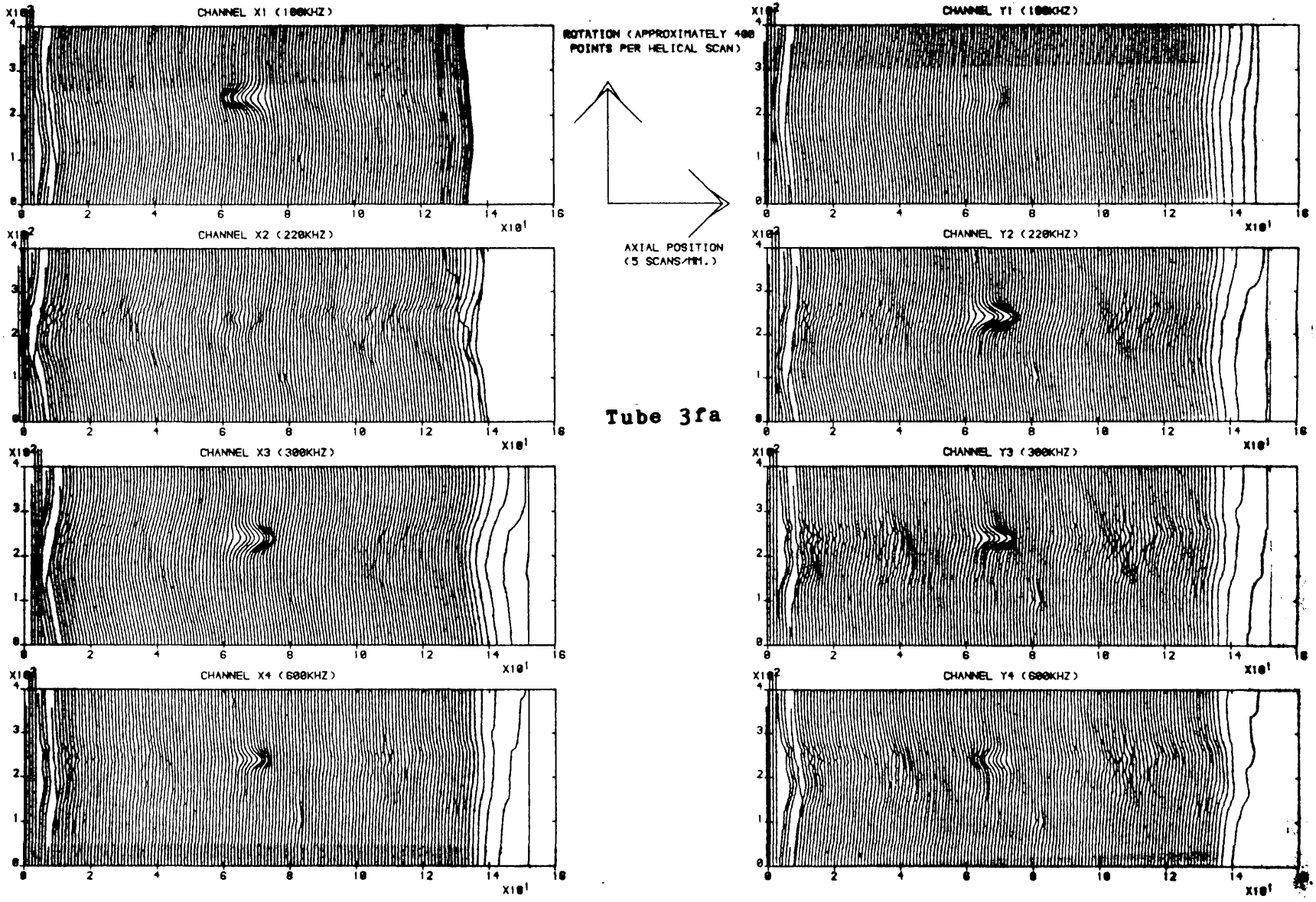

CHANNEL X1 (100KHZ)
CHANNEL Y1 (100KHZ)
ROTATION (APPROXIMATELY 400 POINTS PER HELICAL SCAN)
AXIAL POSITION (5 SCANS/MM.)
CHANNEL X2 (220KHZ)
CHANNEL Y2 (220KHZ)
Tube 3fa
CHANNEL X3 (300KHZ)
CHANNEL Y3 (300KHZ)
CHANNEL X4 (600KHZ)
CHANNEL Y4 (600KHZ)

FIG. 5.24 ZNCH

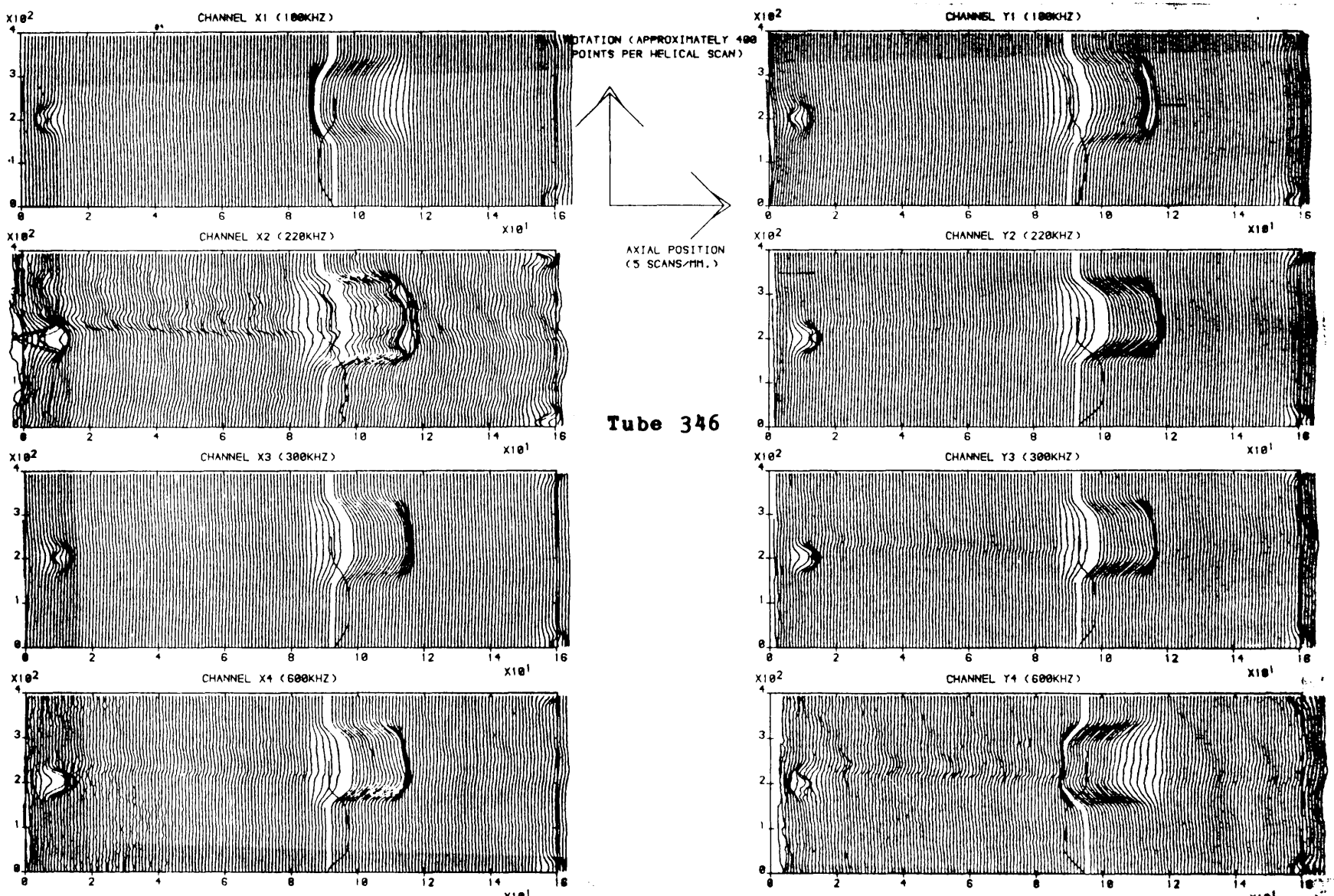



FIG. 5.25 PIGA AND A HOLE

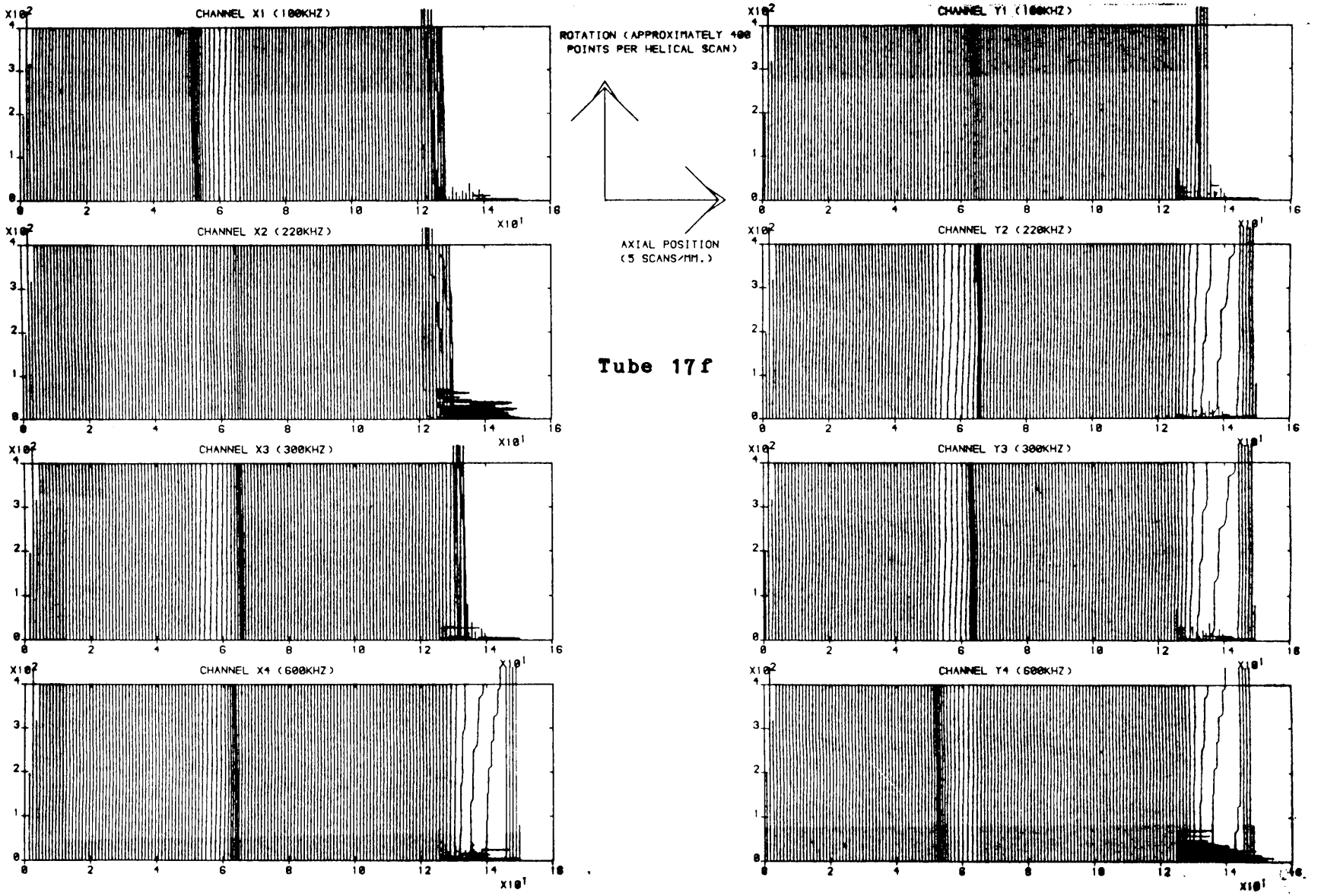



Fig. 5.26 Thinning

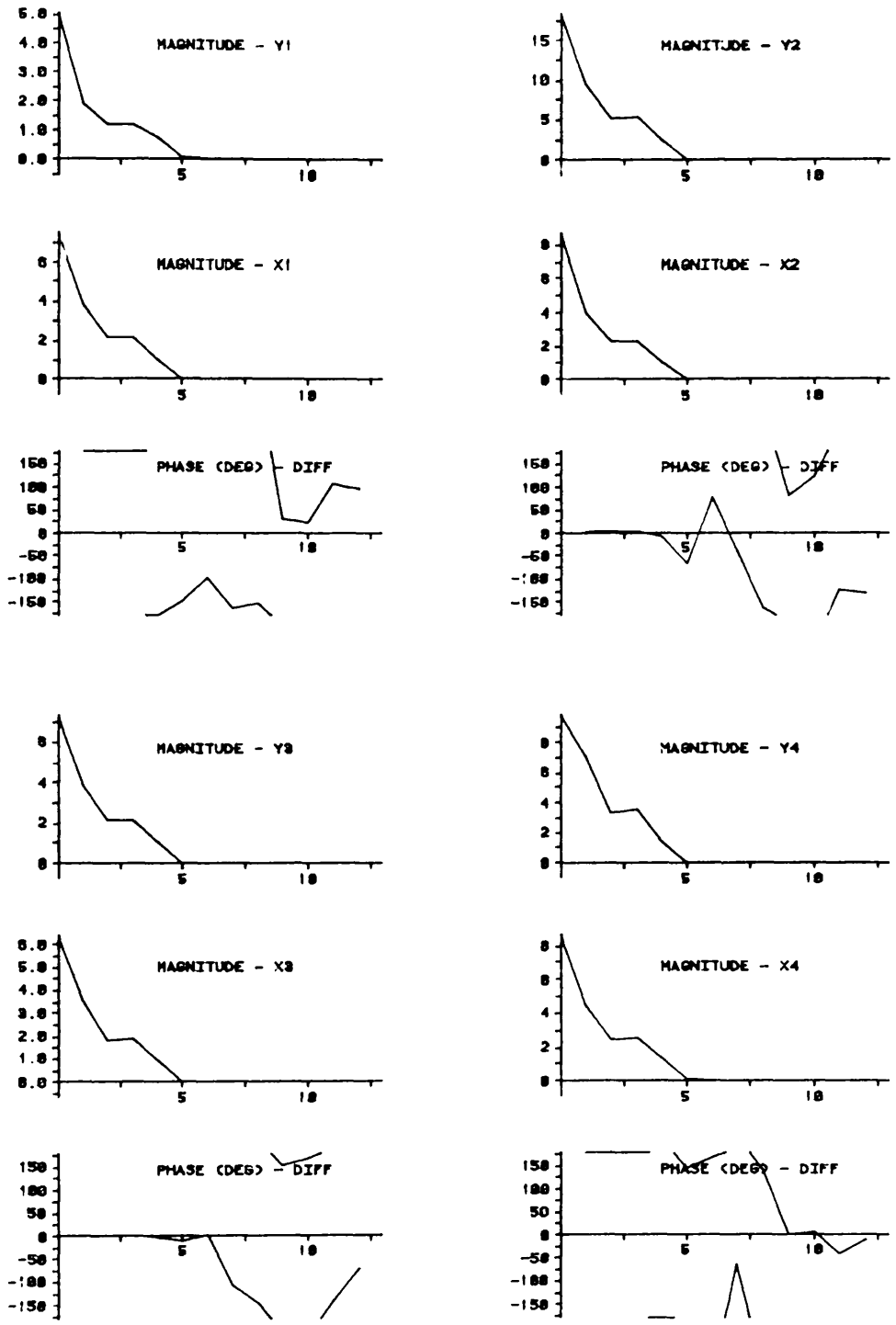



FIG. 5.27

CENTRAL SCAN FROM PATCH OF 80% THRU WALL IGA

INPUT FILE DP0:TUB346.RAW

AMP SPECTRUM OF SCAN 100 THETA = 125 TO 350

PLOTTED ON 11-AUG-84 AT 18:01

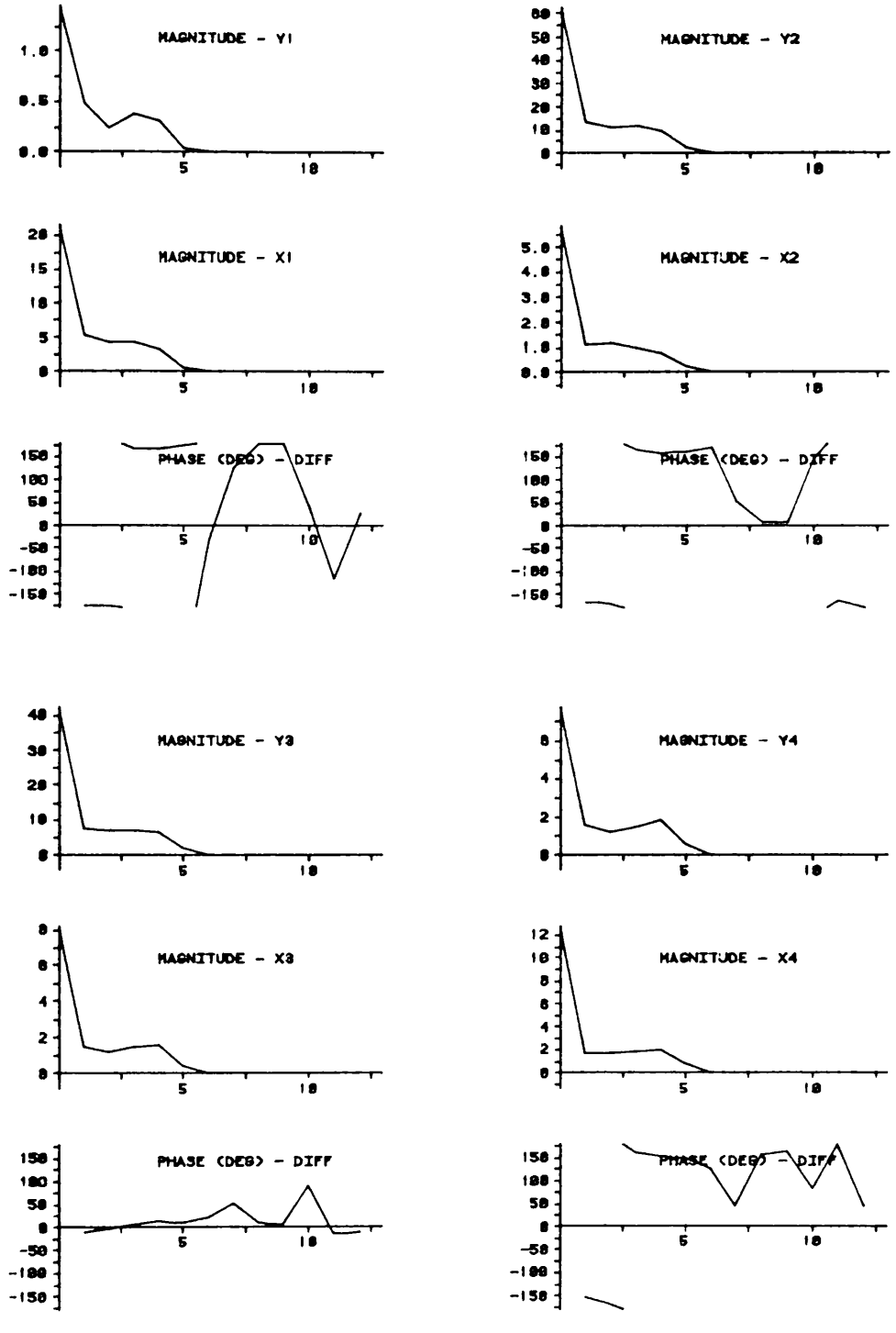



FIG. 5.28

CENTRAL SCAN THROUGH 80% THRU WALL CIRC. NOTCH

INPUT FILE DP0:TUB15C.RAW

AMP SPECTRUM OF SCAN 52 THETA = 130 TO 280

PLOTTED ON 11-AUG-84 AT 17:43

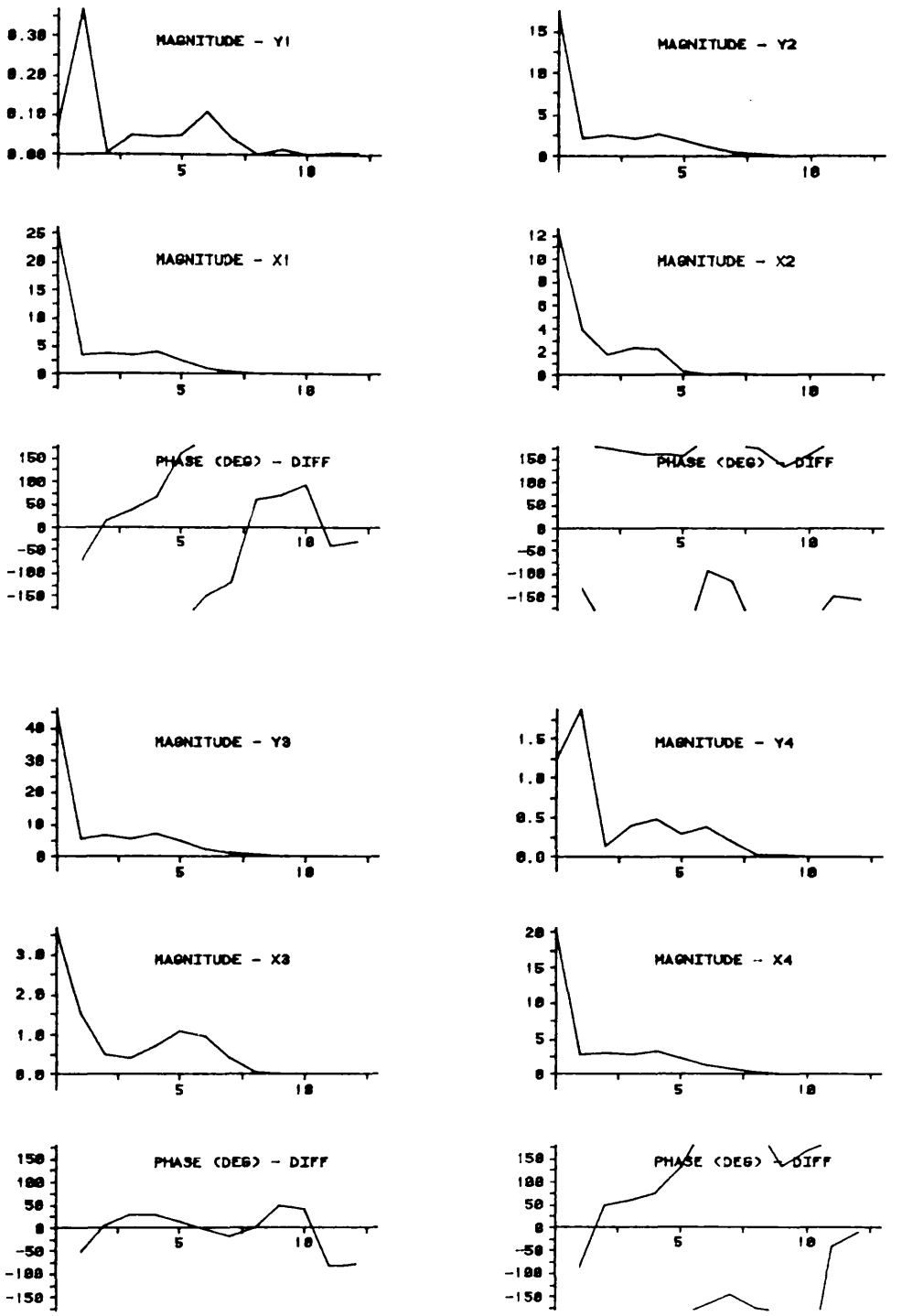



FIG. 5.29

78% THRU WALL AXIAL NOTCH, CENTRAL SCAN

INPUT FILE DP0:TUBE5F.RAW

AMP SPECTRUM OF SCAN 56 THETA = 175 TO 275

PLOTTED ON 11-AUG-84 AT 17:52

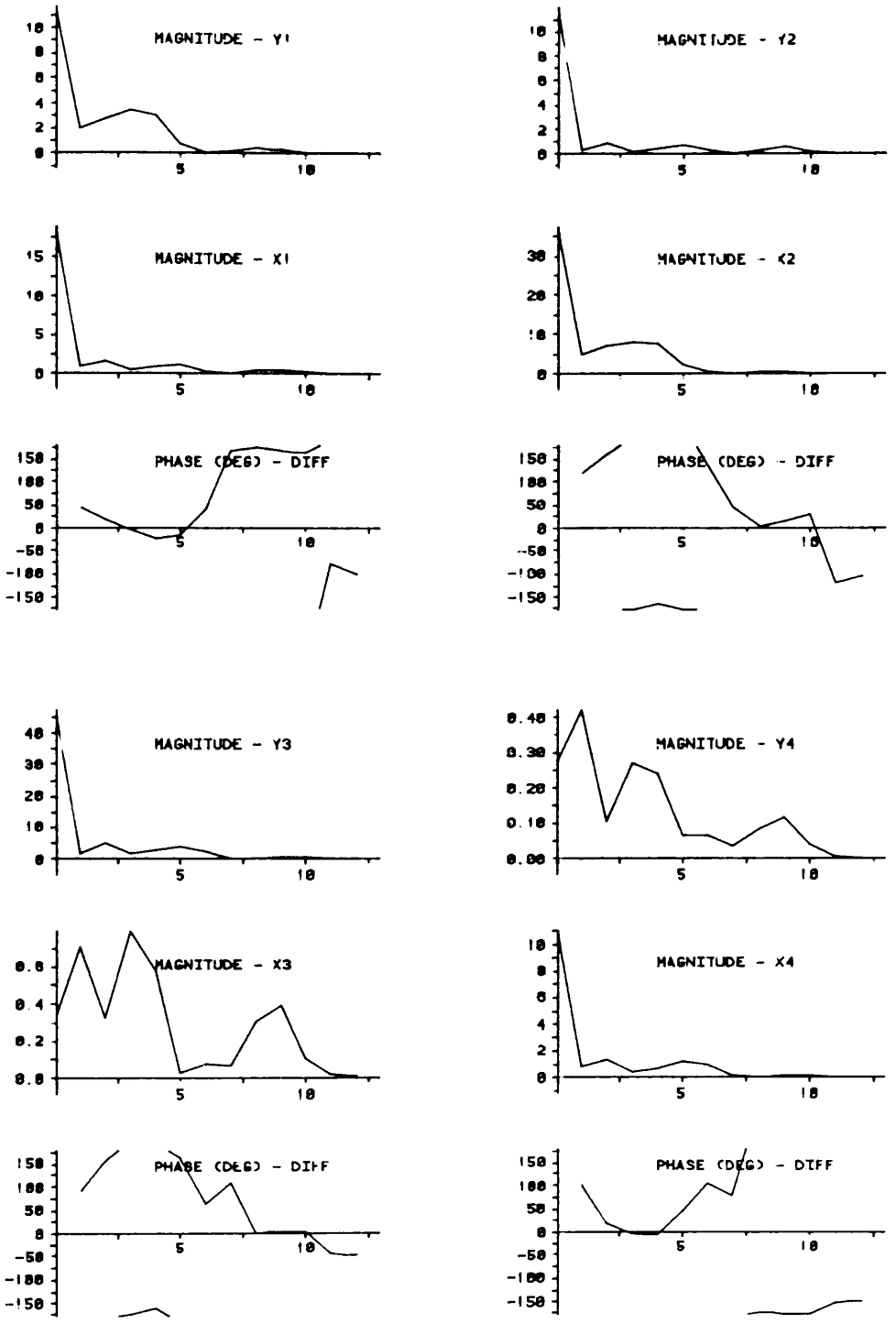



FIG. 5.30

FLAT BOTTOMED HOLE

INPUT FILE DP0:TUB342.RAW

AMP SPECTRUM OF SCAN 10 THETA = 160 TO 260

PLOTTED ON 12-AUG-84 AT 13:16

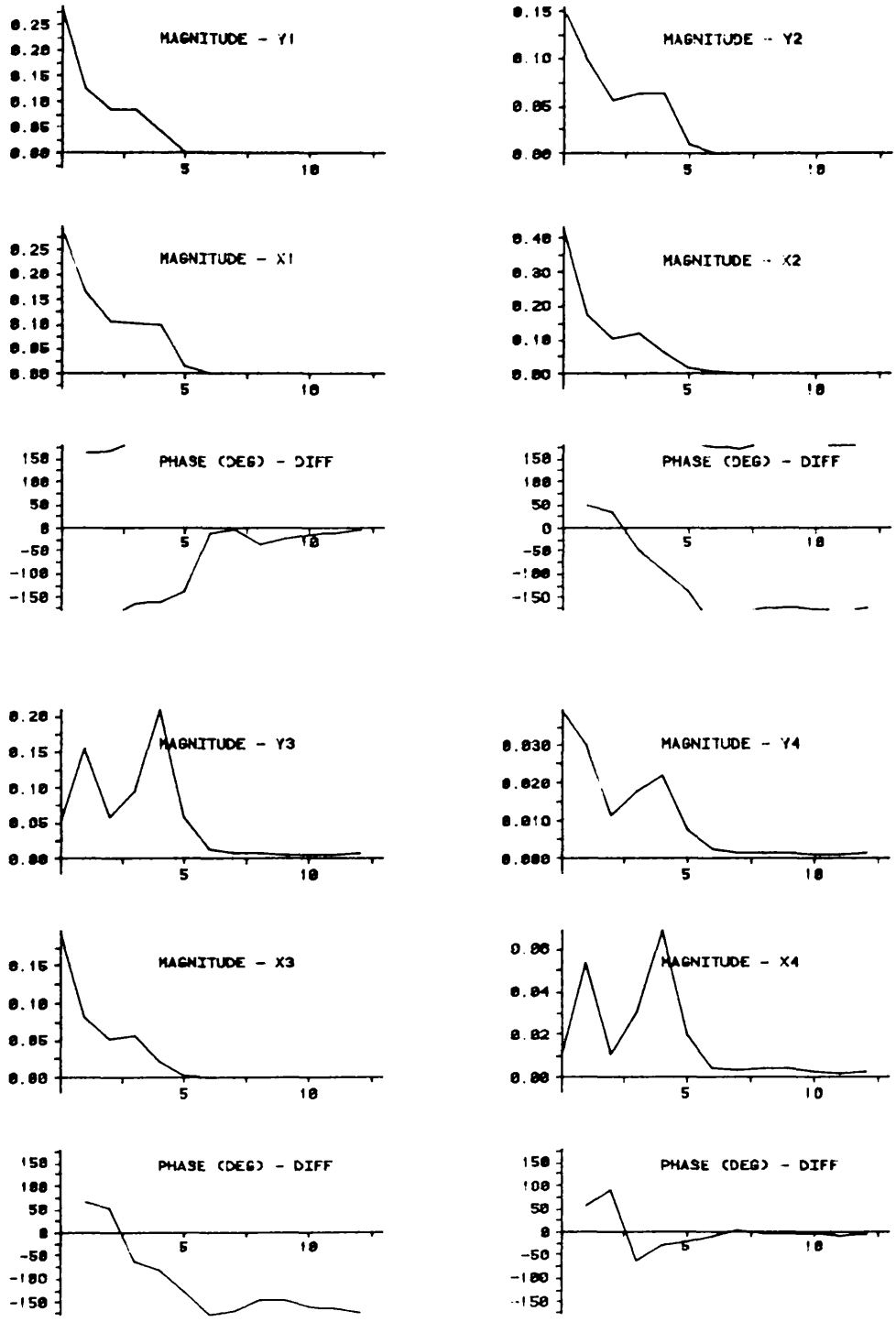



FIG. 5.31

CLEAN TUBING

INPUT FILE DP0:TUB342.RAW

AMP SPECTRUM OF SCAN 40 THETA = 100 TO 350

PLOTTED ON 12-AUG-84 AT 13:31

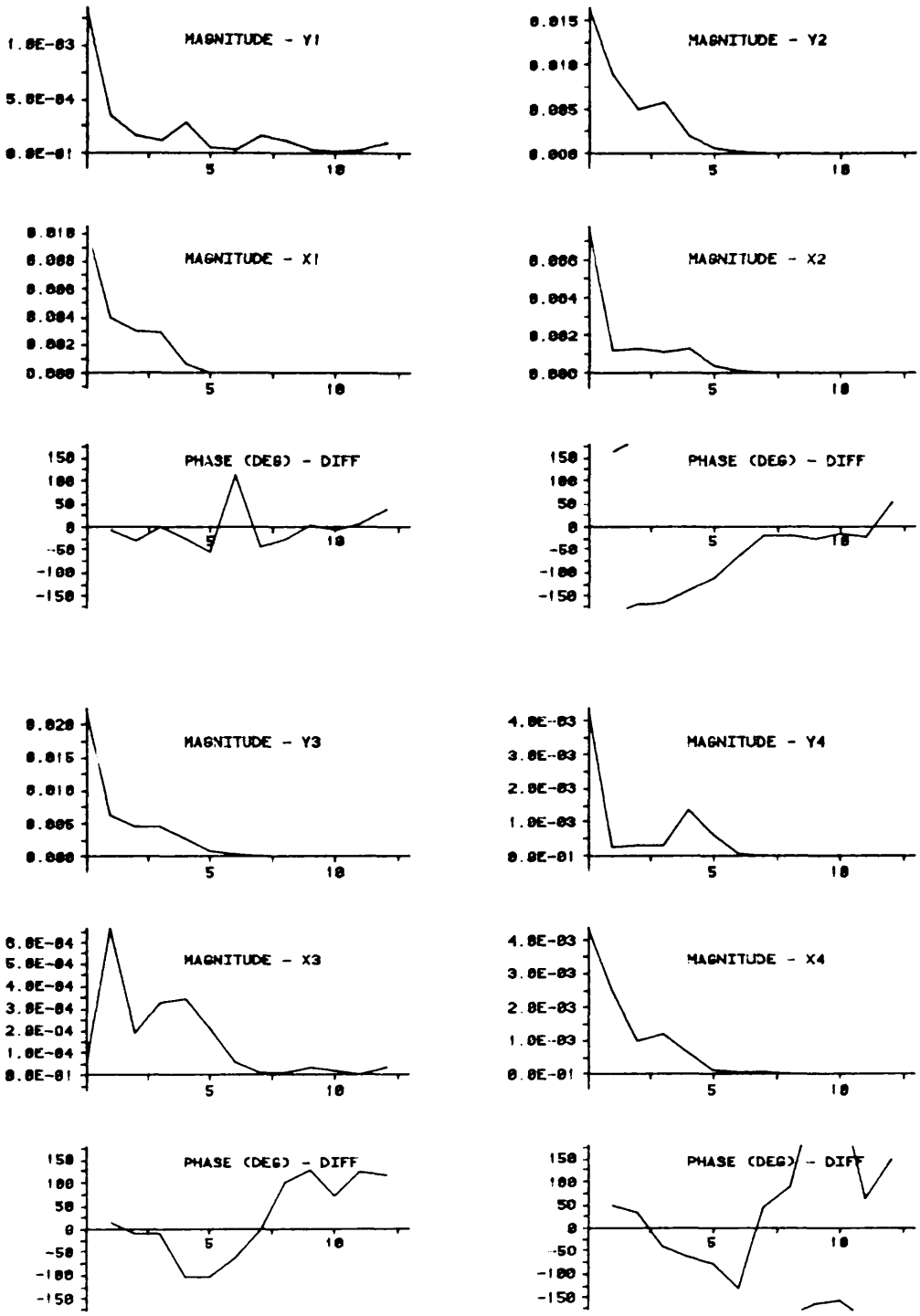


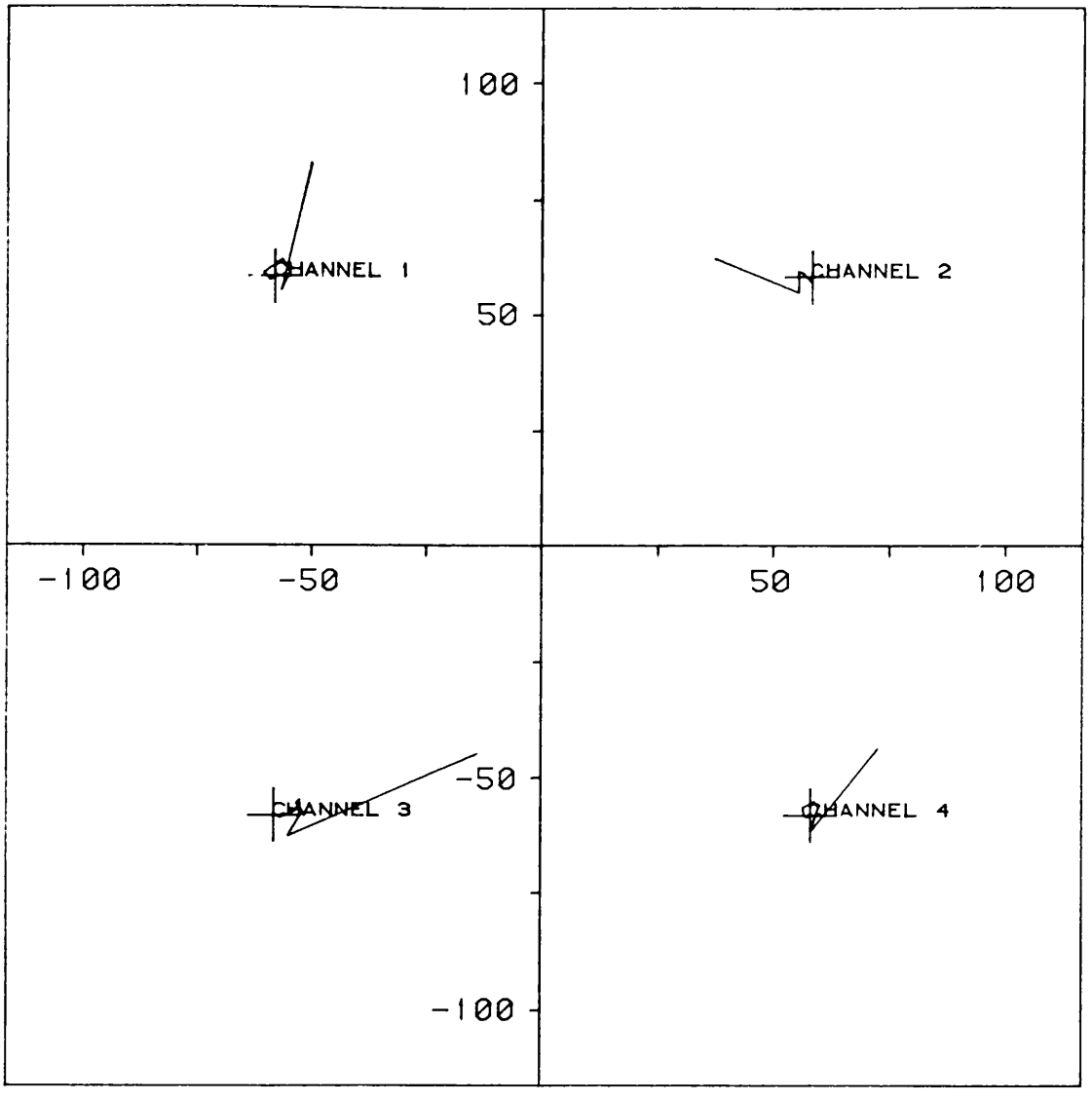

FIG. 5.32

48% THRU WALL UNIFORM THINNING, CENTRE OF BAND

INPUT FILE DP0:TUB18F.RAW

AMP SPECTRUM OF SCAN 65 THETA = 100 TO 350

PLOTTED ON 12-AUG-84 AT 14:21




SPCROT RUN ON 30-AUG-84 AT 15:43
LINE 56 POINTS175 TO 275 FILE DP0:TUBE5F.RAW
BLEMISH CLASS ZNCH

FIG. 5.33 AN ALTERNATIVE REPRESENTATION OF THE MAGNITUDE PLOT IN FIG. 5.29

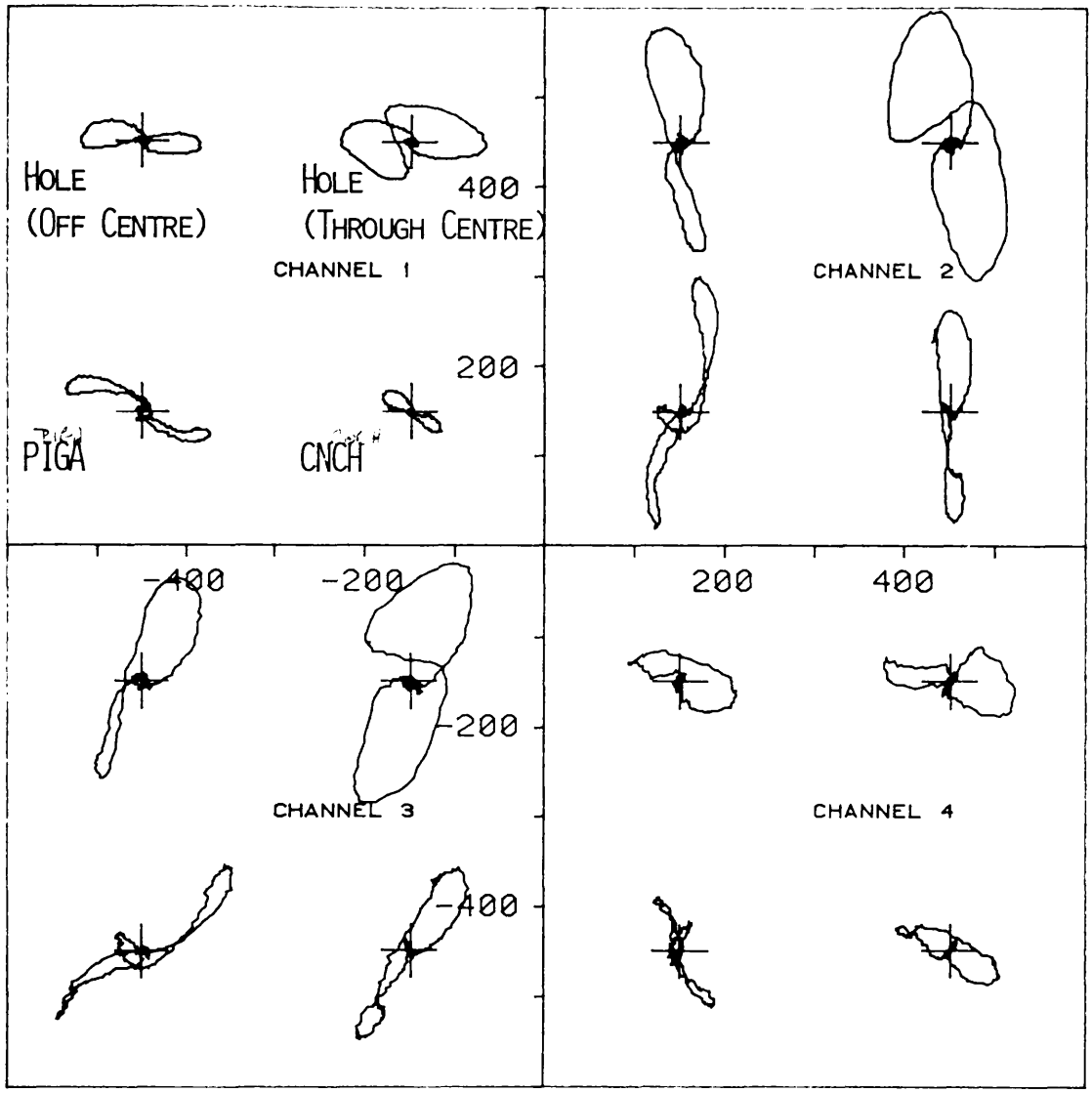



FIG. 5.34 DIFFERENTIATED SCANS

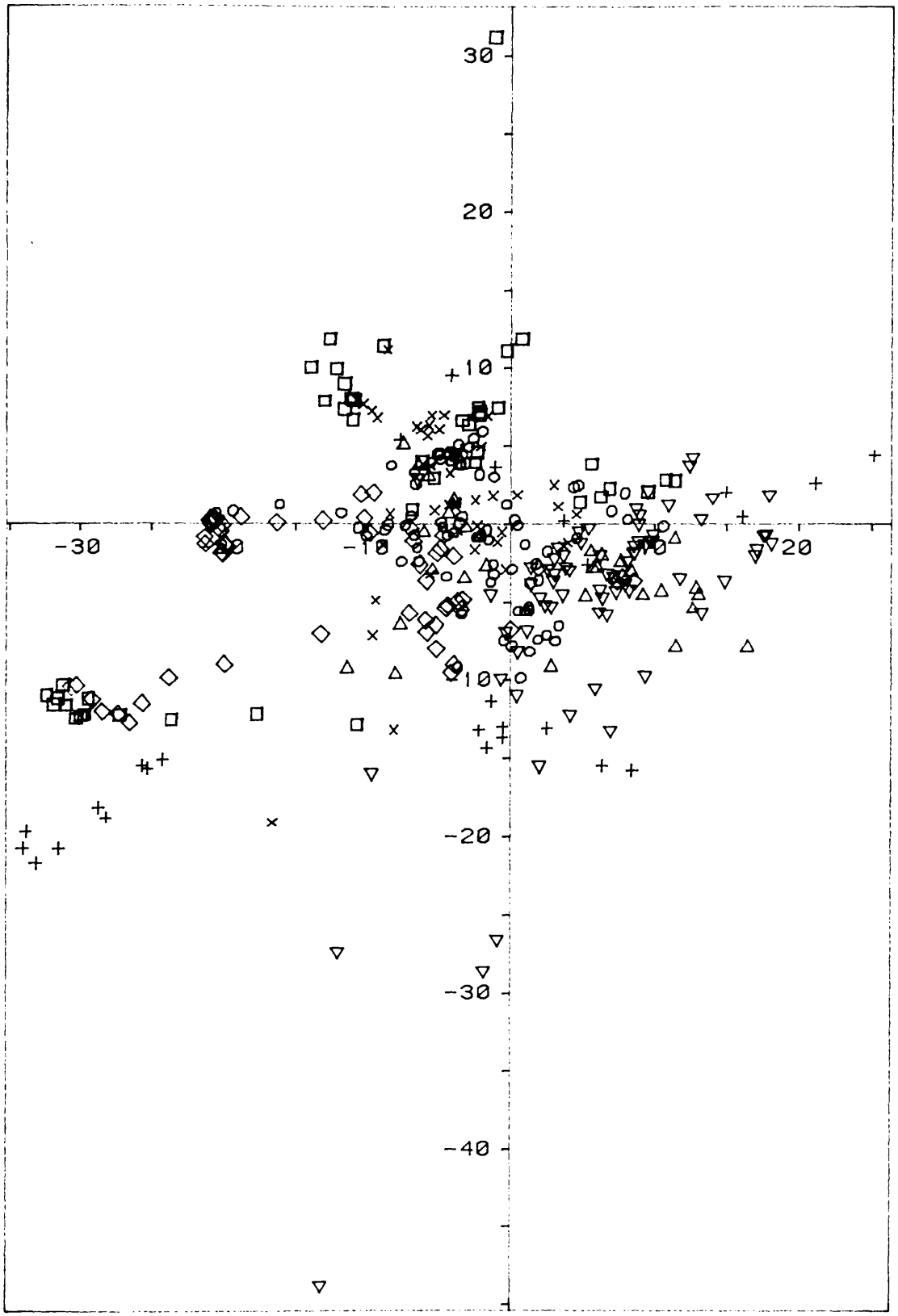
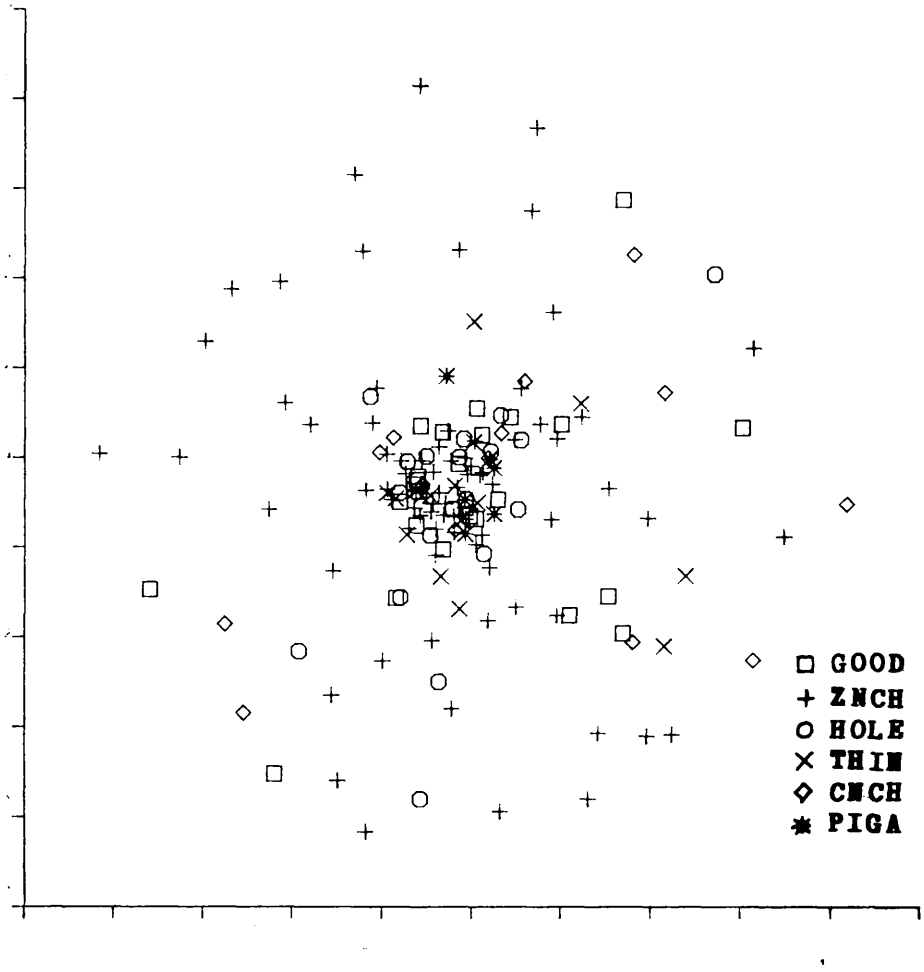


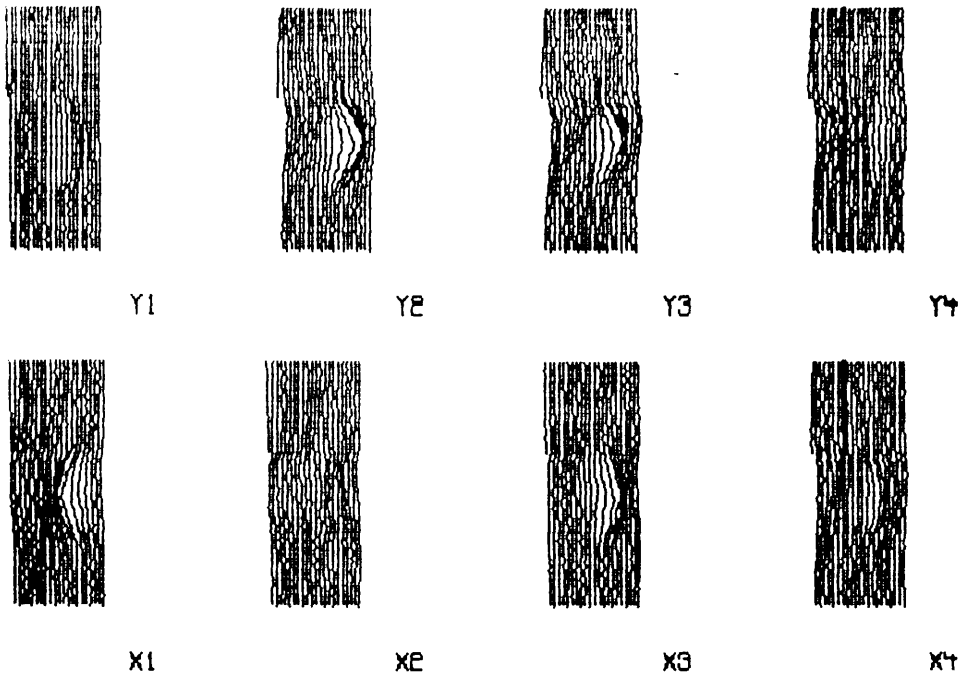

FIG. 5.35

CPLOT

HELICAL SCAN TEST SET, WITH CIRC. NOTCH CLASS IN BOLD, FEATURE TRANSFORMATION TRAINED UPON CNCH VS. ALL OTHER DEFECT TYPES.




5.36 A NON LINEAR MAPPING OF THE HELICAL SCAN SAMPLE TRAINING SET.




TUPLOT RUN ON 24-DEC-85 AT 03:20:16 FROM USE:TUB14C.RAW
PLOTTED LINES 1 TO 22 (SPACE FOR -6 TO 49)
SCALE: 1 SCAN INTERVAL = 20.0 ADC INCREMENTS

**fig 5.37(a)** Tube 14c relogged at approx. 55 sample points /scan, in order to achieve a rectangular sampling grid. The tubeplot has been scaled by selection of the plotting space (-6 to 49, see above) to preserve this equality of resolution in both directions. Comparable plots are 5.23 (old PDP11/45 work) and 5.37(c) (Medelec UV scope plot to confirm accuracy of logged plot).

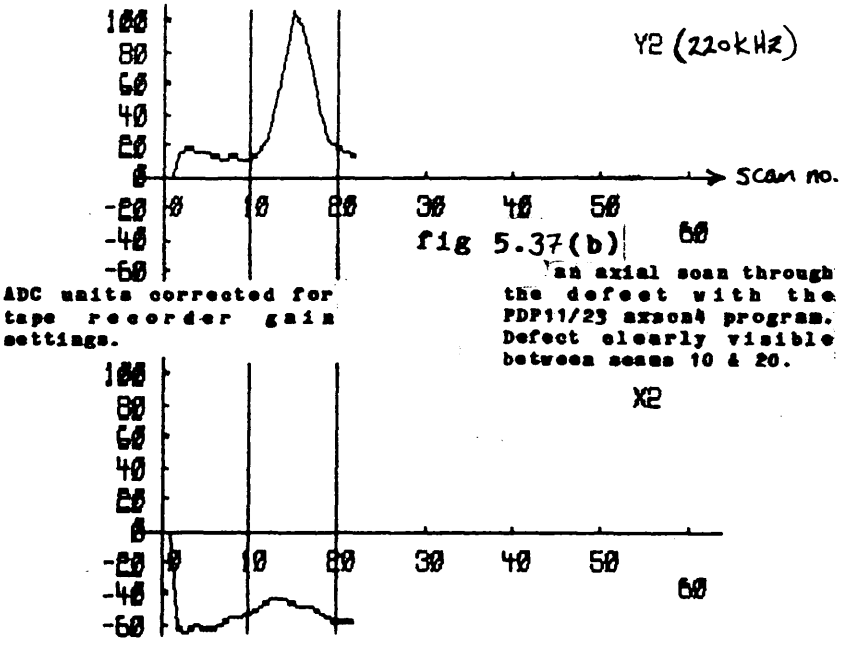
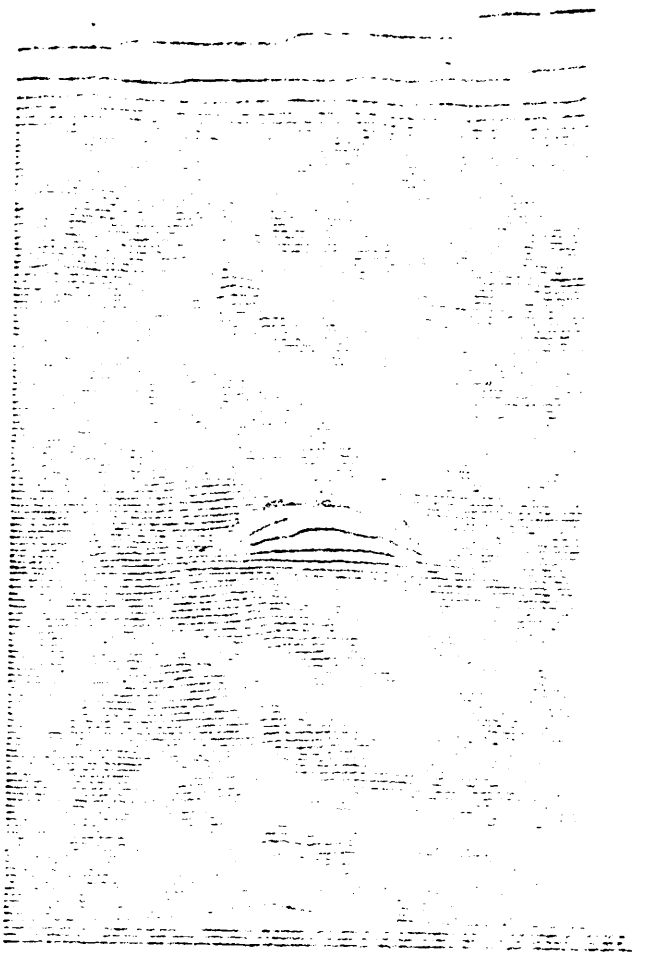



fig 5.37(b) An axial scan through the defect with the PDP11/23 axscn4 program. Defect clearly visible between scans 10 & 20.

PROGRAM AXSCN4: AXIAL SCAN PLOT: 1 CHANNEL PAIR
INPUT FILE USE:TUB14C.RAW
THETA 30 TIME SERIES (VERT SCALE = A.D.C. UNITS)
PLOTTED ON 24-DEC-85 AT 03:28 (+128 = 1 VOLT)



TUBE No. C14 FREQUENCY 220

SCAN No. 1 PHASE 0

~~X~~/Y GAIN 20

POSITION MEDELEC GAIN 1V/DN

~~DIRECT~~/TAPE ATTENUATION 1

DATE 19/4/84

fig 5.37(c) A plot of channel $y_2$(220kHz) made during the logging of tube14c to the Racal tape (to be compared with figs 5.23 & 37(a)).

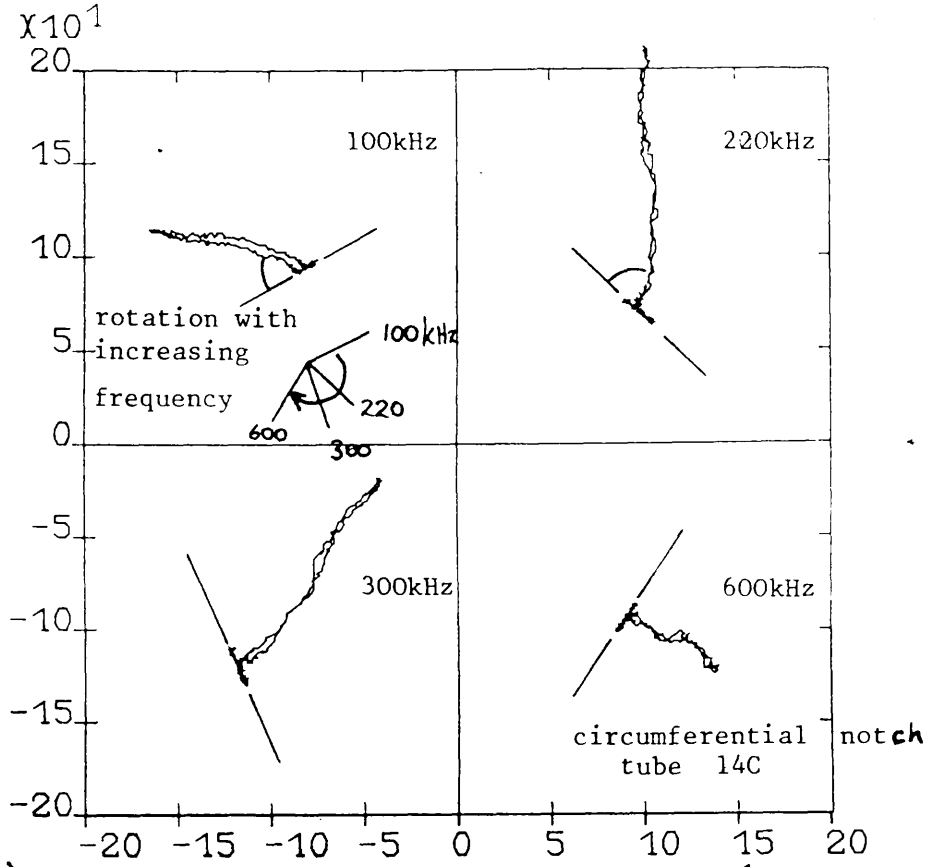



**fig. 5.38(a)** Characteristic Pilger directions at the 4 inspection frequencies. The defect is a 56% through wall Cnch (tube 14c, scan 14). Contrary to what would be expected with accepted eddy current theory, the Pilger loci rotate clockwise with increasing frequency.

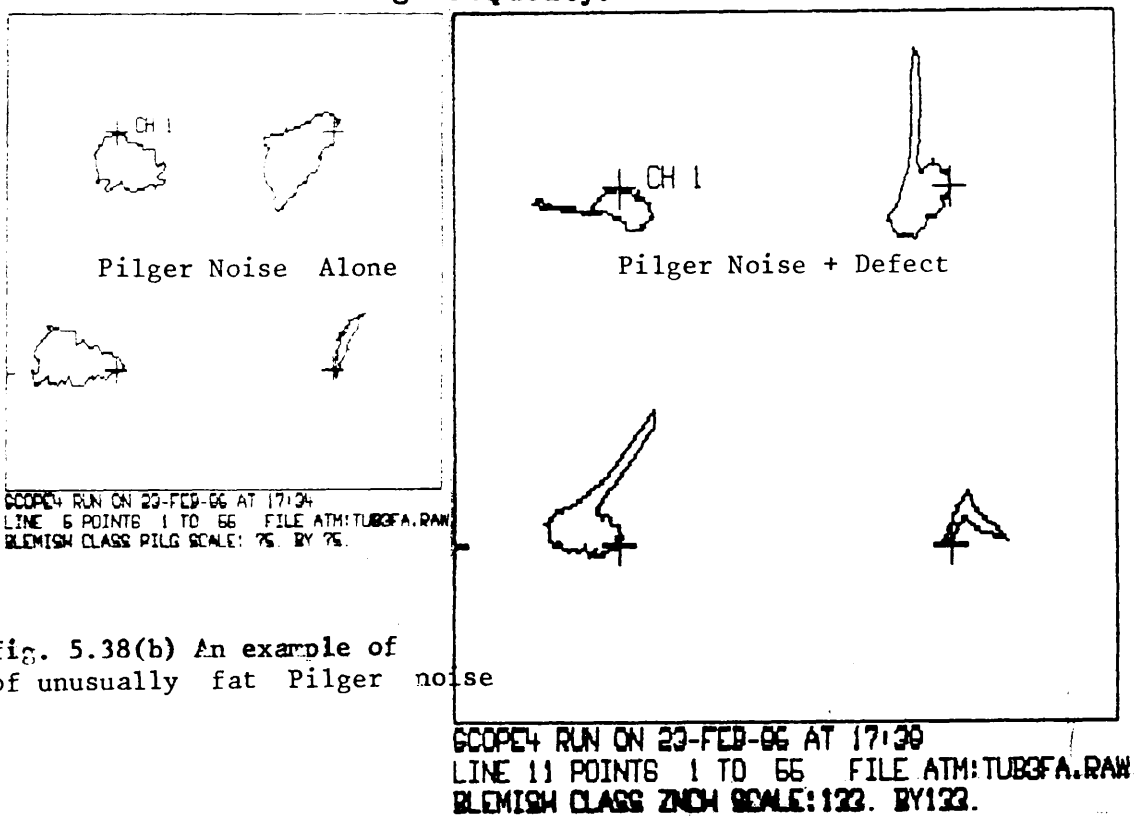



fig. 5.38(b) An example of of unusually fat Pilger noise

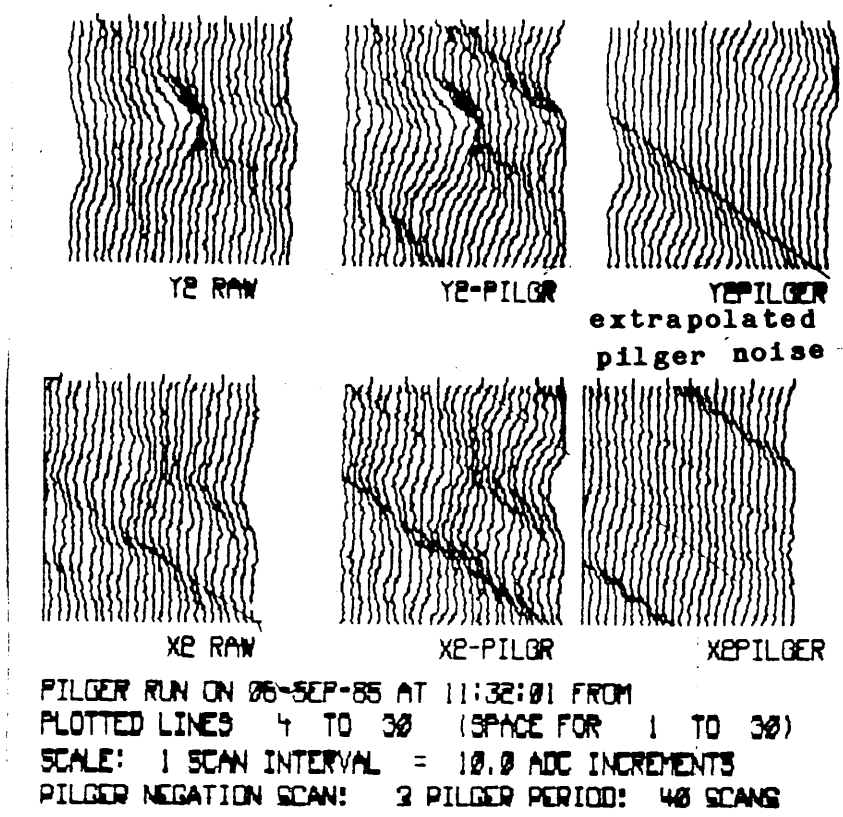



**fig 5.39** The results of picking one scan (the first one) and rotating it at the rate at which the Pilger noise advances around the tube wall, as determined by inspection (1 rev. every 52 scans or 52mm.) As can be seen from the central pair of plots, the method was not very successful. This was found to be because the Pilger noise did not advance angularly in a sufficiently regular fashion for linear extrapolation, such as this method, to work.

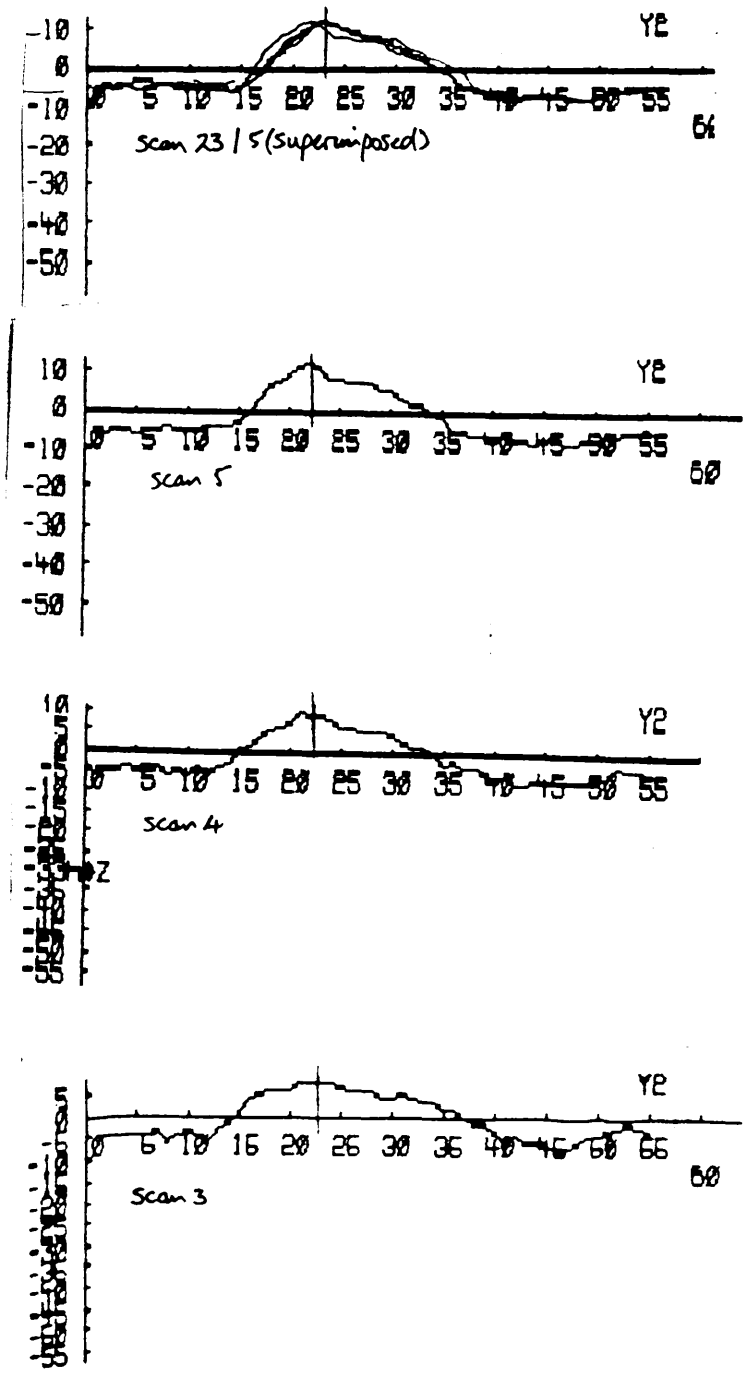



**fig 5.40** Scans 3,4,5, and 23 from channel $y_2$ of the tube in fig 5.39, with 4 and 5 superimposed upon 23, to demonstrate the lack of advancement in the pilger noise. Despite this, the ripples due to the Pilger noise are clearly visible as advancing around the tube wall in the raw $x_2$ and $y_2$ plots in fig. 5.39. From a series of plots such as this, it was concluded that the rotation of the Pilger noise could not be easily tracked by a peak following approach; nor was linear extrapolation sufficient to synthesis it.

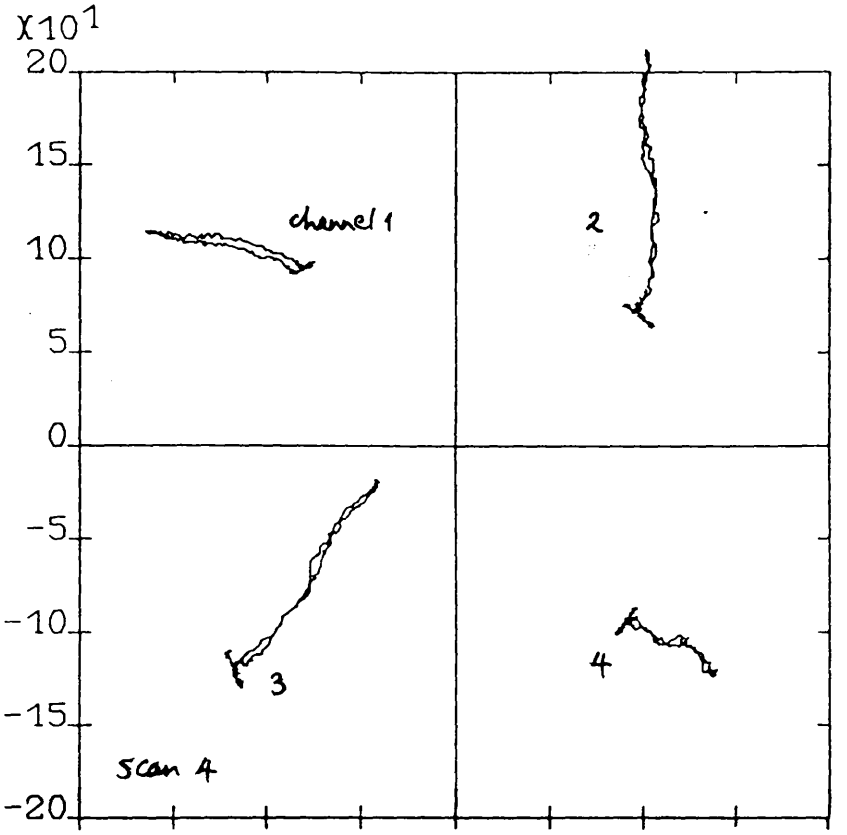
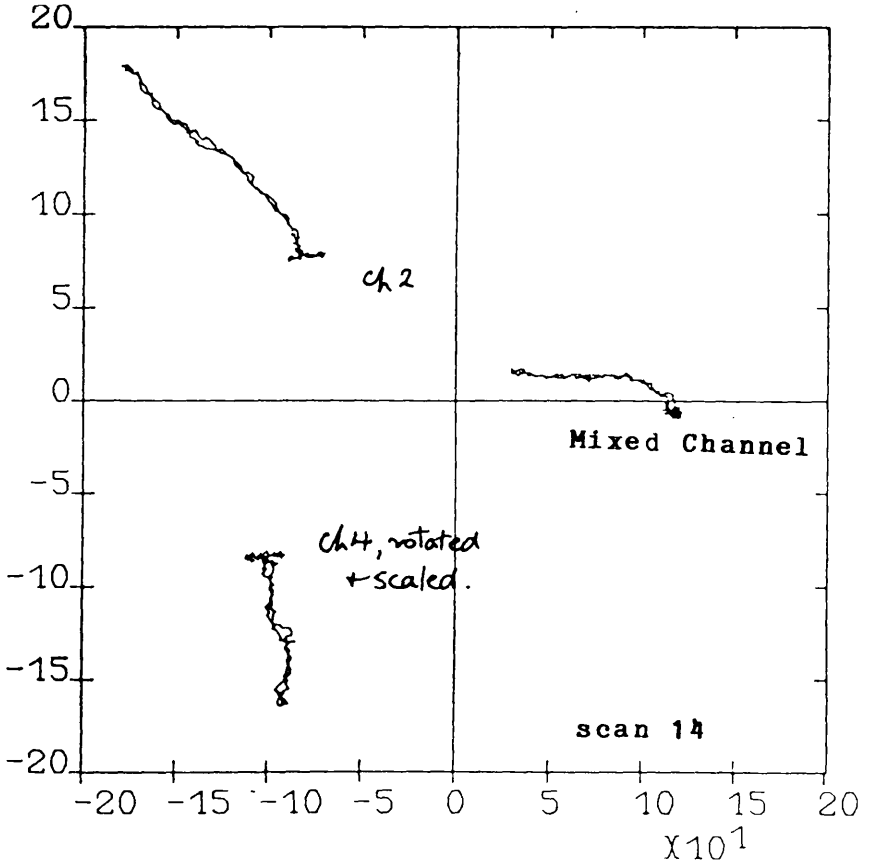



**fig 5.41** Mixing of channel 2 and 4 according to the method of computational Pilger nulling described in method (vii). Channel 4 in scan 4, tube 14c has been rotated and scaled to match channel 2, and the angular rotations and gains required have then been re-applied to scan 14 from the same tube (below), reducing apreciable length Pilger noise locus to a point, and removing spurious fatness from the defect locus.



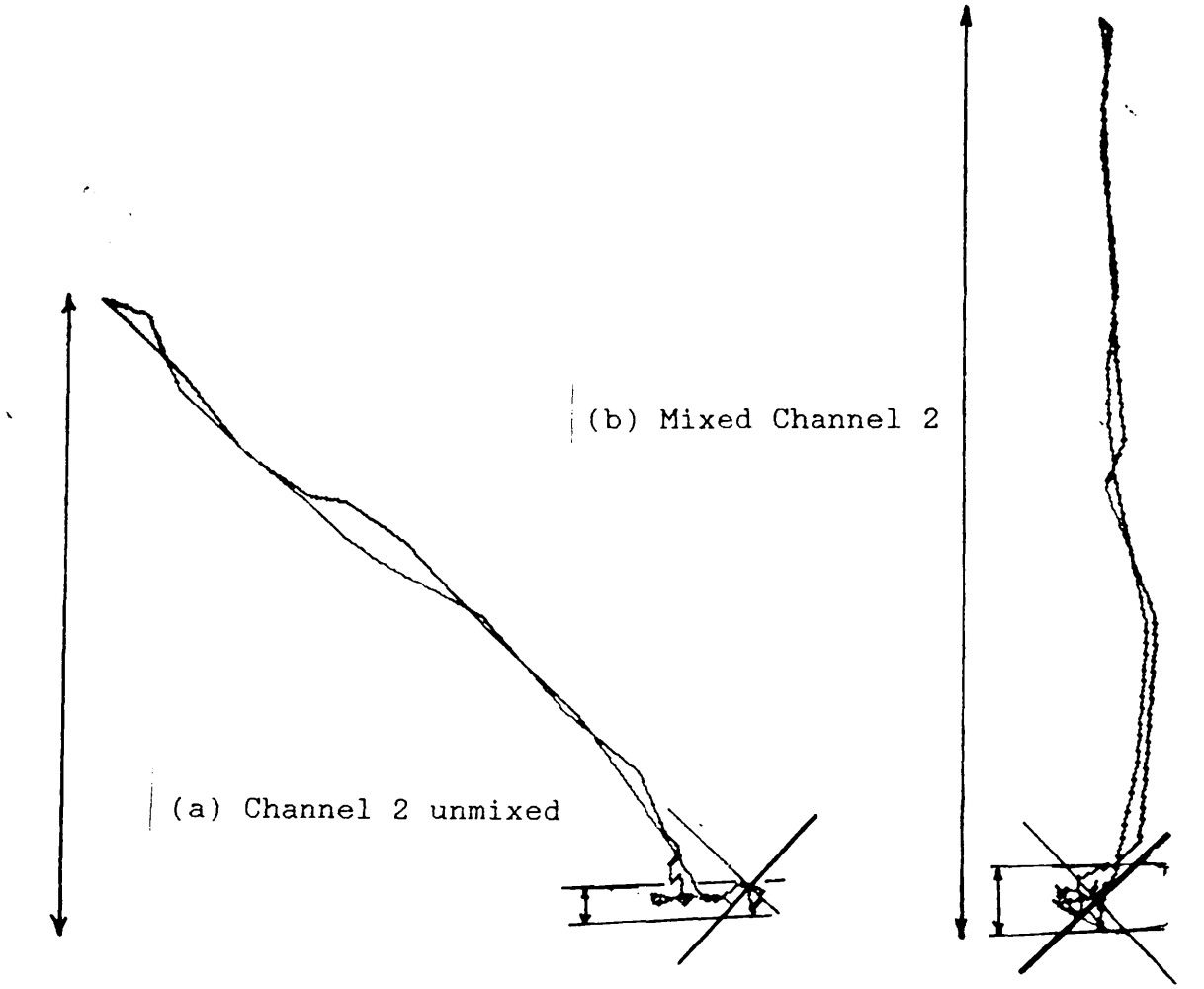
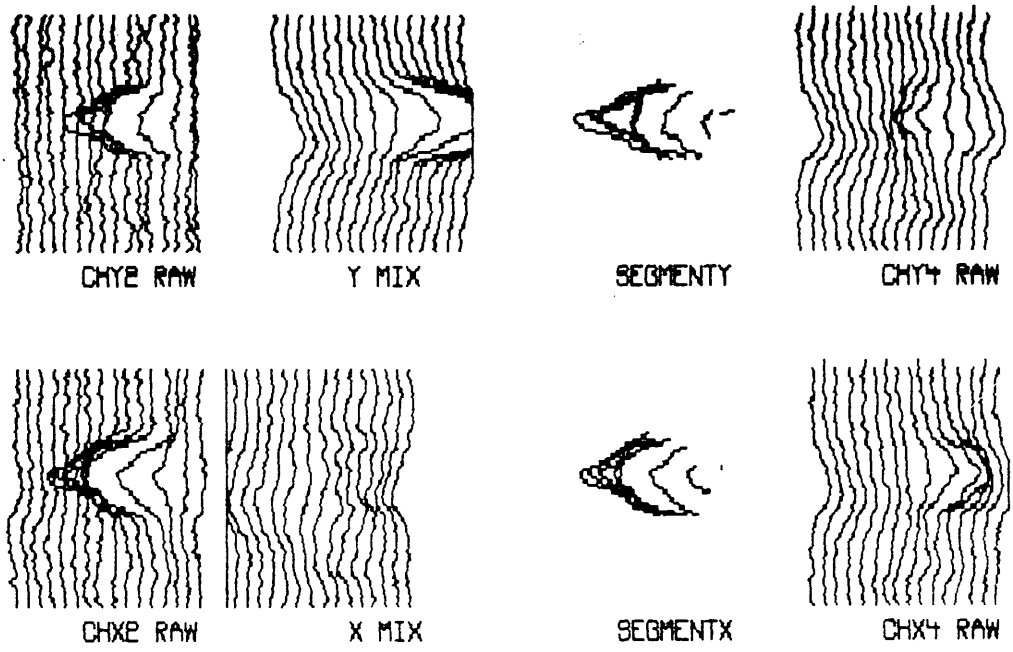



Fig. 5.42 Enlarged plots of scan 14, tube 14c (a) unmixed and (b) mixed with channel 4 as shown in fig. 5.41. The reduction of the mixed locus to one with coincident outward and inward bound trajectories is clear to see and this is desirable given that the defect is a symmterical V-section notch, but a 30% reduction of the signal to Pilger noise ratio from 17.5:1 to 12.5:1 is also apparent.




```
.RUN SEGPLT
SCAN FILSPC? TUB14C.RAW
DO YOU WANT TO USE PREVIOUSLY CALCULATED MIX? N
LINE NUMBER FOR MIX ANULLMENT? 4
SEGMENTION THRESHOLD VALUE? 10.0
PLOTTER MIN & MAX Z COORDS? 1,20
ACTUAL START & FINISH Z COORDS? 5,20
PLOT ALL CHANNELS? Y
SIGNAL VALUE FOR UNIT SCAN HEIGHT? 10.0
IS THE ABOVE CORRECT?Y
 PEN, <CR> TO CONTINUE:
```

**fig. 5.43** Graphical output from a typical run of the segmentation program using method (vi). The raw channel 2 signals are plotted ("segment x" & "segment y") whenever the magnitude of the mixed x and y channels exceeds an annular threshold (10.0) which just encloses the Pilger noise remaining after mixing.

**Table 5.1 The Impedance Plane Featureset**

All definitions of all features assume that the origin is defined as the 1st point of the defect segment (as defined either in the feature extractor driver data files, or by the segmentation algorithm passing the defects to the feature extractor) to avoid including d.c. shift with respect to the origin that the locus might have.

For calculation of features 3-6 and 9-15, the loci is rotated so that it lies with its point of maximum magnitude along the +ve x-axis and a 5th order Chebyshev polynomial is fitted to the outbound trajectory using the NAG library curve fitting routine,.

* 1. $\emptyset$, the overall angle, measured anti-clockwise with respect to the horizontal, at the point of maximum magnitude within the locus,

* 2. $R_{max}$, the loci maximum magnitude,

* 3. the sum of the vertical differences between the inbound and outbound paths, normalised with respect to maximum signal magnitude, $R_{max}$, and the number of points in the lobe, np, so that the feature really measures the "fatness" of the locus,

* 4. As for feature 3, but biased in favour of the area enclosed near to the outermost tip:

* 5. As for feature 3, but biased in favour of area near to the origin:

* 6. As for feature 3, but biased in favour of the area near to the middle:

7 and 8. The angle and radius of the d.c. offset of the first point in the locus

* 9-15. The coefficients of the curve fitted to the outbound trajectory,
(A 5th order polynomial in x was thought sufficient to describe the shape of any loci.)

* 16. A boolean values feature, based upon feature 3,

to indicate if the locus was twisted and thus the outward path crossed over the returning trajectory.

| | | |
|---|---|---|
| * 17. $\emptyset_{ch2}-\emptyset_{ch1}$ | * 20. $R_{ch2}/R_{ch1}$ | "*" = feature selected for use at classifier stage. |
| * 18. $\emptyset_{ch2}-\emptyset_{ch3}$ | * 21. $R_{ch2}/R_{ch3}$ | |
| * 19. $\emptyset_{ch2}-\emptyset_{ch4}$ | * 22. $R_{ch2}/R_{ch4}$ | |

**Table 5.2 The Spectral Featureset**

* **1. and** * **2.** The magnitude and phase of the d.c. ordinate
* **3.** Total spectral power, i.e. the sum of the magnitude of the ordinates from d.c. to fsamp/10 (ordinates 1 to 13 in the 128 point power spectral density periodogram)
* **4.** Concentration of spectral power in the higher frequency ordinates, measured as
* **5.** Ordinate 7/Ordinate 5, assuming that the ordinates are numbered such that the d.c. ordinate is 1.
* **6.** The ratio of power under the bottom two ordinates, to the total spectral power.

**7.** The ratio of area in the x spectra to area y spectra,

**8-14.** Defined as per features 1.-7., but applied to the signals from the 220kHz x and y channels.* (8, 10, 11, 12, 13 used)

**15.-21.** Defined as per features 1.-7., but applied to the signals from the 600 kHz x and y channels.* (15, 17, 18, 19, 20 used)

**22.** Phase Difference between the 100 kHz and 220 kHz channels' d.c. ordinates.

**23.** Phase Difference between the 220 kHz and 600 kHz channels' d.c. ordinates.

**24.** Ratio of spectral areas under 1st 13 ordinates for 220kHz and 100 kHz channels.

* **25.** Ratio of spectral areas under 1st 13 ordinates for 600 kHz and 220 kHz channels.

**26.** Ratio of magnitudes of 100kHz and 220 kHz channels' d.c. components.

* **27.** Ratio of magnitudes of 100kHz and 220 kHz channels' d.c. components.

"*" = feature selected for use at classifier stage.

**Table 5.3: Multifrequency Classification Results**

**Table 5.3(a):** Results with 1-NN classifier on the best subset of **impedance** plane features (results obtained with the same set of points with a **single stage** 7-class classifer are given in the last column).

| | | Class Assigned | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Good | Znch | Hole | Thin | Cnch | Piga | Uiga | total | lin. class |
| True Class | Good | 14 | 0 | 0 | 5 | 0 | 3 | 8 | 14/30(47%) | 57% |
| | Znch | 0 | 39 | 24 | 0 | 2 | 3 | 4 | 39/72(54%) | 56% |
| | Hole | 0 | 3 | 22 | 0 | 1 | 1 | 0 | 22/27(81%) | 22% |
| | Thin | 1 | 0 | 0 | 26 | 10 | 0 | 6 | 26/31(84%) | 32% |
| | Cnch | 1 | 9 | 3 | 0 | 26 | 2 | 3 | 26/44(61%) | 71% |
| | Piga | 0 | 31 | 0 | 0 | 1 | 22 | 1 | 22/55(29%) | 50% |
| | Uiga | 1 | 2 | 0 | 17 | 0 | 2 | 60 | 60/82(73%) | 22% |

overall correct classification rate: 61.5%

**Table 5.3(b)** Results with a 1-NN rule classifier on the best spectral domain features, again with the linear classifier results in the last column.

| | | Class assigned | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Good | Znch | Hole | Thin | Cnch | Piga | total | lin. class. |
| True Class | Good | 21 | 0 | 0 | 8 | 1 | 0 | 21/30(70%) | 13/30 |
| | Znch | 1 | 16 | 5 | 0 | 5 | 0 | 16/27(59%) | 15/27 |
| | Hole | 0 | 2 | 2 | 0 | 1 | 2 | 2/7(29%) | 2/7 |
| | Thin | 4 | 0 | 0 | 27 | 0 | 0 | 27/31(87%) | 28/31 |
| | Cnch | 7 | 2 | 0 | 3 | 9 | 1 | 9/22(41%) | 2/22 |
| | Piga | 0 | 1 | 1 | 0 | 20 | 53 | 53/75(71%) | 16/75 |

total correct classification rate: 59.4%

**Table 5.3(c)** Separability of Banded and Non-banded defects 1-NN rule classifier on helical scan spectral features
Class 1: Non-banded defects: Piga, Znch, Cnch & Holes: 137 scans
Class 2: Banded "defects": Good tubing,Thinning & Uiga: 218 scans

| | | Assigned Class | |
|---|---|---|---|
| | | Nonbanded | Banded |
| True Class | Nonbanded | 132 | 5 |
| | Banded | 28 | 190 |

total correct classification rate: 90.7%

**Table 5.3(d)** Separability of Non-axisymmetric Defects 1-NN rule classifier on impedance plane features helical scan.

| | | Assigned Class | | | | |
|---|---|---|---|---|---|---|
| | | Hole | Znch | Cnch | Piga | |
| Real Class | Hole | 40 | 25 | 3 | 4 | 55% |
| | Znch | 3 | 22 | 1 | 1 | 81% |
| | Cnch | 7 | 3 | 30 | 4 | 68% |
| | Piga | 48 | 0 | 0 | 26 | 35% |

Overall correct classification rate: 60%

**Table 5.4 Relationships between Graphical Plots in this Chapter**

| Class | Helical Scan Plots | Axial Scan Plots | Z-plane Rotating Scan Plots |
|---|---|---|---|
| Piga | 5.1 60%,100,tube 346 | 5.7 40%,tube 342 | 5.13 |
| Cnch | 5.2 80%,52,tube 15c | 5.8 56%,tube 14c | 5.14 |
| Znch | 5.3 78%,56,tube 5f | 5.9 78%,tube 5f | 5.12 78%,56,tube 5f & 5.15 |
| Hole | 5.4 100%,10,tube 342 | no plot | 5.16 100%,10,tube 346 |
| Good | 5.5 n/a,40,tube 342 | 5.10 n/a,tube 9 | no plot (pointless) |
| Thin | 5.6 48%,65,tube 18f | 5.11 48%,tube 18f | no plot (pointless) |

| Class | Z-plane Axial Scan Plots | t8plots | Spectrum Plots |
|---|---|---|---|
| Piga | 5.17 40%,200,tube 342 | 5.22 tube | 5.27 60%,100,tube 346 |
| Cnch | 5.18 56%,200,tube 14c | 5.23 tube 14c | 5.28 80%,52,tube 15c |
| Znch | 5.19 56%,225,tube 5f | 5.24 tube 3fa | 5.29 78%,56,tube 5f |
| Hole | 5.20 100%,200,tube 346 | 5.22,23&24 | 5.30 100%,10,tube 342 |
| Good | no plot | 5.22,23&24 | 5.31 n/a,tube 342 |
| Thin | 5.21 48%,200,tube 18f | 5.26 tube 17f | 5.32 48%,65,tube 18f |

| Class | End-on Polar Plots | Synthetically Differentiated Plots |
|---|---|---|
| Piga | | 5.34 |
| Cnch | | " |
| Znch | 5.33 78%,56,tube5f | |
| Hole | | " |
| Good | | |
| Thin | | |

| Defect type and set membership | Depth, thru wall | Tube File name | Line Nos. Involved | Circum. Co-ords |
|---|---|---|---|---|
| Patch IGA | $HNO_3$ 16% | Tube 7 | 63-74 | 80-280 |
| (training) | $HNO_3$ 10% | Tube 8 | 82-105 | 160-360 |
| | Poly. 25% | Tube 343 | 70-100 | 75-300 |
| | Poly. 69% | Tube 346 | 90-115 | 125-375 |
| (testing) | $HNO_3$ 14% | Tube 9 | 805-105 | 160-360 |
| | $HNO_3$ 12% | Tube 10 | 80-106 | 100-350 |
| | Poly. 54% | Tube 342 | 82-108 | 50-250 |
| ZNCH | 34% | Tube 2FB | 55-65 | 200-300 |
| | 78% | Tube 5F | 52-61 | 175-275 |
| | 75% | Tube 11 | 39-56 | 190-270 |
| | 31% | Tube 8FA | 34-67 | 50-250 |
| | 60% | Tube 9FA | 43-59 | 190-290 |
| | 59% | Tube 3FB | 45-58 | 190-270 |
| | 100% | Tube 6F | 62-72 | 175-275 |
| | 58% | Tube 3FA | 60-75 | 200-300 |
| | 72% | Tube 4F | 33-45 | 170-260 |
| | 69% | Tube 10F | 46-61 | 190-270 |
| | 56% | Tube 9FB | 37-52 | 180-270 |
| CNCH | 100% | Tube C01 | 45-50 | 180-280 |
| (training) | 100% | Tube C03 | 50-55 | 150-280 |
| | 50% | Tube C08 | 55-59 | 170-290 |
| | 100% | Tube C02 | 50-56 | 160-280 |
| | 45% | Tube C09 | 55-65 | 150-275 |
| | 80% | Tube 15C | 49-55 | 125-325 |
| Testing | 75% | Tube C06 | 55-60 | 170-270 |
| | 50% | Tube C11 | 50-58 | 190-300 |
| | 56% | Tube 14C | 60-70 | 150-325 |
| | 34% | Tube 13C | 60-70 | 175-325 |
| Uniform | 48% | Tube 18F | 57-75 | 1-400 |
| Thinning | 48% | Tube 17F | 58-65 | 1-400 |
| (THIN) | 32% | Tube 16F | 58-75 | 1-400 |
| | 16% | Tube 13F | 57-65 | 1-400 |
| HOLES | 100% | Tube 346 | 7-14 | 150-250 |
| (training) | 100% | Tube 343 | 2-9 | 150-250 |
| | 100% | Tube 353 | 5-12 | 175-260 |
| | 100% | Tube 352 | 152-158 | 160-260 |
| | 100% | Tube 342 | 8-14 | 150-250 |
| (testing) | 100% | Tube 354 | 152-160 | 150-250 |
| | 100% | Tube 26 | 105-115 | 160-275 |
| circumferentially | 81% | Tube 352 | 66-90 | 1-256 |
| uniform IGA | 27% | Tube 353 | 70-95 | 1-256 |
| (training) | 43% | Tube 26 | 39-61 | 1-256 |
| (testing) | 63% | Tube 24 | 22-56 | 1-256 |
| | 68% | Tube 354 | 68-100 | 1-256 |
| | 42% | Tube 106 | 72-85 | 1-256 |

Table 5.5 The complete set of defects in [illegible]

# 6 Conclusions and Further Work

In this first half of this chapter the main topics discussed are the reliability and extendability of the results, and the tasks necessary to develop the techniques described in chapters 3,4 and 5 for use outside of a laboratory. In the latter half, the software development and defect sample requirements for further experimentation are quantified.

## 6.1 Single Frequency Work Conclusions

### 6.1.1 Accuracy of Manual Classification

The classification error rates were sufficiently close to completely correct classification that the accuracy of the manually attributed reference class labelling becomes significant. **Doctor**(1983) has reported that on routine in-service inspections, an error rate of 5% for missed defects and wrong categorisation decisions is typical, and even amongst the best inspection teams, the error rate was still about 2%. This was true even when they knew that their work was being supervised and checked and thus were strongly motivated to perform well.

The technician who did the original division of the samples into classes of defect, classified the sample set using a UV plot of the entire set of scans from each defect to decide upon which type it was. It would therefore be unlikely for one single scan to be wrongly labelled unless the whole parent blemish was similarly misclassified.

It is however possible that the manually attributed class membership of some blemishes was wrong since many of the misclassified single scans were from blemishes of a uniformly low amplitude and consequently these were as likely to confuse a human operator as a machine, even when the human operator had the advantage of looking at a plot of the entire defect. In connection with this it should be noted that although defects tend to be a manifestation of only one phenomenon over a group of scans, it is possible for an axial crack to develop from intergranular cracking. However this requires years of exposure to working temperatures and pressures, and since all of the defects here came about during manufacture and had not had any long term exposure to elevated temperatures or pressures, or any great stress or vibration, such a transmutation at the ends of any defect is unlikely. On the other hand it was impossible to determine that the sample set was entirely free from such transmutative specimens without sectioning all of the defects containing misclassified scans.

In summary then, individual scans and complete blemishes of low amplitude seem to have been responsible for the observed non-zero misclassification rate, rather than any innate deficiency in the featureset or classifiers invoked.

### 6.1.2 Extendability of Methods

The segmentation method developed in chapter 4 could be applied to the signal from any other differential probe provided that the signal is first collected in a buffer long enough to contain the entire defect in the direction of the scan. The method is based upon the swings in the signal at the extremities delineating the edges of the defect. In order that the buffer be

long enough, the maximum number of samples which ever occur between the defect end points must be determined. Also, the method is limited in that defects may not overlap; in fact it presumes that the minimum distance between defects must be greater than the maximum distance between the defect extremities in a scan. In the single frequency study, this was not a serious limitation because the tubing had never seen active service and thus could not have accumulated volumetric defects such as fretting induced thinning. Otherwise, the maximum defect width parameter in the segmentation algorithm would have had to have been increased to a length eqivalent to several rotations. The combination of elliptical thinning and a nearby crack, or a dent and a crack could not then be separated by the algorithm, but such defects do not occur in 316 tubing in the types of heat exchanger in which it is used, and in any case, an ID differential probe with as small a tip as the one used (and with a maximum depth of penetration less than the tube wall thickness, taking into account the cooling fins) could not have picked up any signal from such defects. However, it is not obvious how the segmentation scheme could be modified to overcome the above limitations were it to becomes necessary to do so because, for instance, of the occurence of overlapping defects.

The main variation between one type of non-ferromagnetic metal and another from the point of view of eddy current testing is in the conductivity. For single frequency inspection, most big inter-material differences can be cancelled out by alterations to the excitation frequency, and so in principle similar defects in different materials can be made to produce similar loci. So although the segmentation threshold levels (and the classifier weight vectors) are specific to the signals from finned 316 stainless steel tubing,

the methods could probably be easily applied to other alloys provided that they are non-ferromagnetic and that they are being inspected from within the tube, in search of similar classes of ID defects: not unreasonable conditions especially if the tubing is from a heat exchanger of some description.

### 6.1.3 Syntactic Classification

It is unlikely that the performance of the adhoc featureset which has been developed could be improved on the available defect set whilst continuing to treat the scans through a defect individually. Good Fisher Discriminant ratios have been achieved and almost complete linear separability of the classes is apparent in the transformed feature axis plots.

Were a syntactic approach to be used, it would probably have to be upon the entire set of scans from a defect, because within one scan, the shape of the figure contains more or less the same primitive constituent curves in the same order every time (the essential difference between loci is not so much the order of primitives but the degree of what can be broadly described as curvature within each, a variable which really best suits statistical pattern recognition). Having decided that the only course to take with syntactic pattern recognition is to consider whole defects as one single sample, the practical problem of finding enough defects to train a classifier upon become acute. Consequently, syntactic classification was impractical in this project.

There are two other points which require to be made concerning this type of classifier, neglecting its impracticality for a moment.

**(a)** A syntactic approach might improve classification rates, because of the improvement in classifier performance observed when voting was used over complete blemishes because of the tendency of misclassification to take place at defect extremities, but since in practice majority voting would probably be used anyway, that is not what is at issue here.

**(b)** Syntactic methods tend to be more computationally intensive and difficult to program than statistical methods, and few reports of their successful application to situations requiring real-time respeonses have been published.

Thus a decision was made not to divert effort into syntactic classification in view of the fact that statistical pattern recognition appeared to be sufficiently powerful for the task in hand. Probably the most that would be gained if more sophisticated syntactic methods were employed would be improved individual scan misclassification rates for tail end scans.

Although syntactic methods may appear here to be being equated with image processing, this is not so. Treating a sequence of scans syntactically like "the words in a sentence" is not suitable for this application for the reasons given above, distinct from image processing (as briefly experimented with in Chapter 4) which was not pursued because of a lack of suitable hardware i.e. a system fast enough, small and portable enough and cheap enough to be incorporated in a site portable system.

### 6.2 Multifrequency Work

Chapter 5 described in detail the stages in a multifrequency defect classification system. It is appropriate in this chapter to explain the reasons behind the choice of methods involved in the system of chapter 5, and to stress that the system described in that chapter was not the final system to be devised, but that it was the most advanced one which was fully tested on the tubing set. In fact, the final system is that of fig. 3.3, but that one has not yet been fully freed from bugs. The transform-based feature extractors in Chapter 3 and the feature selection work was all done latterly on the PDP11/23, after most of the work for chapter 5 had been done but problems with the FFT and FD methods (that became apparent only during the testing stages) were not resolved before this thesis was written.

The original program suite has been almost completely rewritten in order to use the signals from the Zetec multifrequency equipment. Ad-hoc geometrical and spectral impedance plane featuresets from individual scans have been devised in order to express every detail of the loci that might possibly have been an essential difference between one type and the others, and the resulting sets of feature vectors have been tested for classification power with nearest neighbour and linear classifiers.

Analysis of the class membership of the defects which were misclassified would have been greatly aided if the 2-space plotting program of fig. 5.35 had been capable of flashing all of the points belonging to any one tube, in a manner analogous to the way in which an illuminated

public street map highlights all of the theatres in a town upon the press of a button. Due to the way in which they work, neither the Teklib graphics library on the PDP11's nor the Ginograf library on the GEC4180 was capable of this, but the hardware/software development system suggested for future use in the next section is.

### 6.2.1 Feature Set Development

After the rather pessimistic results were obtained in Chapter 5, a program of work was undertaken with various non-adhoc featuresets, and with a systematic featureset evaluation and selection scheme, which is described in Chapter 3.

The essential characteristic of the featureset sought was that it be insensitive to any difference in size or bulk angle between loci (which are primarily dictated by defect volume and depth respectively). It was desirable that it be sensitive only to genuine differences in shape. By satisfying this criterion it seemed reasonable to expect that if the scans still clustered according to defect depth, then there must be a genuine difference in shape between a defect of one depth and another both belonging to the same class (this was not expected to happen because it seemed from inspection of figs. 5.12-16 that there was a much stronger uniformity of shape across classes than across loci of similar depths but of different types).

It appeared from the literature that Fourier Desriptors would be ideal because other authors claimed to have found them to be insensitive to

nonconstant arc length sample spacing, while their use of that same arc length-based, rather than time-based, strategy in the transform supposedly allowed compression of the shape information into a smaller number of coefficients than the FFT could. The attraction of insensitivity to nonconstant sample spacing along the locus was that this would allow a scan based feature extractor to be used on a banded defect without reordering the sample points into an axially aligned scan. For instance, consider the point describing a locus which moves out to the extremity of the locus and stays there for a number of scanning revolutions and then collapses back to the origin. This happens with banded, axially symmetric defects such as uniform thinning and IGA which are consequently not well-suited to time domain based spectral featuresets such as those based upon the FFT. The lack of any increments in the arc length would prevent these additional points from altering the coefficients coming out of the FD-based transform and thus producing a result any different from that which would have been obtained had the cursor stayed out at the extrema for only one sample point.

The only one transform-based feature extractor to have worked so far (the FFT) was not run on the full multifrequency defect scan set because the feature vector produced was too large for the size of the scan set, and problems have been experienced in compressing the shape information into a few measurements, as explained in section 3.3.

### 6.2.2 Feature Selection Development

While assessing feature subset selection as approach to overcoming

the sample set size problem, revealed a deficiency became apparent in the range of practical criterion functions which are useful in practice in pattern recognition. One which overcame the problem was devised and has been tested upon synthetic data. A real featureset has not yet become available for it to be tested upon because the featureset data files from chapter 4 and 5 could not be transferred to the PDP11/23, and there was no point in testing it upon feature extractors studied in Chapter 3 until they were known to be able to accurately describe shape in sufficiently few coefficients. Also the feature selector could not have been sensibly run upon any of the featuresets because for instance for a typical, 32 dimensional feature vector, it was impossible to accomodate the workspace necessary on the PDP11 using Dec Fortran (greatest when using virtual memory, and even then maximum real array size is 32768 ).

### 6.2.3 Further Work

A priority for further development on should be the development of a curve parameterisation scheme for the loci from absolute probes which can describe the curve with few enough parameters for statistical validity of any classification results given the sample set available. It will then be possible to determine if samples from the same defect type but of different depths can be grouped together. If this is successful, it will be possible to use **Roberts'**(1985) empirical phase/depth curves once the class of an unknown defect has been decided to determine its depth from its phase as measured at $R_{max}$. Such a means of parameterising a locus might be by first mixing out any Pilger noise, and then fitting a function of $\phi$

to the values of r in the resulting trajectory which should be the same for both the inbound and outbound traces. Alternately, a Joukowski aerofoil (which has already been used to imitate the signal from an absolute probe) could be fitted to the point set, and here it might not be necessary to remove Pilger noise by mixing beforehand. In both cases, in view of the problems encountered in Chapter 3, prescaling of all loci would be necessary, to achieve uniform perimeter length, followed by resampling in order to get the same number of points in all loci, with an equal arc length sample spacing.

## 6.3 Suggestions for Further Work

### 6.3.1 Size of the Requirement for More Defects

One of the prime requirements which will be necessary prior to further experiments with the stages beyond the feature extractor is more defects, becasue the upper limit upon the feature vector dimensionality at the classifier is at present being dictated by the size of the sample set rather than by the number of features that it is desired to include.

It is possible to predict the minimum number of samples that will be required for further work, from the dimensionality of the feature vector. The scans of the Inconel 600 tubing contain 55 points. This is a resolution equivalent to that of the signal coming out of the eddy current equipment. For circumferential scans through any nonbanded defect, a segment of length 16 samples has been found to be adequate. For the

special case of "broad" defects like IGA, a segment which captures only the step up on to the defective area contains enough shape information to indicate the defect type. Sixteen samples is also a preferred length for a segment to be used as the input to FFT-type transform-based feature extractors. In the worst case, where it is not possible to truncate the 32 dimensional vector which results from using all of the 16 complex valued ordinates directly as a feature vector, a minimum of 66 samples per class would be required to assure the determination of a valid linear hyperplane when training a linear classifier. The number would not be substantially different for a k nearest neighbour classifier.

For the legitimate use of single scans as raw input patterns, the class that would require the largest number of whole defects would be that of circumferential cracks and notches, with eleven whole defects being needed, because they typically encompass only 6 scans each when a circumferential scanning pattern is employed.

If the segmentation scheme were to perform reorientation of the scanning direction sympathetic to individual defect orientations as per method (vii) in Section 5.6.3then individual holes or pits would become the lowest yielder of scans for each complete defect. In that case, a minimum of 6 holes would be required in the defect set because a hole typically spans 11 scans.

Were the same feature set to be used to estimate depth (or even if the classifier were to switch to an alternative set of features of the same rank, but with features more attuned to the description of depth than type), then the above class size minima would apply to each of the four

commonly accepted depth regimes individually, within each defect type and so for instance, 24 holes or circumferential cracks would be needed in the training set.

However, this unlikely to be necessary in practice because once defect classification by type has been performed, through wall depth can be calculated from a set of curves which have been fitted to point plots of defect depth vs. phase ($\phi$ at $R_{max}$) for each type of defect (**Roberts**,1985). The importance of knowing the type of defect before applying one particular depth phase curve lies in the fact that the parameters of the (depth/phase) curves vary with the volume of the defect (which is dependent upon the defect type).

### 6.3.2 Production of More Defect Loci

As pictured in chapter 4, attempts have been made at producing real cracks. All of the defects used in this thesis have been artificially produced to some degree. The defects used in almost all other applications of pattern recogniton to NDT suffer from the same criticism, because of the time required to induce natural defects and then measure them (see for example **Lemaignan**,1983 or **Foster**,1984). Due to the cost and difficulty of manufacturing defects, a defect set sufficiently big for pattern recognition purposes is not likely to become available unless access to a working steam generator is obtained. Fortunately, there are other ways to overcome this problem.

One way is to use finite element (FE) synthesis as **Satish and Lord**(1983)

have done. They have computationally simulated the signal from a defect by inching a probe past a defect whilst calculating the magnetic field and eddy current intensity at every point in a mesh laid over the cross section of the defect, and summing the contributions from all of the nodes in the mesh to predict the coil impedance at each point along its path. Three dimensional FE simulations have become available, though they require to be run on supercomputers. Such simulations will be necessary for the simulation of surface probes interacting with defects which are not symmetrical about the main axis of the probe.

Another approach might be to use existing real defects from sources other than PWR steam generator tubing. There are lots of boilers and heat exchangers containing sets of real defects, such as the heat exchangers found in oil rigs currently lying in disuse, where the defects are in materials and tubing diameters different to those of immediate interest. More defects might be added to the sample set by recording the signals from defects in Inconel 600 and from defects from these other materials providing that they are similarly non-ferromagnetic, in order to find a means of transforming the loci from defects in one material into those of defects in the other. It is unlikely that this method will be of any benefit where IGA is concerned (since this is a corrosion phenomena peculiar to the corrosion resistant Nickel alloys now being used in PWR's), but it may helpful for cracking and pitting-type defects.

There is a need to develop the defect set from one which contains only defects in straight tubing to include defects coincident with other artefacts which affect the EC signal. The curved section of an inverted-U

tube, and tube support plates and tube sheets, especially at the point where the tubes emerge from these sheets, have all caused problems in the past. Detection is difficult at these points because any small (usually circumferential crack-type) defects will be masked for instance by the edge of the ferromagnetic tube sheet, and the variation in probe/tube wall coupling due to the tube roll. Also, in practice detection can be further complicated by external tube wall copper deposits and thick deposits of electrically conductive sludge that may be lying around the outside of tubes on the tube sheet.

In an attempt to test inspection under more real (and difficult) conditions, a length of heavy steel tubing which fits tightly around the short lengths of Inconel has been used to simulate a tube sheet by the sponsors, and the existing set of defects could now be relogged with the entry point into the sleeving placed progressively at various positions with respect to each defect in each tube. Automatic mixes could then be developed which remove these signals, leaving only the mixed defect signals to be recognised.

However, although the effect of a tube passing through a support plate can be imitated by a thick tightly-fitting collar, in order to gather samples that accurately reproduce the signals from the exit of a tube from the top of the tube sheet, the tube will have to be expanded as shown in fig. 1.4 to form a tube roll region and some means of simulating the usual layer of conductive sludge will have to be provided too.

### 6.3.3 Software Development Problems

Conversion of all programs to Pascal from Fortran, and away from DEC hardware has become desirable, for several reasons. Before any other software development is attempted, a move should be made away from Fortran to a language such as Pascal in order to take advantage of such facilities as run time array bound checking. Also, it is essential to upgrade the support for graphical output (and input) if the program suite is to become user friendly enough to be useable in an interactive manner by anyone other than the programmer of the pattern recognition system himself. Furthermore, whatever hardware/software environment is chosen for further development, the hardware should sufficiently standardised that the software can distributed to inspection teams by floppy disk, so that it can be used immediately without any need for it to be modified.

Turbo Pascal on an IBM PC (or compatible clone) would seem particularly suitable, because it satifies all of the criteria above and is also fast and cheap. Also, debugging can be accomplished with ease, and recompilation (typically 30 seconds for programs of the size necessary), and consequently program debugging and development can be speeded up. On the PDP11/23, a modification to the feature extractor (a program of typical size) took about five minutes to recompile and link.

The reason that the DEC PDP11 family of processor boards and peripheral I/O cards has been so popular for real-time applications in the past (and continued to be so right up to the initiation of this project) was largely the availability of worldwide product support. A

change to more modern hardware would be worth the effort given the availability of support for the hardware advocated.

Moving to TurboPascal would involve dropping the DEC operating system RT11, and the DEC PDP11/23 hardware because TurboPascal is not compatible with either. It is unfortunate that this change of systems will require another rewriting of the program suite, this time into Pascal from Fortran, (not just as before to run under a slightly different operating system). However, in the long run (ie over several months of program development) the overhead will be repaid due to the current frequency of particularly awkward bugs which are difficult to fix under RT11 primarily due to the lack of memory because the application has outgrown the hardware, rather than because of the lack of any particular debugging tool.

There are further advantages. Converting the system in to a ROM-based form incorporated in a single board computer within an eddy current instrument (which might be the aim of further work on this topic) will be more easily achieved with the IBM PC's open architecture and consequently freely available hardware documentation than for the equivalent series of PDP11-derivative single boards computers, and is certainly possible without another change of hardware. Any advantage in speed that the PDP11 might have over the PC's 8086 could be easily outweighed with the inclusion of an 8087 fast maths processor.

Also, the 8086 has a superior memory addressing capability. Although TurboPascal does impose some maximum memory size constraints (a maximum of 64kbytes each for the code space, stack/heap space and

data space respectively), these are a considerable improvement compared with the limits for the PDP11/45 running RSX11m (64kbytes), and for the 11/23 with RT11 (44kbytes). Furthermore these limits are not as rigid as the PDP11's, because a PC can address up to 640 kbytes of RAM, and under version 3.2 of MS-DOS and above, most of the spare 440 kbytes can be used as a "RAM disk", a means of using memory to mimic an extra floppy disk drive for disk access intensive programs, whilst not incurring the usual disk access time penalties. This could be used at run time to hold a very large file of direct access records to make practical the manipulation of large matrices for feature selection, for example, which could not normally be accomodated in the 64 kbyte data space. There also exist procedures in TurboPascal to allow the heap to expand up to 512k, and an IBM PC AT or cloned equivalent can address 2.5 Megabytes of RAM, thus implying that a large part of a hard disk could be mapped into memory for really big matrix manipulation tasks.

### 6.4 Summary

The recommendations within this chapter concerning the immediate future of the project may be reduced to two matters which must be attacked before anything else. Firstly, there is still a problem concerning the availability of defects to fuel this study. This should be resolved before further work commences, as it constrains the type of approaches that can be taken to the classification problem in general.

The second suggestion is one regarding the attractiveness of a change of hardware despite the desirability of continuing with the PDP11/23 system

where all of the program suite is currently based. It is suggested that in the interests of avoiding environmental constraints upon the ease of development of programs, that all further program development should be on an IBM PC AT or equivalent, in order that the programs may be written in Turbo Pascal.

If both of these recommendations are followed, and the general strategy of Chapter 3 bears fruit (i.e. a feature extractor and selector combination which can be allowed to train unsupervised upon any set of loci), then a standalone sytem for use by operators with little interpretational skill could be produced within the next three years.

### 6.5 References

1. **Singh, G.P. and Udpa, S.**,"The Role of Digital Signal Processing in NDT",pp177-189,NDT Intl 19(3) 177-189 (1986).

2. **Houserman, H.,Lamb, L. and Kitson, W.**,"Digital Data Analysis of Steam Generator Tubing",6th ASME Conf. on NDE in the Nuclear Industry,465-480, Zurich, Dec. 1983.

3. **Doctor, S.R. and Heasler,P.G.** "A Pipe Inspection Round Robin Test",6th ASME Conf. on NDE in the Nuclear Industry,563-568, Zurich, Dec. 1983.

4. **Lemaignan, C.**,"Creation of an Internal Crack in Tubes for NDE Calibration",6th ASME Conf. on NDE in the Nuclear Industry,825- 827, Zurich, Dec. 1983.

5. **Foster D.G.**,"PWR Steam Generator Tubes - Rotating Probe Eddy Current Testing", BPL Report No. (62)/85/23, May 1985.

6. **Roberts, P.T.E.**,"Assessment & Development of Rotating Probe Eddy Current Techniques for PWR Steam Generator Tubing",BPL Report No. E/85/073 September 1985.

# 7 Appendix: Logging System Development

## 7.1 The Mk 1 (Single Frequency) System

The original system used for logging the 316 tubing defects is shown in fig.A.1. It was developed by Dr. Macleod and Dugald Campbell, the senior Departmental Technician. It logged the tube to floppy, which was then transferrred from the single board computer system to the PDP11 minicomputer using a pair of file transfer programmes, one for each machine, both called **$micro**. Files were transferred in 4-bit wide nybbles rather than bytes, so there existed a program on the PDP11, **Nybtran**, which reformatted the data back into bytes. The last program in the data logging process was **Chopscan**, which was introduced to chop inordinately long scans (which occured occasionally because of a missed sync. pulse) into 400 sample long sections and also to equalise th length of all of the scans to 400 points. The circograph equipment generated a once per revolution pulse. The logging software was made to detect this, and to insert a byte signifiying an end of scan (80H) into the stream of byte-valued samples. The inclusion of this byte made it possible to plot the helical scan as a set of separate scans such that their beginnings are aligned, in the form that looks as if the tube had been cut open along its length and lain out flat.

The logging program was written in Microsoft Fortran, on the Z-80 SBC100-based system, with all of the real time sampling and sync. pulse detection performed in an assembly language subroutine. The program stored the samples within the 56 kbytes available to the user and after it was

stopped by user intervention from the keyboard, saved these to disk. Storing blocks of samples to disk without interrupting the regularity of the sampling process would have required a dual processor system with one processor to control disk operations during sampling (which is not possible on the S100 bus). The system did not require any buffering of the signal in order to slow the signal down (unlike the multifrequency system described next), so the system could have been connected directly to the circograph had it not been for transportation problems. However, storing the samples in RAM did limit the maximum tube file size to 56 kbyte (the equivalent of 70 single frequency scans with 400 points/scan in each).

### 7.2 The Multifrequency Systems

Although it was desirable to have a data logging system that could tackle a long length of tubing, such as one whole inverted U-tube, this required that any emptying of the sample buffer to disk should have no effect upon the sampling process, which must continue without sample loss or delay. Consequently a system such as a pdp11 which stops work to service a hard disk (because it is serviced with the highest priority), even if for only a few milliseconds, is not tolerable.

A system was developed which used the SBC100 to control the sampling process, and the PDP11 to store the data asynchronously. The databytes were to have been have been streamed to the PDP11/45 via a handshaking parallel interface, which would have blocked the data and dumped it periodically to its hard disk, thus achieving an arrangement which would be able to process an unlimited length of tube, the

overcoming the SBC100's memory size limitation without losing samples due to disk accesses (because the SBC100 system had enough RAM to buffer any samples taken until the PDP11 became free again).

The approach failed to work because the program on the PDP11/45 which was to have controlled communication with the SBC100 and the blocking of the data onto its own hard disk could not be debugged, because it could not be loaded into memory once the DEC macro11 debugger ODT was already there. The fault did not occur in the program sections separately, only while they were together, so debugging the program in smaller chunks was not a solution, so unfortunately this approach had to be abandoned.

### 7.3 The Mk2 System

Detection of the sync. pulse was not reliable, and this together with some other problems such as an inability to use the Zetec MIZ12 with the circograph probe forced a number of changes in the tube logging hardware, the most major of which were

(i) the sampling based pulse detection system was converted to a monostable based design so that the sync pulse could not be missed by the pdp11, because once the monostable was set, it had to be cleared, and this only happened when it was read by the pdp11.

(ii) A stepper motor driven tube inspection rig was built by Babcock, in order to overcome the shuddering and probe lift-off experienced when using the Circograph.

The Mk II (multifrequency) logging system centred around a 256 byte circular

buffer, which received the 8 ADC samples from the 8 channels after every sampling interval, and between these occasions was continually emptied via an RS232 link to a 600kbyte cartridge datalogger running at 9600baud. The sync pulse monostable was checked at this point, and reset when found to be set. Since the "most recent data-in" pointer to the buffer could not be allowed to lap the "next data to be sent out" pointer, the maximum sampling (ie buffer filling) rate was dictated by the serial port output rate (approximately 960 bytes/sec at 9600 baud). This proved to be too slow for direct connection to the Zetec, and so required the signal to be slowed down by a factor of 16, which was done by recording it at 15 inches/sec (i.p.s.) and then replaying it at 15/16 i.p.s.. The standard tape speed for logging used by Babcock was 7.5ips, so the entire deifect set had to be re-recorded at 15 i.p.s.in order to get the necessary slow down factor of 16:1, because the slowest speed was only 15/16 ips.

This slowing down gave a scanning rate of one scan every 6.5 seconds, or at a 60Hz sampling rate, 390 samples/scan, a rate which had been deliberately chosen to approximate that used with the single frequency tubing datafile set. Ultimately, this sampling rate proved to be higher than necessary by a factor of about 7, especially since the probe used was an absolute type, giving a more slowly varying signal. However the probe type change did require the use of the Racal 14 channel FM tape recorder, to record signals with a d.c. component, a facility which the the more conventional tape recorder used in the first system did not have.

The high sampling rate detrimentally affected filesizes and transfer sometimes by as much as a factor of 7 compared with the old single frequency

programs. For instance the individual file transfer times from the SBC100 to the bigger departmental computers were:

(i) 80 minutes for a typical 0.8 mbyte file to the 4180, followed by 3 hours for the 8 channel tubeplotting program to produce a plot file, and 2 hours to actually produce this plot.

(ii) 20 minutes for the same file to the PDP11/45; but where high enough quality plots such as those in figures 5.22 to 26 were not possible.

A non-system file format was used to store the tube data files on the PDP11 (to save diskspace), and was incorporated in the **$micro** parallel file transfer program, and so unfortunately the files could not be read using normal Fortran I/O statements. The files had to be read using a custom set of assembly language routines. Corresponding routines to write files did not exist, and because the documentation concerning the $micro file format and these custom subroutines was very poor, no attempt was made to devise any. Consequently no attempt has been made to rewrite any of the tube files after segmentation.

### 7.4 The Mk3 System

Upon transferring the work to Babcock's PDP11/23, the logging program was rewritten again. This was partly because it was more convenient to be able to log the tubes on the premises, and partly because transferring of the archived data files on the 4180 to the 11/23 would have been difficult given the 11/23's lack of any form of magtape reader or modem with which to communicate with the University, and also because of the sheer size of the datafiles, 10 of which would have entirely filled the hard disk.

The mkIII program/system which was thus created used a slightly modified version of the synchronising pulse detection monostable, but reverted to the previous "memory filling/ single dump to disk" strategy, because the computer did not have any spare RS232 ports to drive a cartridge data logger.

The logging program had to be completely rewritten in Macro11 in order to use RT11 system subroutines to drive the 11/23's ADC and parallel interface cards. The documentation for this version is held by the computer services group at **Babcock's Research Centre.**

mk 2

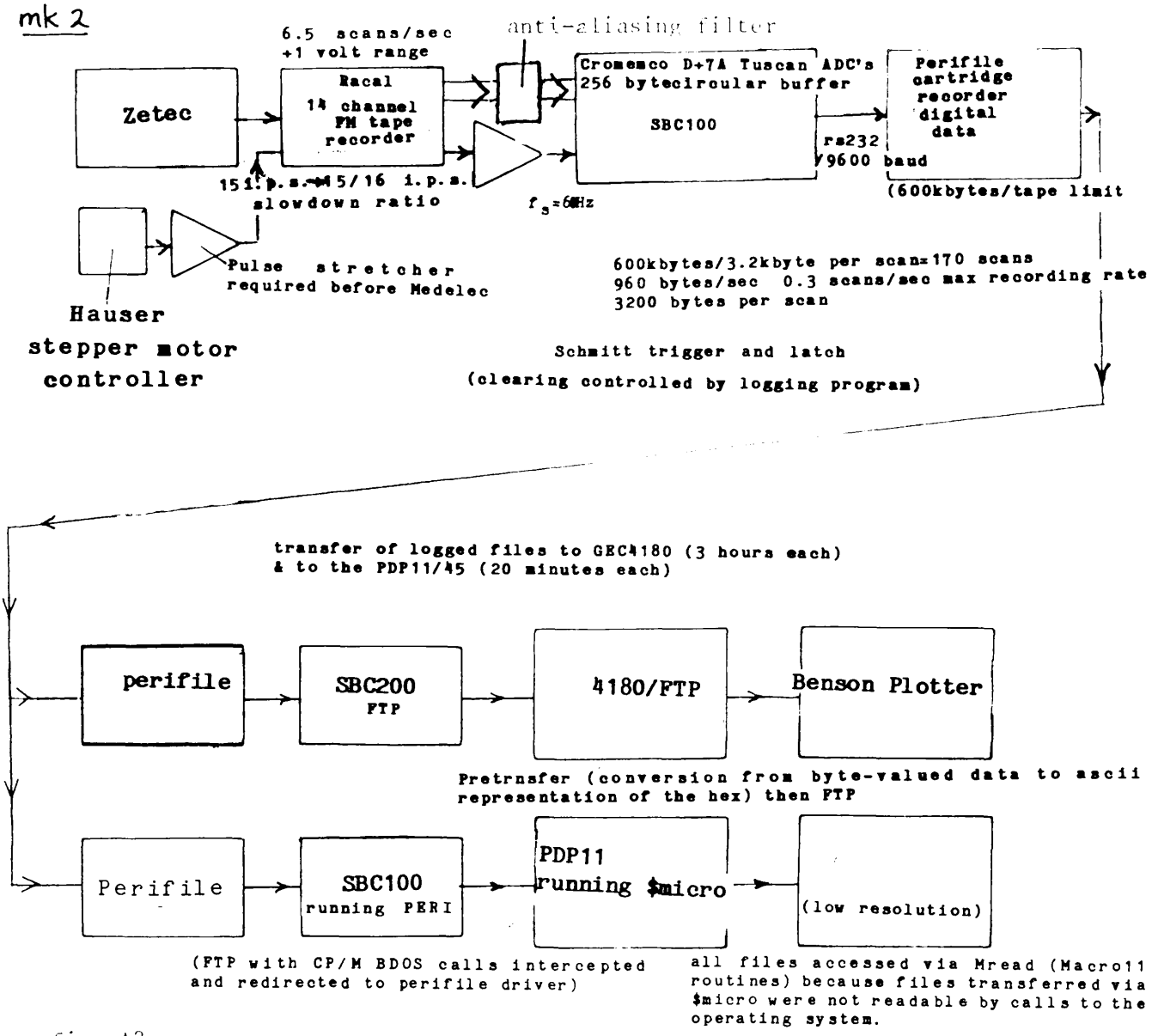



fig. A2

Mk 3

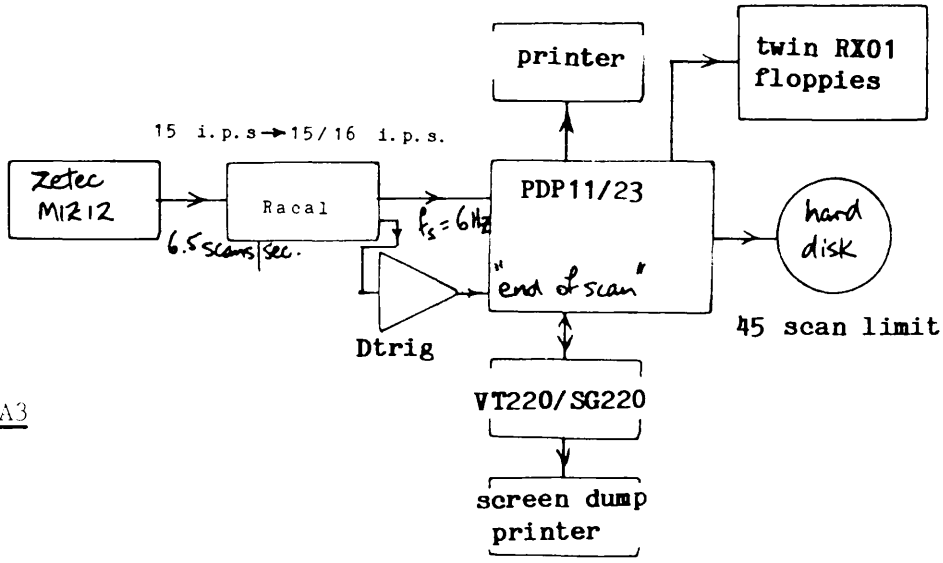



fig. A3

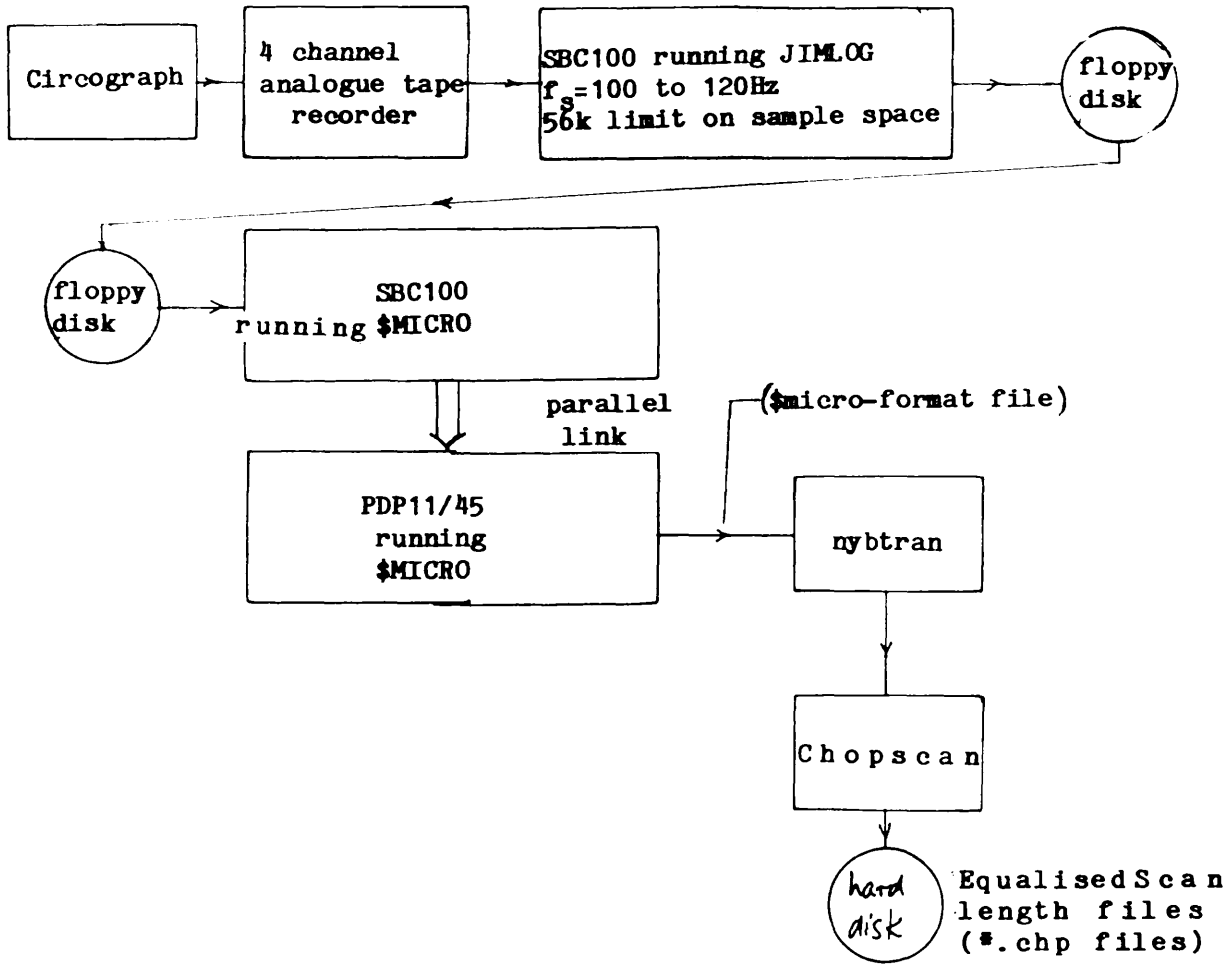
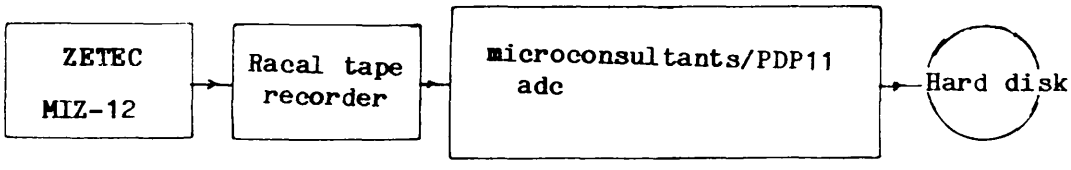



fig A.1 The original logging system which was used to record all of the single frequency scans of the tubes used in Chapter 4.




**MFEC system mk1**