

IOWA STATE UNIVERSITY

Digital Repository

Retrospective Theses and Dissertations

Iowa State University Capstones, Theses and
Dissertations

1996

Invariance transformations for processing NDE signals

Shreekanth Ammanji Mandayam

Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/rtd>



Part of the [Artificial Intelligence and Robotics Commons](#), and the [Electrical and Electronics Commons](#)

Recommended Citation

Mandayam, Shreekanth Ammanji, "Invariance transformations for processing NDE signals" (1996). *Retrospective Theses and Dissertations*. 11120.

<https://lib.dr.iastate.edu/rtd/11120>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company

300 North Zeeb Road, Ann Arbor MI 48106-1346 USA

313/761-4700 800/521-0600

Invariance transformations for processing NDE signals

by

Shreekanth Ammanji Mandayam

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Department: Electrical and Computer Engineering

Major: Electrical Engineering (Communications and Signal Processing)

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Members of the Committee:

Signature was redacted for privacy.

Iowa State University
Ames, Iowa
1996

UMI Number: 9620979

**UMI Microform 9620979
Copyright 1996, by UMI Company. All rights reserved.**

**This microform edition is protected against unauthorized
copying under Title 17, United States Code.**

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

To Shama and Parijatha

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	xi
CHAPTER 1. INTRODUCTION	1
1.1 Objective and Scope of the Dissertation	3
1.2 Research Contribution	4
CHAPTER 2. PARAMETERS THAT AFFECT MFL SIGNALS	7
2.1 Magnetostatic Flux Leakage NDE	7
2.2 MFL Inspection of Gas Transmission Pipelines	8
2.2.1 Steps in MFL signal interpretation	11
2.2.2 State-of-the-art in pipeline inspection procedures	12
2.2.3 Governing equations	13
2.3 Effect of Magnetization Characteristic	13
2.3.1 Grade of pipe material	14
2.3.2 Stress	16
2.4 Effect of Probe Velocity	18
2.5 Effect of the Origin of Sampling	19
CHAPTER 3. INVARIANT PATTERN RECOGNITION: A CON-	
TEXT FOR INQUIRY	21
3.1 Invariance by Feature Selection	23

3.2	The Concept of Invariance in Mathematics	24
3.2.1	Applications of invariant theory	27
3.2.2	Invariant groups	29
3.3	Invariance by Signal Compensation	33
3.3.1	Velocity effects compensation	33
3.4	Invariant Pattern Recognition Techniques for Characterizing MFL Signals	34
CHAPTER 4. UNIVERSAL APPROXIMATION: A TOOL FOR INVARIANCE TRANSFORMATIONS		36
4.1	An Introduction to Approximation Theory	37
4.1.1	Basic terminology	38
4.1.2	Basic theorems	38
4.2	The Interpolation Method for Invariance Transformation	39
4.3	The Permeability Invariance Scheme	41
4.4	Phenomenological Methods	44
4.4.1	The Zatsepin and Shcherbinin model	44
4.5	Non-phenomenological Methods	46
4.5.1	Simple functions	48
4.5.2	Feedforward networks for universal approximation	49
4.6	Radial Basis Functions	52
4.7	Fuzzy Inference Systems	53
4.8	Multiresolution Networks: Wavelet Basis Functions	58
4.9	3-D Permeability Invariance Scheme	67
4.10	Issues in Permeability Invariance	68

4.10.1	The necessity of using two signals (features)	68
4.10.2	Robustness of the transformation	72
CHAPTER 5. INVARIANCE BY SIGNAL COMPENSATION		74
5.1	Velocity Effects Compensation	74
5.1.1	Analytical methods	74
5.1.2	Deterministic methods	76
5.1.3	Stochastic methods	78
5.1.4	Fixed scheme	82
5.1.5	Adaptive scheme	87
5.2	Pre-processing MFL Signals for Velocity Effects Compensation	90
5.2.1	Analytical method: Compensation for variation in sampling origin	90
5.2.2	Stochastic method: Signal registration	92
5.3	3-D Velocity Invariance Scheme	98
CHAPTER 6. RESULTS AND DISCUSSION		100
6.1	Permeability Invariance	100
6.1.1	Radial basis function networks	100
6.1.2	Fuzzy inference systems	104
6.1.3	Wavelet basis function networks	104
6.1.4	3-D permeability invariance	109
6.2	Velocity Invariance	109
6.2.1	Deterministic filter	112
6.2.2	Stochastic filters: Linear and quadratic	112
6.2.3	Adaptive filters	114

6.2.4	Preprocessing MFL signals prior to velocity effects compensation	116
6.2.5	3-D velocity effects compensation	124
CHAPTER 7. CONCLUSIONS	126
7.1	Summary	126
7.2	Conclusions	128
7.2.1	Permeability invariance	128
7.2.2	Velocity invariance	131
7.3	Future Directions for Research	134
7.3.1	Improvements in the invariance algorithms	134
7.3.2	MFL signal characterization	135
BIBLIOGRAPHY	137
APPENDIX A. UNIVERSAL APPROXIMATION BY MULTI-LAYER FEEDFORWARD NETWORKS	144
APPENDIX B. BEST APPROXIMATION PROPERTY OF MULTILAYER FEEDFORWARD NETWORKS	150
APPENDIX C. FUNCTIONAL EQUIVALENCE BETWEEN RBF NETWORKS AND FUZZY INFERENCE SYSTEMS	153
APPENDIX D. GAUSSIAN WAVELETS AND THE ADMISSIBILITY CONDITION	155

LIST OF FIGURES

Figure 1.1:	The inverse problem in nondestructive evaluation	2
Figure 1.2:	An overview of the research topics in this dissertation	5
Figure 2.1:	Illustrating the origin of leakage fields [3]	9
Figure 2.2:	Schematic of MFL based pipeline inspection system	10
Figure 2.3:	Axial and radial leakage field profiles for a rectangular slot .	10
Figure 2.4:	Variation of magnetization characteristic with grade of pipe material	15
Figure 2.5:	Variation of B_z with magnetization characteristic and depth for a 2 inch defect	16
Figure 2.6:	Variation of magnetization characteristic with stress for an X-42 pipe	17
Figure 2.7:	Variation of axial flux densities with stress for an X-42 pipe .	17
Figure 2.8:	Influence of pig velocity on the MFL signal	19
Figure 2.9:	Variation of axial flux densities due to different sampling origins	20
Figure 3.1:	Illustration of coordinate transformation	26
Figure 4.1:	Variation of P_z and P_r with magnetization characteristic . .	42
Figure 4.2:	Flow-chart of the permeability invariance scheme	43

Figure 4.3:	Strip dipole configuration	45
Figure 4.4:	Permeability invariant feature derived using Zatsepin model .	47
Figure 4.5:	Permeability “invariant” feature derived using simple functions from the <i>TableCurve</i> © software. All defects are 2 inches long. - - - : 20% deep; — : 50% deep; - - - - : 80% deep . . .	50
Figure 4.6:	A typical multilayer feedforward neural network	51
Figure 4.7:	RBF network for the permeability invariance scheme	53
Figure 4.8:	Fuzzy inference system	54
Figure 4.9:	Triangular and Gaussian membership functions	56
Figure 4.10:	The input-output map of the permeability invariance scheme	59
Figure 4.11:	The Mexican hat wavelet and its dilated and translated versions	62
Figure 4.12:	Wavelet basis function network	65
Figure 4.13:	Gaussian scaling and wavelet functions	67
Figure 4.14:	Variation of the axial component of the flux density for different B-H curves for a 3-D defect	69
Figure 4.15:	Variation of the radial component of the flux density for different B-H curves for a 3-D defect	70
Figure 4.16:	Variation of the circumferential component of the flux density for different B-H curves for a 3-D defect	71
Figure 5.1:	A simple velocity compensation filter model	75
Figure 5.2:	Variation of parameters a (—) and b (- -) with velocity	76
Figure 5.3:	Velocity effects compensation using analytical filter. — restored, - - desired	77
Figure 5.4:	Adaptive filter	87

Figure 5.5:	Interpolator-decimator cascade to compensate for variations in the origin of sampling	93
Figure 5.6:	Misalignment in 1-D axial scans	94
Figure 5.7:	3-D MFL signal obtained at 2 mph	95
Figure 5.8:	Perfectly aligned 3-D MFL signal, obtained at 7 mph	96
Figure 5.9:	Misaligned 3-D signal, obtained at 7 mph	97
Figure 5.10:	3-D velocity effects compensation scheme	99
Figure 6.1:	Permeability invariant signals derived using RBF network . .	101
Figure 6.2:	Permeability invariance using the axial and circumferential components of the MFL signal. - - - 20%, - - - 50%, — 80% .	103
Figure 6.3:	Permeability invariance scheme using only the radial component	105
Figure 6.4:	Data vectors and cluster centers determined using gradient descent in FIS: \times = data points, \circ = cluster centers	106
Figure 6.5:	Interpolation amongst magnetization characteristic for FIS .	106
Figure 6.6:	Input-output mapping of the WaveNet for training and testing data	108
Figure 6.7:	Depth interpolation results using WaveNet. - - - 20%, - - - 50%, — 80%	109
Figure 6.8:	Typical input MFL signals for the 3-D permeability invariance scheme	110
Figure 6.9:	3-D permeability invariant MFL signals	111
Figure 6.10:	Velocity effects compensation using deterministic filter	113
Figure 6.11:	Velocity effects compensation using fixed linear filter. — re- stored, - - desired, - - distorted	115

Figure 6.12: Second order terms of the quadratic Volterra filter	116
Figure 6.13: Eigenvalues of the auto-correlation matrix	117
Figure 6.14: Mean squared error with iteration.— training data, - - testing data	117
Figure 6.15: Velocity effects compensation using gradient descent algorithm: training data — restored, - - desired	118
Figure 6.16: Velocity effects compensation using gradient descent algorithm: testing data — restored, - - desired	119
Figure 6.17: Results of the spatial shift invariance scheme	120
Figure 6.18: Cross-correlation vectors for perfectly aligned and misaligned 1-D MFL scans	121
Figure 6.19: Cross-correlation matrix for perfectly aligned 2-D MFL scans	122
Figure 6.20: Cross-correlation matrix for misaligned 2-D MFL scans . .	123
Figure 6.21: Typical input signal to the 3-D velocity effects compensation scheme	124
Figure 6.22: Desired output signal of the 3-D velocity effects compensation scheme	125
Figure 6.23: Restored signal obtained from the 3-D velocity effects com- pensation scheme	125

ACKNOWLEDGEMENTS

Many thanks to Dr. Lalita Udpा for guiding me through my graduate program, every step of the way. I am fortunate in having an advisor with an unfailing enthusiasm for every new idea. Thanks for leading me into fascinating explorations in areas that I would not normally have traversed and making my graduate experience richer, rewarding and fulfilling. The many hours she spent going through this dissertation and suggesting improvements are sincerely appreciated.

It has been a pleasure to be a part of Dr. Satish Udpा's research team during the course of this project. This student has greatly benefited from Dr. Udpा's breadth of knowledge and research acumen. For all the help, encouragement and the many opportunities given to me, my deepest gratitude.

Sincere thanks to Dr. William Lord for his help and guidance at many stages of this research work. His advice on technical writing is much appreciated. It has been a privilege to work with Dr. Lord during this project.

Dr. Yushi Sun's expertise in the magnetics area has greatly contributed to the success of this research project. My heartfelt thanks for his help and advice.

Thanks to Dr. Eric Bartlett for being on my research committee and suggesting several improvements in this dissertation to make it more clear and readable. His advice is much appreciated.

Thanks to Dr. William Meeker, Jr. for being on my research committee.

The financial support of the Gas Research Institute, Chicago, IL, is gratefully acknowledged. This has been a very enjoyable project, in a large part due to the camaraderie of the original “fab four” team — thanks to Marian, Gopi and Jiatun. Those were happy times!

My deepest of thanks to my family — father Shama, mother Parijatha and brother Sreedhar, for their unswerving support and endurance during my very long graduate program. Couldn’t have done it without you!

CHAPTER 1. INTRODUCTION

Inverse problems in nondestructive evaluation (NDE) address the crucial issue of defect characterization wherein the information/features in NDE signals are characterized in terms of shape, size and location of defects in the material. Electromagnetic NDE methods span the entire electromagnetic spectrum, based on the frequency of the excitation source. Magnetostatic flux leakage (MFL) and electrostatic potential drop NDE are dc (zero frequency) methods. Eddy current and remote field NDE are quasi-static techniques. X-ray radiographic, microwave and ultrasonic NDE are wave-based high frequency methods. As one traverses the frequency spectrum, the entropy or the amount of information contained in the signal increases [1] [2]. Compared to wave or diffusion based methods, static methods such as MFL yield signals with the least amount of entropy, thereby making the inverse problem or signal characterization more difficult.

The challenge in signal characterization is the development of invariance algorithms for processing the raw signal so that it is insensitive to the variations in test parameters. Such processing ensures the possibility of developing subsequent classification schemes and in some cases renders such schemes simple and efficient. Techniques for performing the aforementioned tasks fall under the category of invariant pattern recognition algorithms. Invariance transformations are becoming a subject

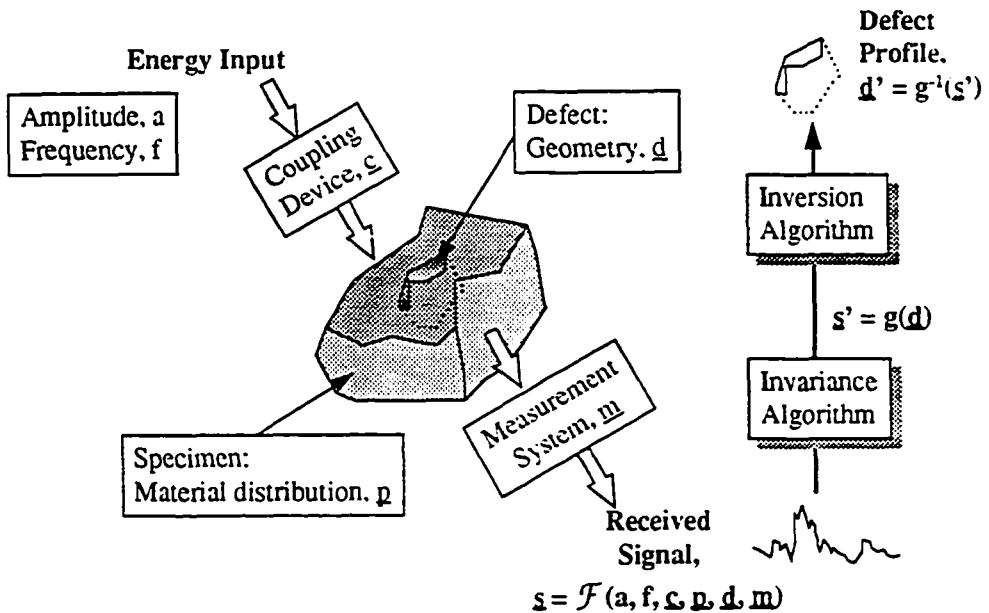


Figure 1.1: The inverse problem in nondestructive evaluation

of increasing activity in pattern recognition and forms the bulk of signal and image processing steps prior to classification. Figure 1.1 summarizes the inverse problem in electromagnetic nondestructive evaluation.

Parameters that affect a flaw signal can originate from varying experimental conditions or artifacts introduced due to the measurement system. The objective of the invariance transformation is to isolate the flaw signal from the unwanted effects introduced by the measurement systems. For example, transducer frequency invariant signals are desired in the case of ultrasonic NDE. In radar applications, images invariant to object rotation, scaling and translation are required. In magnetostatic flux leakage NDE, the application discussed in this dissertation, invariance to a variety of parameters is desired. These parameters may be related to the characteristics

of the test object such as magnetization history, residual stress etc. Or, they may describe the test condition such as probe velocity and probe orientation. Accurate signal characterization in terms of defect geometry is possible if the MFL signal is made invariant to these parameters.

1.1 Objective and Scope of the Dissertation

A systematic procedure for developing a generalized invariance scheme that addresses the problem of rendering nondestructive evaluation signals insensitive to selected test parameters, is presented in this dissertation. Various methods of achieving signal invariance in general are addressed, with specific application to invariance issues in magnetostatic flux leakage nondestructive evaluation. While designing an invariance scheme it is important to ensure that the technique compensates for undesirable parameter variations, such as permeability, velocity etc., while retaining at the same time, sensitivity to defect related parameters, such as depth, width, etc. A signal may exhibit completely different functional dependencies with respect to different parameters. Mathematically, we can say that the parameters describing the signal are all orthogonal to each other. In such a case it is relatively easy to isolate the influence of an undesirable parameter on the signal and thereby compensate for it. For example, in a case where the undesirable effects and the defect-induced changes occupy different frequency bands, simple band-pass filtering will suffice. The majority of invariant feature selection schemes employing shape or statistical descriptors operate with signals of disparate functional dependencies. However, in some cases, the changes in the signal due to variations in the experimental parameters, is similar to the corresponding changes due to varying defect parameters. The former effects are

to be compensated for, whereas the latter effects need to be preserved and analyzed. Here, the parameters describing the signal are non-orthogonal. The compensation procedures in these cases are therefore more complex.

In this dissertation, invariance techniques to compensate the MFL signal for operational variables such as permeability of the test specimen and velocity of the scanning device form the primary focus of the investigation. A wealth of signal processing techniques have been developed over the years. This study attempts to determine the optimal combination of these techniques that is best suited for making MFL signal interpretation more accurate.

1.2 Research Contribution

This dissertation presents the following original contribution in the area of invariant pattern recognition. A novel method, that recasts the invariance transformation as a problem in interpolation of multidimensional data, is developed. Subsequently, various universal approximation methods are explored to determine the optimal technique to solve the interpolation problem.

Figure 1.2 gives a global perspective of all the topics that are addressed in this dissertation.

This dissertation is organized as follows:

Chapter 1 commences with a statement of the problem and provides a preview of various aspects of this investigation. An introduction to magnetic flux leakage methods of nondestructive evaluation is given in Chapter 2. The effect of parametric variations on the MFL signal is studied and the need for invariance transformations is demonstrated.

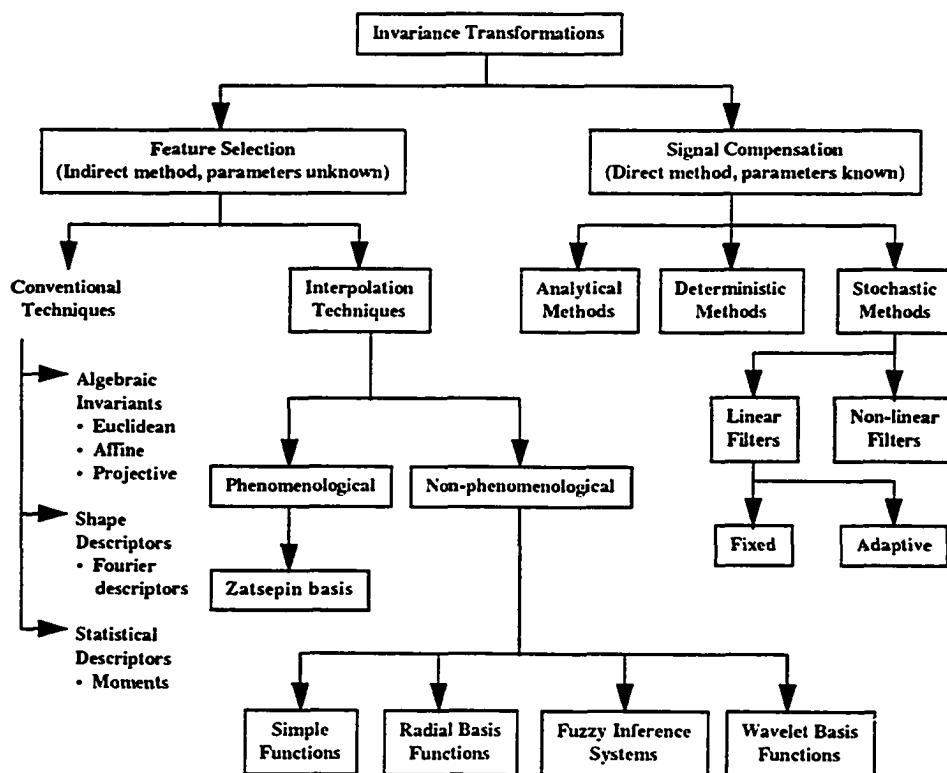


Figure 1.2: An overview of the research topics in this dissertation

Chapter 3 is comprised of a brief survey of existing invariant pattern recognition techniques. The limitations of these methods with respect to the requirements in the current application area are discussed.

Universal approximation techniques are introduced in Chapter 4. The invariance transformation is recast as an interpolation problem and the permeability invariance scheme is developed. Phenomenological methods based on analytical models of the MFL testing situation are investigated. Non-phenomenological methods based on various types of feedforward neural networks are used as tools for performing universal approximation based invariant pattern recognition.

Signal compensation mechanisms are studied in Chapter 5. Deterministic, stochastic and adaptive algorithms for velocity effects compensation are developed. Pre-processing schemes required prior to signal compensation are described.

Chapter 6 presents the results of implementing the invariance transformation algorithms that have been developed. The dissertation concludes with Chapter 7 wherein the research accomplishments are summarized and future research directions are discussed.

CHAPTER 2. PARAMETERS THAT AFFECT MFL SIGNALS

2.1 Magnetostatic Flux Leakage NDE

Magnetic flux leakage (MFL) methods of nondestructive evaluation (NDE) are widely used in industry for inspecting the integrity of ferromagnetic objects. The leakage flux method is conceptually a two step procedure — in the first step, the test object is magnetized using direct current or a permanent magnet. In the presence of anomalies, the flux is redistributed resulting in “leakage” magnetic flux in the surrounding region. In the second step, the leakage flux is measured and detected using flux sensitive devices such as Hall-effect probes or pickup coils.

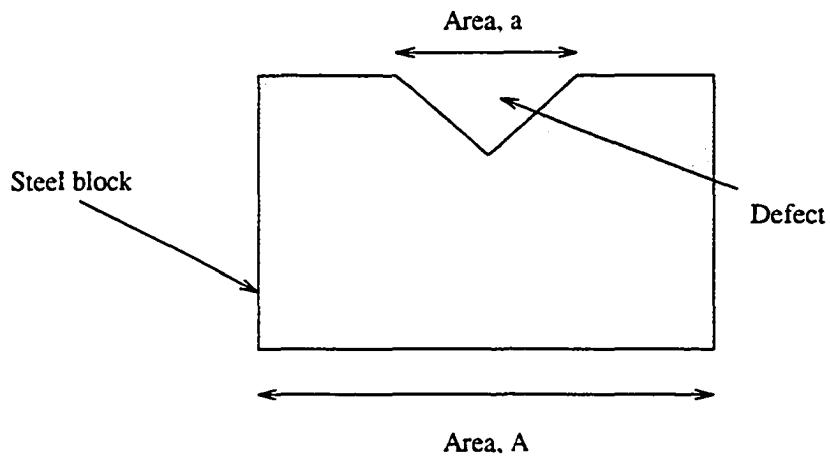
A brief description of the underlying physics responsible for the origin of the leakage field follows [3]. Consider a ferromagnetic block of cross-sectional area A , with a surface-breaking defect of area a , as shown in Figure 2.1(a). The cross-sectional area of the undamaged portion of the block is $(A - a)$. When the block is placed in a uniform magnetizing field of intensity, \mathbf{H} , the induced flux density is B_1 , indicated by the point Q in Figure 2.1 (b). The corresponding permeability is indicated by point P . The total flux through the undamaged part of the block is given by $\phi = BA$. If the same flux is constrained to pass through the reduced area around the flaw, the flux density in this region increases to $B_2 = B_1 A / (A - a)$. Consequently, in the flaw region, the quiescent point on the flux density curve shifts

to Q' and the permeability decreases to P' . Due to the increase in flux density and a reduction in the permeability, some of the flux leaks into the surrounding region. This leakage flux can then be detected by Hall-effect sensors.

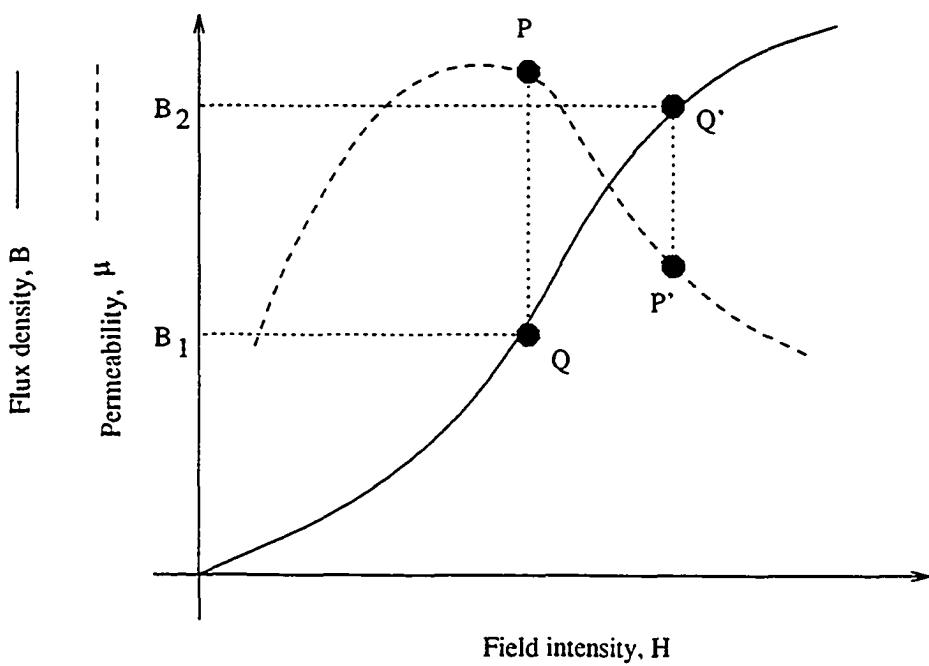
2.2 MFL Inspection of Gas Transmission Pipelines

An example of the application of MFL methods in industry is the inspection of gas pipelines that are buried underground. Assuring the integrity of the 280,000 mile network of gas transmission pipelines is crucial for the safe and economical supply of one of the nation's cheapest forms of energy — natural gas. The pipes are inspected with a magnetizer-sensor assembly known as a "pig", shown schematically in Fig. 2.2.. The pig is comprised of a strong permanent magnet to magnetize the pipe-wall, a circumferential array of Hall-effect sensors to measure the leakage flux and a microprocessor based data acquisition system to digitize and store the data. The inspection procedure involves moving the pig through a section of the pipeline network, either by pulling or by utilizing the pressure of natural gas. At the end of the pigging operation the pig is retrieved and the data acquired is analyzed.

The Hall-effect sensors measure the axial and radial components of the magnetic flux density. The signals measured in an axial scan of a rectangular slot in the pipe-wall are as shown in Figure 2.3. The shape of the entire signal contains considerable information about the defect profile. For example, the peak-peak amplitude of the MFL signal is a good measure of the defect depth and the peak-peak separation distance characterizes the defect's axial length [4].



(a) Ferromagnetic block with defect



(b) Magnetization characteristics

Figure 2.1: Illustrating the origin of leakage fields [3]

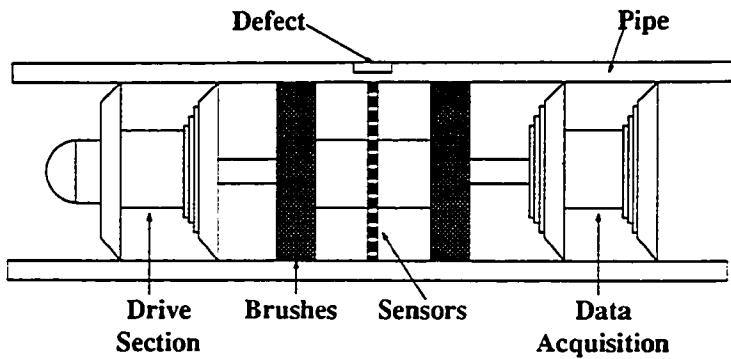


Figure 2.2: Schematic of MFL based pipeline inspection system

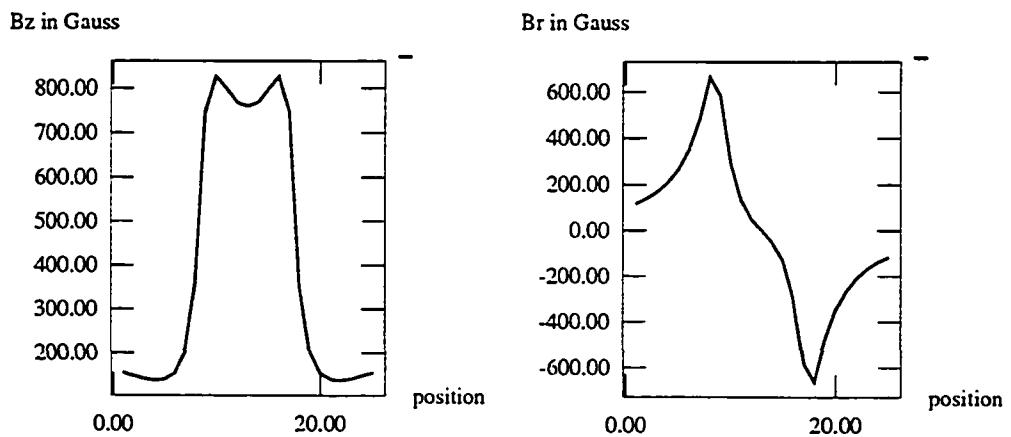


Figure 2.3: Axial and radial leakage field profiles for a rectangular slot

2.2.1 Steps in MFL signal interpretation

Current MFL inspection tools detect and store the axial component of the magnetic leakage flux density. Every pigging operation results in the collection of a voluminous amount of MFL data. A systematic procedure to analyze this data consists of the following three steps:

1. *Signal Identification:* Leakage flux arises at every region of the pipe where there exists a local variation in the magnetic behavior. Such local variations include changes in the geometry and/or permeability of the region surrounding the tool. As a result, MFL indications are obtained not only due to defects present in the pipe-wall, but also because of welds, joints, transitions, T-sections, valves, etc. The first step in the analysis process involves separating the benign indications from the potentially dangerous ones. This process is referred to as signal identification.
2. *Signal Compensation:* After an MFL signal has been identified as a “defect signature”, a direct characterization of the signal is still fraught with difficulty. The signal received by the sensors is dependent on a variety of test parameters some of which cannot be measured or accurately controlled. These parameters include the pipe material, pipe-wall thickness, axial and hoop stresses in the pipe, velocity of the pig, spatial shifts in the sensor orientation over repeated runs, remanent magnetization in the pipe etc. All of these parameters alter signal levels and distort the shape of the signal. Therefore, defect characterization must be preceded by appropriate signal compensation techniques, that render the raw MFL signal invariant to the operational parameters.

3. Signal Characterization: The processed MFL signal can be analyzed in one of two ways. A simple characterization scheme will predict the mean length, depth and width of a defect, conveying little information about its actual shape. In most situations, defects caused by corrosion have complicated shapes. A complete defect characterization algorithm, is able to predict the exact profile of the defect, indicating the variation in defect depth along the axial and circumferential directions, on the pipe.

2.2.2 State-of-the-art in pipeline inspection procedures

Current pipeline inspection operations culminate in the manual inspection of the data collected, before remedial measures, if required, can be taken. The MFL data is typically displayed on a CRT using appropriate software. Trained operators scroll through the data, flagging the MFL indications. Features that are flagged as defects are analyzed for depth, length and width predictions using simple algorithms based on calibration curves. In summary, the state-of-the-art in MFL signal interpretation does not include steps towards compensating the signal; identification is carried out manually; and characterization involves the use of simple calibration approaches. The research work presented in this dissertation describes various methods for MFL signal compensation. A systematic study of the effects of different operational variables on the MFL signal, commencing with an examination of the governing equation for the MFL process is presented next.

2.2.3 Governing equations

The differential equation governing the electromagnetic processes in the moving pig is given by [5]

$$\nabla \times \frac{1}{\mu}(\nabla \times \mathbf{A}) = \mathbf{J} - \sigma \frac{\partial \mathbf{A}}{\partial t} + \sigma \mathbf{v} \times (\nabla \times \mathbf{A}) \quad (2.1)$$

where \mathbf{A} is the magnetic vector potential, μ is the permeability, σ is the conductivity, \mathbf{v} is the velocity of the probe and \mathbf{J} is the source current density or its equivalent. The magnetic flux density \mathbf{B} is given by

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (2.2)$$

Choosing a cylindrical (r, θ, z) coordinate system with the z -axis coincident with the axis of the pipe, the leakage flux density, \mathbf{B} can be decomposed into the following components — the axial flux density, B_z , the radial flux density, B_r , and the circumferential flux density, B_θ . Essentially, it is seen that the leakage field is dependent on the material characteristics (permeability and conductivity) of the pipe, the velocity of the scan, the geometry of the test situation, as described by the differential operators, and boundary conditions. The following sections describe the effect of some of these parameters and the cause for their variation in a real world test situation.

2.3 Effect of Magnetization Characteristic

As seen in Equation (2.1), the MFL signal is dependent on the permeability of the pipe-wall material. In general, this is not a constant and varies as a function of the applied field intensity as described by the magnetization characteristic or B-H curve. The pipe material E-H curve itself is not the same for all MFL pigging

operations. It is dependent primarily on the grade of pipe material. Also residual stresses in the pipe-wall alter the magnetization characteristic which in turn affect the MFL signal.

2.3.1 Grade of pipe material

A majority of the gas pipeline system in the United States has been installed in the last forty years; but some segments that were constructed prior to World War II are still in operation. The pipe materials used are by no means consistent and range in grade from ones designated Grade-B to X-70. The B-H curves of these materials show different magnetization histories and saturation levels. A study was conducted to investigate the sensitivity of a defect signal to a systematic variation of magnetization properties. Two B-H curves, one saturating at 13 KG and the other at 26 KG, were chosen as those representing the upper and lower bounds of the variation in the pipe materials. A family of curves that lie between these extreme values was generated using functional interpolation given by

$$B(H) = (1 - t)B_1(H) + tB_2(H) \quad (2.3)$$

where $0 \leq t \leq 1$ is a parameter that indexes a particular B-H characteristic, $B(H)$, lying between $B_1(H)$ and $B_2(H)$. Note that when $t = 0, B = B_1$ and when $t = 1, B = B_2$. The family of curves $B(H)$, shown in Figure 2.4 can be considered to encompass all variations in B-H characteristics of the pipe materials.

The MFL signals for each magnetization characteristic were obtained using a computational model. The pipeline geometry, including the permanent magnet was modeled using a non-linear axisymmetric finite element method [6]. MFL signals were

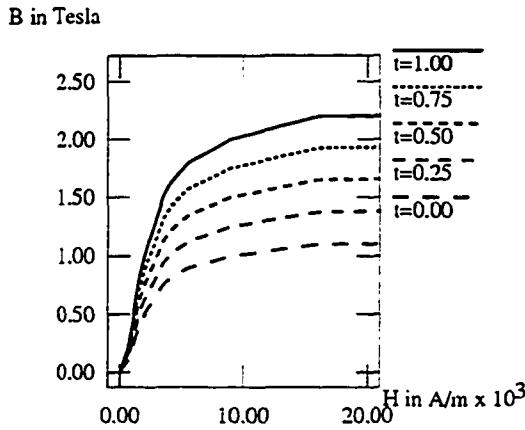


Figure 2.4: Variation of magnetization characteristic with grade of pipe material

obtained for a variety of rectangular defect profiles of varying depth and constant width under a controlled variation of material characteristics as shown in Figure 2.4. It is important to note that the pipe-wall magnetization cannot be controlled experimentally with any reasonable degree of accuracy and the data for this study has to be generated using a numerical model. Figure 2.5 shows the variation of the axial component of the leakage flux density (B_z) due to a rectangular defect, for four different magnetization characteristics. The defect chosen is 2 inches long and 50% (of wall thickness) deep. Figure 2.5 also shows the change in B_z with defect depths, for a single B-H characteristic. The spread in the range of the flux density values due to variation in the magnetization for a single defect indicates that a raw MFL signal can be easily misinterpreted in terms of the defect depth. Hence, the MFL signal must be made insensitive to magnetization history, before being subjected to defect characterization.

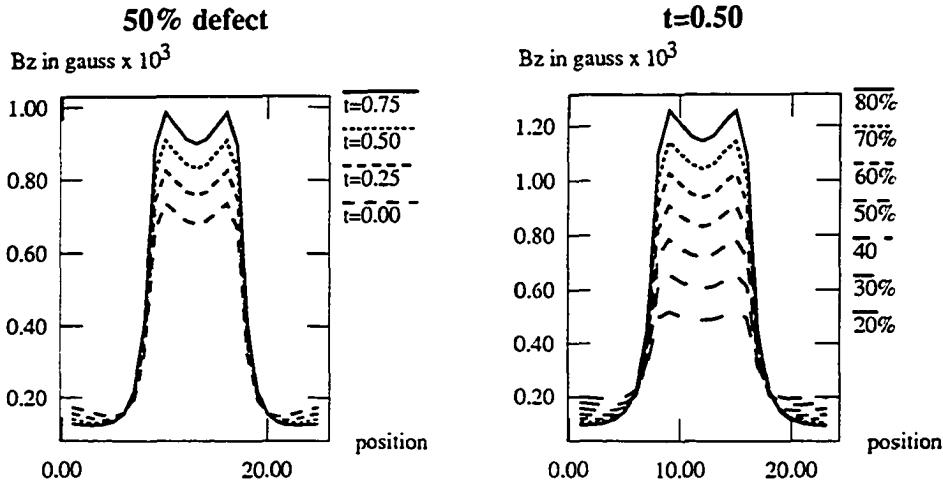


Figure 2.5: Variation of B_z with magnetization characteristic and depth for a 2 inch defect

2.3.2 Stress

A gas pipeline in operation is subject to a variety of compressive and tensile, axial and hoop (circumferential) stresses. Stress alters the magnetic behavior of an object. The effect of an applied stress causes some of the domain walls in the material to become free. These free domains move in such a manner that the magnetization curve approaches the ideal magnetization curve (anhysteretic) [7]. While compressive stresses vary the B-H curve significantly, tensile stresses are seen to have very little effect [8]. The change in magnetization characteristic for an X-42 grade pipe for different axial and hoop stresses has been studied at Southwest Research Institute and one such plot is shown in Figure 2.6 [9]. The corresponding axial leakage field intensities, obtained from a finite element simulation is shown in Figure 2.7. While the change in flux density values is not as pronounced as in the case of variation in pipe material grade, some variation nevertheless exists.

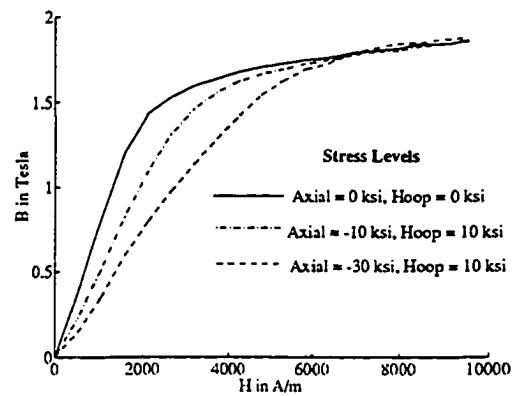


Figure 2.6: Variation of magnetization characteristic with stress for an X-42 pipe

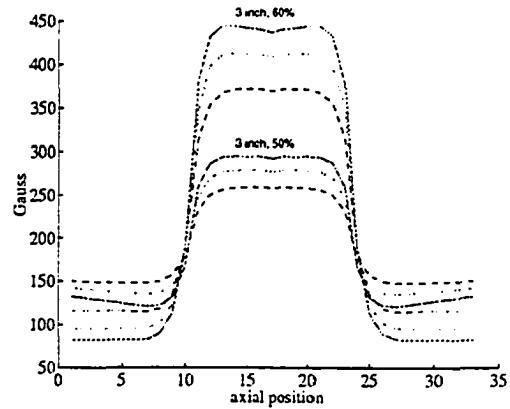


Figure 2.7: Variation of axial flux densities with stress for an X-42 pipe

2.4 Effect of Probe Velocity

The leakage field that is seen by the Hall elements is altered due to the relative motion between the fixed defect and the moving magnetizer-sensor assembly. The last two terms on the right hand side of Equation (2.1) become prominent at sensor velocities greater than 2 mph. At high sensor speeds, the current inducing effects of tool velocity come into play. The tool velocity influences the MFL signal in the following two ways:

1. Motionally induced currents in the pipe-wall and under the pole shoes alter the signal dc levels. These currents occur as a result of the $\mathbf{v} \times (\nabla \times \mathbf{A})$ term in Equation (2.1). When the flux density, $\mathbf{B} = \nabla \times \mathbf{A}$, and the tool velocity \mathbf{v} are parallel to each other, $\mathbf{v} \times \mathbf{B} = 0$. As the pig moves in the pipe, at the regions under the pole shoes, there exist components of \mathbf{B} that are normal to \mathbf{v} which are responsible for current generation.
2. Motionally induced currents around the defect distort the signal shape. These currents arise due to the $\sigma \frac{\partial \mathbf{A}}{\partial t}$ term in Equation (2.1). In the presence of a defect, the field, or equivalently, the potential, \mathbf{A} , in the vicinity of the defect changes with $\frac{\partial}{\partial t}$ and hence the velocity.

Figure 2.8 shows experimental axial component MFL signals from a 4 inch long, 50% deep rectangular defect obtained at pig velocities of 2 mph and 8 mph. It can be seen that the MFL signal obtained at a velocity of 8 mph has a lower dc level and is asymmetrically distorted.

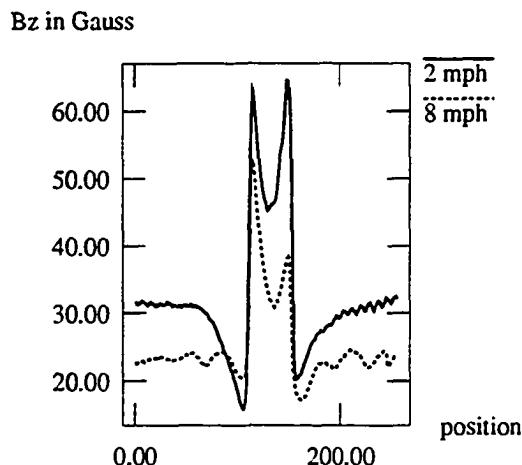


Figure 2.8: Influence of pig velocity on the MFL signal

2.5 Effect of the Origin of Sampling

As the pig moves along the axis of the pipe, the MFL signal is recorded at discrete spatial steps. Successive runs of the pig cannot ensure that identical points along the pipe-walls are being scanned. The MFL signal, for identical defects can therefore look different due to variation of the origin of the sampling. This becomes a problem not as much for defect characterization, as it is for the construction of a signal restoration filter to compensate for the effects of tool velocity. As is described in Chapter 5 the coefficients of the restoration filter are derived from a deterministic or stochastic analysis of the MFL signals at two different velocities. In order to facilitate this filter design process, it is essential that the signals at the two velocities be obtained from identical locations on the pipe. This, however, may not be always possible. Figure 2.9 shows two MFL signals obtained at different axial positions for

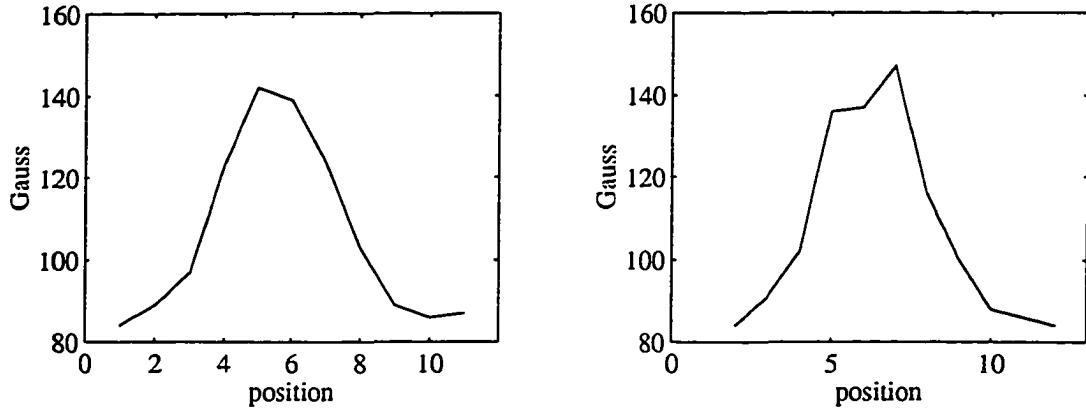


Figure 2.9: Variation of axial flux densities due to different sampling origins

the same defect. It is apparent that a preprocessing scheme to compensate for this experimental artifact is required before subsequent signal processing.

Consequently, there exists a need for eliminating the effects of the operational parameters described in the previous sections, before the MFL signal can be interpreted. An in-depth discussion of the issues involved in invariance transformation is presented in the next chapter.

CHAPTER 3. INVARIANT PATTERN RECOGNITION: A CONTEXT FOR INQUIRY

The previous chapter demonstrates the need for invariance transformations as a prior step to interpreting real world signals that are subject to varying test conditions. The field of invariants is not new and a variety of techniques have been developed over the years, rooted in rigorous mathematical theory. Engineering disciplines routinely employ signal processing systems to achieve invariance to parametric variations before signal classification using pattern recognition techniques. Artificial neural networks are an important class of pattern recognition methods.

Neural networks are densely interconnected systems of simple computational elements connected by links with variable weights [10]. The network acquires knowledge through a learning process, which is facilitated by a learning algorithm. The learning algorithm modifies the weights of the network in an orderly manner so as to achieve a desired design objective [11]. The supervised learning paradigm involves the modification of these weights by applying a set of training samples, each of which consists of a unique input signal and the corresponding desired response. When the network weights are adapted so as to provide the desired response, the network is said to be trained. Subsequently, the network is tested with samples that are not part of the original training set. The power of the neural network lies in its ability to generalize

and interpolate from the training data samples so as to enable output predictions for the test data.

Mathematically, the neural network is capable of a non-linear, multidimensional functional mapping from an input space onto an output space; the parameters of this mapping are the network weights. A pattern recognition problem requires the assignment of an input signal into one of several predetermined categories or classes. This necessitates the estimation of decision boundaries in the input signal space to partition the data samples into a known number of classes. The learning paradigm described earlier enables the generation of complex, arbitrary decision boundaries. As a result, neural networks are eminently suitable for pattern recognition applications.

Two types of networks that are widely used in these applications are the multilayer perceptron (MLP) networks and the radial basis function (RBF) networks. MLP networks are primarily used in pattern classification applications. RBF networks are used principally in functional approximation problems. The architecture and capabilities of both these networks are discussed in Chapter 4.

An invariant pattern recognition system, consisting of an artificial neural network architecture, can fall into one of the following three categories [12]. In the first method, the structure of the neural network itself can be designed such that it can recognize that disparate signals do indeed belong to the same class. The second method is essentially a brute-force method, accomplished by presenting the representatives of a large class of transformations during the training phase so that the neural network learns that these signals are equivalent. The third approach relies on identifying features of the original signals that are invariant to the desired parametric variations. These features are then input to the neural network for characteriza-

tion. The primary requirements of the feature selection schemes employed for defect characterization are:

1. Features of signals originating from the same defect, but varying test conditions, must be identical.
2. Features of signals from different defects must be different.

In other words, the feature selected must be invariant only to the operational variable and sensitivity to defect geometry must be preserved. The feature selection technique is by far the most popular and has found extensive use in diverse applications, ranging from scene analysis [13] to eddy current NDE [14]. In most situations, a feature selection scheme can operate in complete ignorance of the parameter for which invariant signals are required.

The aforementioned methods are signal interpretation techniques. An altogether different approach is to build a signal compensation scheme, that attempts to “correct” the signal before it is classified, thereby undoing the unwanted effects of parametric variation. Such a scheme must ensure that only the undesirable parameter variations are compensated for, leaving valuable information unaltered. Typically, compensation schemes require the value of the undesirable parameter as an input.

In this dissertation, feature selection and signal compensation are the two mechanisms that have been used to achieve parameter invariant signals.

3.1 Invariance by Feature Selection

Feature selection techniques are more widely used because, typically, they do not require that the parameter responsible for the signal variation be specified. A survey

of the literature shows that typical features extracted are shape recognition features such as Fourier descriptors and statistical features such as moments, kurtosis etc. In this dissertation, an entirely new feature extraction method based on interpolation techniques has been developed for processing MFL signals. It is shown that this method is sufficiently general so that it can be used in many other applications. But first, a brief survey of the classical invariant feature selection techniques is presented. The following review starts with a discussion of the algebraic theory of invariants, which forms the foundation for conventional invariant pattern recognition algorithms.

3.2 The Concept of Invariance in Mathematics

One of the most fundamental concepts in mathematics, along with number, set, function, transformation, etc. is the idea of an *invariant* [15]. Mathematically, an invariant represents a quantity, that is related in a definite manner to an object, which remains unaltered following a coordinate transformation of the object [16].

To explore this concept further, a few mathematical preliminaries need to be addressed. A mathematical object is defined by constructing a one-to-one correspondence between the points of the object, P , and real numbers (x_1, x_2, \dots, x_n) . These real numbers, (x_1, x_2, \dots, x_n) , are called the coordinates of the point in an n -dimensional coordinate system. This correspondence can be symbolically depicted as

$$P \leftrightarrow (x_1, x_2, \dots, x_n) \quad (3.1)$$

Suppose, $P \leftrightarrow (y_1, y_2, \dots, y_n)$ is another one-to-one correspondence over the same geometric set of points, P . From the correspondences in the two coordinate systems, $P \leftrightarrow (x_1, x_2, \dots, x_n)$ and $P \leftrightarrow (y_1, y_2, \dots, y_n)$, one can write a single one-

to-one correspondence $(x_1, x_2, \dots, x_n) \leftrightarrow (y_1, y_2, \dots, y_n)$ between the points of the arithmetical space of n -dimensions. The correspondence $(x_1, x_2, \dots, x_n) \leftrightarrow (y_1, y_2, \dots, y_n)$ is called a coordinate transformation and can also be expressed as

$$\begin{aligned} x_\alpha &= \phi_\alpha(y_1, y_2, \dots, y_n) \\ \text{or} \\ y_\alpha &= \psi_\alpha(x_1, x_2, \dots, x_n) \end{aligned} \quad (3.2)$$

where $\alpha = 1, \dots, n$ and ϕ_α and ψ_α are functions of the coordinates (y_1, y_2, \dots, y_n) and (x_1, x_2, \dots, x_n) respectively. Owing to the one-to-one correspondence between the points in the two coordinate systems, the functions ϕ_α and ψ_α are invertible. Such coordinate transformations are also known as permutations or displacements of the arithmetic space.

As an illustration of coordinate transformation, consider a 2-D Cartesian coordinate system OXY (see Figure 3.1 (a)). If (x, y) are the coordinates of a point in this system, a coordinate transformation of the form

$$\begin{aligned} x &= x' \cos \omega - y' \sin \omega + h \\ y &= x' \sin \omega + y' \cos \omega + h \end{aligned} \quad (3.3)$$

defines a new coordinate system, $O'X'Y'$ where (x', y') are the coordinates of the point in the new system (see Figure 3.1 (b)).

The equation interrelating the coordinates of the points in the mathematical object serves to describe some property of the object in relation to the external coordinate axes. Any arbitrary function of the coordinates of the points will yield a quantity that defines its disposition relative to the specified coordinate axes.

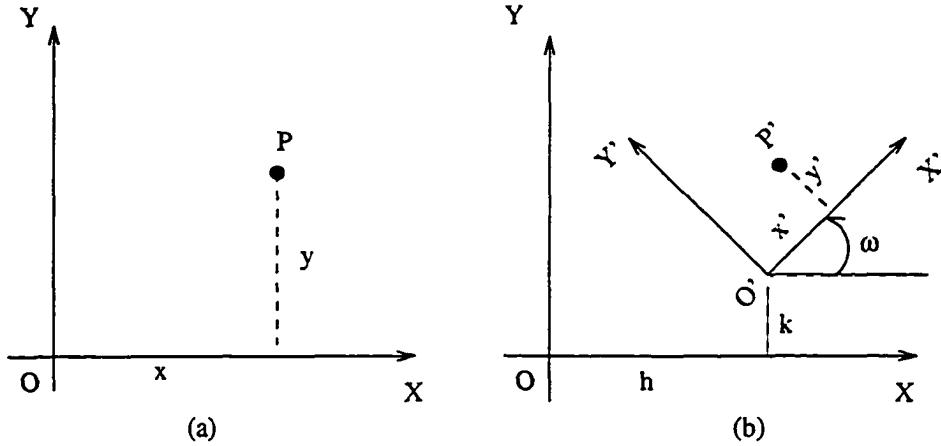


Figure 3.1: Illustration of coordinate transformation

Those functions of the coordinates of the points which are equal to quantities inherent in the object and express its properties, without any relation to the coordinate axes or its disposition thereof, are termed invariants. The defining equation for an invariant function, I , for the coordinate transformation depicted in Equation (3.2) is

$$I(x_1, x_2, \dots, x_n) = |M|^q I(y_1, y_2, \dots, y_n) \quad (3.4)$$

where $|M|$ is the determinant of the transformation and q is known as the weight. If $q = 0$, then, I is called an *absolute* invariant; otherwise, the invariant is termed *relative*.

Revisiting the example for coordinate transformation in 2-D Cartesian coordinates, the function of the coordinates of two points given by

$$d = (x_1 - x_2)^2 + (y_1 - y_2)^2 \quad (3.5)$$

is an absolute invariant with respect to the coordinate transformation defined in

Equation (3.3). To illustrate —

$$(x_2 - x_1) = (x'_2 - x'_1) \cos \omega - (y'_2 - y'_1) \sin \omega$$

$$(y_2 - y_1) = (x'_2 - x'_1) \sin \omega + (y'_2 - y'_1) \cos \omega$$

Therefore,

$$(x_2 - x_1)^2 - (y_2 - y_1)^2 = (x'_2 - x'_1)^2 - (y'_2 - y'_1)^2$$

Geometrically, the invariant, d , represents the square of the Euclidean distance between the two points (x_1, y_1) and (x_2, y_2) . Intuitively, it is apparent that this measure must remain unchanged regardless of the choice of the coordinate system.

The study of invariants deals with the search for such functions of coordinates and owes its origin to a number of problems in number theory, algebra and geometry. The development of invariant theory began with the discovery of an invariant measure belonging to a binary function (a function in two variables) in 1841 by G. Boole. The theory was further developed by A. Cayley (1846) in his work *Memoir on Hyperdeterminants* [17]. Continuing this work, J. Sylvester introduced fundamental definitions and formalized the theory. This laid the foundations for the calculus of modern algebra [18].

3.2.1 Applications of invariant theory

3.2.1.1 Mathematics Invariants can be constructed intuitively by observing the geometrical or physical characteristics of a mathematical object. The Euclidean distance measure, described earlier, is a case in point. This measure is constructed by utilizing a hypothesis underlying all observations — the concept of a

rigid body. Rigidity can be expressed as the necessary condition that there should exist a relation between every pair of coordinates belonging to the body, which is independent of the position of the body. Alternatively, one can temporarily abandon all geometric interpretations and begin by constructing the invariant function and later determine its geometric relevance. Such invariant constructions lay the foundations of different types of algebras and geometries. From a purely mathematical (geometrical) point of view, invariant theory provides a framework to test whether any equation or magnitude can represent a true geometrical entity. The requisite condition is that the quantity must remain invariant under the relevant group of transformations.¹

3.2.1.2 Theoretical physics Invariant theory has played (and continues to play) a vital part in the development of the fundamental theories of physics. The special theory of relativity postulates that the laws of nature have a form that is invariant under transformations of the Lorentz group, which is comprised of the variables forming the space and time coordinates. Out of this condition arises the requirement that the speed of light must remain constant regardless of the choice of the coordinate system. Theories describing the nature of matter, structure of the universe and its origin rely on the study of “symmetries”, which are in some sense invariants to the physical laws governing them.

¹A group of coordinate transformations is a set of coordinate transformations with the following properties:

1. The resultant of every two transformations belongs to the set.
2. The inverse of every transformation exists and belongs to the set.

3.2.1.3 Engineering In engineering applications, algebraic invariant theory forms the basis for invariant pattern recognition systems. This is accomplished by first determining a one-to-one correspondence between the coordinate transformations and the variation of a physical parameter. Naturally, invariant functions under coordinate transformations become invariant features to the variation of the physical parameter. Various groups of coordinate transformations represent various types of alterations in the physical parameter. Classical invariant pattern recognition relies heavily on group theory and a brief description of some of the invariant groups, along with their engineering applications follow.

3.2.2 Invariant groups

3.2.2.1 Euclidean invariants Euclidean groups contain transformations of the form

$$\begin{aligned}x' &= x \cos \omega - \epsilon y \sin \omega + h \\y' &= x \sin \omega + \epsilon y \cos \omega + h \\ \epsilon &= \pm 1\end{aligned}$$

This group is also known as the group of motions and reflections. Invariants in the Euclidean group are called metric invariants and form the basis of Euclidean geometry. These invariants include the distance between two points, the ratio in which an interval is divided by a point, the length of a vector, angle between two straight lines, etc. Metric invariants are used in rudimentary pattern recognition applications.

3.2.2.2 Affine invariants An affine transformation maps a point $\mathbf{p} = (x, y)$ into $\mathbf{p}' = (x', y')$ by the relation [19]

$$\mathbf{p} \rightarrow \mathbf{p}' = \mathbf{A}\mathbf{p} + \mathbf{v} \quad (3.6)$$

where \mathbf{A} is a general nonsingular 2×2 matrix and $\mathbf{v} = [v_1, v_2]^T$ is any vector in \mathbb{R}^2 . Affine transformations map straight lines into straight lines, intersecting lines into intersecting lines and parallel lines into parallel lines. Mutual locations of two lines are unaffected by the transformation. Affine invariants are invariant functions of the affine transformation. For example, consider the affine transformation given by

$$\begin{aligned} x' &= px + qy + h \\ y' &= rx + sy + k \end{aligned}$$

It can be shown that given three points (x_1, y_1) , (x_2, y_2) and (x_3, y_3) , the equations

$$\begin{aligned} x_3 &= \frac{x_1 + \lambda x_2}{1 + \lambda} \\ y_3 &= \frac{y_1 + \lambda y_2}{1 + \lambda} \end{aligned}$$

are invariant with respect to the affine transformation.

Affine invariants are used in computer vision applications for recognizing that different perspective views of the same object are equivalent [19].

3.2.2.3 Projective invariants A projective transformation is a general case of the affine transformation with the mapping given by

$$\mathbf{p} \rightarrow \mathbf{p}' = \frac{1}{w_1 x + w_2 y + 1} \mathbf{A}\mathbf{p} + \mathbf{v} \quad (3.7)$$

where all quantities are defined as before. It can be seen that this formula reduces to the affine transformation for $w_1 = w_2 = 0$. Projective transformations map points onto points and preserve the sum and intersection of subspaces and the independence of points. A projective invariant of classical importance is the cross-ratio, known variously as the double or anharmonic ratio [20]. This relates an ordered set of four points lying on a line, $(x_1, y_1), (x_2, y_2), (x_3, y_3)$ and (x_4, y_4) . The cross-ratio, defined by

$$\frac{(x_3 - x_1)(x_2 - x_4)}{(x_2 - x_1)(x_3 - x_4)} = \frac{(y_3 - y_1)(y_2 - y_4)}{(y_2 - y_1)(y_3 - y_4)} \quad (3.8)$$

is invariant to a projective transformation of the line.

The projective invariant cross-ratio is extensively used in image analysis and object recognition for problems involving perspective projections, partial obscurations etc. Identification of aircraft images is one popular application [21].

3.2.2.4 Moment invariants The moments of order $(p+q)$ of a two-dimensional function $f(x, y)$ are defined by [22]

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy; \quad p, q = 0, 1, 2, \dots \quad (3.9)$$

The central moments μ_{pq} are defined as

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) d(x - \bar{x}) d(y - \bar{y}); \quad (3.10)$$

where $\bar{x} = m_{10}/m_{00}$ and $\bar{y} = m_{01}/m_{00}$. It has been proven that given algebraic invariants of the form (3.4), the central moments of order p have a similar invariant

$$I(\mu'_{p0}, \dots, \mu'_{0p}) = |J||M|^q I(\mu_{p0}, \dots, \mu_{0p}) \quad (3.11)$$

where $|J|$ is the absolute value of the Jacobian of the transformation. Therefore, algebraic invariants can be used to construct moment invariants. From the second

and third order moments, a set of seven functions, invariant to the coordinate transformations of translation, rotation and scaling have been derived.

Statistical descriptors such as moment invariants are applied primarily to images. A typical pattern recognition application is in character recognition [22]. Further development of this work has led to the introduction of complex moments [13]. Complex moments are thought to not only provide in-class invariant features but also enhance the ability to distinguish between classes, especially in the presence of noise. Yet another type of invariant moments, known as Zernicke moments are constructed by mapping an image onto a set of orthogonal functions [23]. The magnitude of the Zernicke moment is invariant to the rotation of the object. Furthermore, reconstruction of the original image is simple using the Zernicke moments.

3.2.2.5 Fourier descriptors Fourier descriptors are shape descriptors derived from the Fourier series, that parameterize the shape of the signal. The Fourier descriptors approach for parameterizing closed contours is based on the fact that any closed curve (signal) can be represented by a cumulative angular function or a complex contour that repeats every 2π radians. This periodic function, in turn, can be expressed as a Fourier series. Appropriate functions of the Fourier series coefficients remain invariant under translation, rotation or scaling of the original signal. These functions form invariants for parametric variations in translation, rotation and scaling.

Fourier descriptors have been used in handwriting recognition and character analysis applications [24]. They have also been used to characterize impedance plane trajectories in eddy current NDE systems [14]. Drift in the zero and gain settings of

the eddyscope results in translation and scaling of the impedance plane trajectory, to which Fourier descriptors of the signal remain unaffected.

3.3 Invariance by Signal Compensation

In this approach, signal compensation systems treat the effects of the test parameters as distortion that needs to be removed. Accordingly, signal compensation is achieved using restoration filters. Designing such a scheme involves computing the filter kernel or filter coefficients, which are dependent on the value of the distortion parameter.

3.3.1 Velocity effects compensation

Here, the problem of signal restoration is considered as an inverse filtering problem, wherein the clean image is distorted by passing through a degradation filter [25]. One example frequently encountered in image processing is blurring by uniform linear motion relative to the object or the imaging device. The image can be restored by convolving the degraded signal with the “inverse” kernel function. Fourier domain filters offer a widely used solution method [26]. The fundamental frequency-domain, linear, space-invariant imaging equation is

$$G(u, v) = F(u, v)H(u, v) + N(u, v) \quad (3.12)$$

where $G(u, v)$ is the Fourier transform of the blurred image, $F(u, v)$ the Fourier transform of the original image, $H(u, v)$ the transfer function of the degradation filter and $N(u, v)$ represents an additive noise process.

For example, consider an image, $f(x, y)$, that is subject to uniform linear motion

in the x -direction. Let the velocity be given by, $x_0(t) = at/T$, wherein the image has been displaced by a distance a in the time interval $t = T$. One can analytically determine a degradation filter for this process, given by

$$H(u, v) = \frac{T}{\pi ua} \sin(\pi ua) \exp(-j\pi ua) \quad (3.13)$$

where (u, v) are the corresponding Fourier domain coordinates for the space coordinates (x, y) .

3.4 Invariant Pattern Recognition Techniques for Characterizing MFL Signals

The principal motivation behind the research presented in this dissertation is the search for suitable invariant pattern recognition algorithms to enhance the accuracy of interpretation of MFL signals. Invariance techniques for the compensation of the two operational variables, namely, pipeline magnetization characteristic (permeability) and the velocity of the inspection tool, form the primary focus of this investigation. Observations of the effects of these operational variables on the MFL signal lead to the conclusion that feature selection techniques are appropriate for compensating for pipeline permeability and restoration filters are suitable for velocity effects compensation. The feasibility of employing the conventional invariance schemes presented earlier in this chapter is now discussed.

Classical invariant pattern recognition algorithms are based on the algebraic theory of invariants. The algebraic theory of invariants is primarily concerned with obtaining functions that are invariant to coordinate transformations. In order to use this invariant theory, the variations of the test signal or image are mapped onto the

variation in the arithmetic coordinates; the variations in the operational parameter is made to correspond to the coordinate transformation function. For NDE applications, it is crucial that a one-to-one correspondence exists between the operational variables and the coordinate transformation function. Otherwise, if the variation in a defect related parameter and an undesirable operational variable is represented by the same coordinate transformation, it is inevitable that an invariant function (feature) derived for this transformation will destroy any defect related information. This is precisely the case in the variation of the MFL signal with pipe-wall permeability and defect depth. Since these variations are so similar (see Figure 2.5), both of them can be represented by a single coordinate transformation. Consequently the conventional invariant feature selection techniques such as Fourier descriptors and moment invariants fail to achieve permeability invariance and simultaneously preserve defect related information.

In image processing applications, image distortion due to relative motion between the camera and the image can be removed by constructing an analytical expression for a restoration filter, once the relative velocity is known in magnitude and direction. However, in the case of MFL NDE a similar analytical expression compensating for velocity distortion of the MFL signal will not suffice. This is because, the moving magnetizer-sensor assembly physically alters the signals seen by the sensor array, according to Equation (2.1). Analytical expressions for the restoration filter are therefore more difficult to compute.

The following chapters describe the development of novel invariant pattern recognition techniques that can successfully overcome the limitations of the conventional algorithms.

CHAPTER 4. UNIVERSAL APPROXIMATION: A TOOL FOR INVARIANCE TRANSFORMATIONS

Universal approximation or interpolation techniques can be employed as tools to construct invariance transformations explicitly. This is a sophisticated method that has the ability to render the signal invariant to one parameter yet sensitive to another, even when both parameters affect the signal in an identical manner. For example, in the case of MFL NDE, the variation of the signal is identical with respect to two parameters — defect depth and magnetization characteristic (see Figure 2.5). While elimination of the latter variation, whose value is not known, is desired, the change in signal amplitude with respect to depth must be preserved.

In this approach, the invariance transformation is recast as a problem in interpolation of scattered multidimensional data. The field of computational mathematics is rich with sophisticated techniques for data interpolation that can now be employed for achieving invariance. Of all these techniques, feed-forward neural networks have triumphed as the ones possessing the widest range of application. Multiquadric surface interpolation, as in a radial basis function (RBF) network, has been proved to guarantee non-singular matrices that yield unique surfaces that interpolate distinct data points [27]. Recent developments such as fuzzy inference systems (FIS) [28] and wavelet transform based networks (WaveNets) [29] can also be viewed as interpo-

lation (universal approximation) techniques. A brief introduction to approximation theory is presented in this context.

4.1 An Introduction to Approximation Theory

Approximation theory is primarily concerned with selecting a particular member from a prescribed class of functions that is, in some measure, “close” to a certain fixed function [30]. Linear approximation theory deals with the class of polynomial functions and non-linear approximation theory deals with rational functions. Typical measures of distance (norms) between two functions $f(x)$ and $g(x)$ are:

1. *Supremum norm*, defined by $\max |f(x) - g(x)|$.
2. *Infimum norm*, defined by $\min |f(x) - g(x)|$.
3. *Mean squared norm* or the finite energy metric,
defined by $\left(\int_a^b \{f(x) - g(x)\}^2 dx \right)^{1/2}$

Assigning properties intuitively to the distance measures yields the abstract concept of a metric space. If $x, y, z \in X$ and d is a distance measure associated with X , then X is a metric space if the following properties hold:

1. *Reflexivity*: $d(x, x) = 0$.
2. *Positivity*: $d(x, y) > 0$ if $x \neq y$.
3. *Symmetry*: $d(x, y) = d(y, x)$.
4. *Triangle inequality*: $d(x, y) \leq d(x, z) + d(z, y)$.

4.1.1 Basic terminology

Some basic terminology in approximation theory is now introduced.

Dense: A set $F \subset D$, where D is a metric space, is said to be *dense* if for any $\epsilon > 0$ and any $g \in D$, there exists an $f \in F$, such that $|g - f| < \epsilon$.

Compact: If X is a metric space, $K \subset X$, is said to be *compact* if every sequence of points in K has a subsequence which converges to a point of K .

Closure: A set F in a metric space is said to be *closed* if the limit of every convergent sequence in F is also in F .

Best approximation: An element, $u_0 \in M$, where M is a non-empty set in a metric space, E , is said to be a *best approximation* to $f \in E$, with respect to a prescribed norm, if [31]

$$|f - u_0| = \min_{u \in M} |f - u|.$$

Existence set: In the previous definition, if a best approximation to f from M exists for all $f \in E$, then M is known as an *existence set*.

Universal approximator: An approximation scheme is called a *universal approximator* for a larger set of functions, Φ , if for any $\phi \in \Phi$ and any $\epsilon > 0$, there exists a function a in the set of approximating functions, A , $a \in A$, such that $|\phi - a| < \epsilon$.

4.1.2 Basic theorems

The three basic theorems in approximation theory are paraphrased below:

4.1.2.1 Theorem 1 Every compact set is an existence set [30].

4.1.2.2 Theorem 2 Every existence set is a closed set [32].

4.1.2.3 Theorem 3: Stone-Weierstrass Theorem If F is a set of continuous real-valued functions on a metric space D , $x \in D$ and F satisfies the following conditions:

1. *Identity*: The constant function $f(x) = 1$ is in F .
2. *Separability*: For any two points $x_1 \neq x_2$ in D , there exists an f in F such that $f(x_1) \neq f(x_2)$.
3. *Algebraic closure*: If f and g are any two functions in F , then fg and $af + bg$ are in F for any two real numbers a and b .

Then F is dense in D [33].

A function is called a universal approximator if it can approximate any other function in its metric space, to an arbitrary degree. The theorems listed above prove useful for deciding whether a particular candidate function is capable of universal approximation. The next section describes the exact manner in which the invariance transformation is recast as an interpolation problem.

4.2 The Interpolation Method for Invariance Transformation

Given two signals, X_A and X_B , characterizing the same phenomenon, two distinct initial features, $x_A(d, l, t)$ and $x_B(d, l, t)$, are chosen, where t represents an operational variable (for instance, permeability) and d and l represent defect related parameters (depth and length, respectively). x_A and x_B are chosen such that they have dissimilar variations with t . A systematic procedure is developed to obtain a feature, h , which is a function of x_A and x_B and invariant to the parameter, t . For

simplicity, x_A and x_B are considered to be dependent on only three parameters d , l and t . We need to find a function, f , such that

$$f\{x_A(d, l, t), x_B(d, l, t)\} = h(d, l) \quad (4.1)$$

Given two functions g_1 and g_2 , a sufficient condition to obtain a signal invariant to t can be derived as

$$h(d, l) \diamond g_1(x_A) = g_2(x_B) \quad (4.2)$$

where \diamond refers to a homomorphic operator. Then the desired t -invariant response is defined as

$$f(x_A, x_B) = g_2(x_B) \diamond g_1^{-1}(x_A) = h(d, l) \quad (4.3)$$

To implement this procedure, the functions h , g_1 and g_2 need to be obtained. Since h is a user-defined function, it can be chosen conveniently; for example, a linear combination of d and l . The function g_2 could be used to serve as a “conditioning” function, chosen to better condition the data. For example, if x_B contains widely spread values, g_2 can be chosen to be a logarithmic function. Having chosen h and g_2 arbitrarily, a suitable functional form is assumed for g_1 , whose coefficients are to be determined. This is done by solving a set of simultaneous equations at discrete points, (d_i, l_j, t_k) ; $i : 1$ to m ; $j : 1$ to n ; $k : 1$ to p , in the data space. That is,

$$h(d_i, l_j) \diamond g_1\{x_A(d_i, l_j, t_k)\} = g_2\{x_B(d_i, l_j, t_k)\} \quad (4.4)$$

should be solved exhaustively. Invariance is possible using this method only if a unique solution to (4.4) exists, which is dependent on an appropriate choice of g_1 . The implementation of this approach for permeability invariance is presented in this context.

4.3 The Permeability Invariance Scheme

Magnetization invariant signals can be derived using the following procedure. Two signals X_A and X_B that describe the same defect are represented by the axial (B_z) and the radial (B_r) component of the leakage flux density. As shown in Figure 2.5, changes in B-H curve primarily affect peak-peak amplitude of the flux leakage signal. Figure 4.1 shows the variation of peak-peak axial flux density, P_z , and the peak-radial flux density, P_r , with change in magnetization characteristic, for defect depths ranging from 20% of wall thickness to 80% wall thickness. Either of these peak-peak signal values cannot uniquely characterize a defect depth. We now have two quantities, $P_z(d, l, t)$ and $P_r(d, l, t)$, both characterizing the same defect, having dissimilar variations with respect to t . With the homomorphic operator \diamond represented by multiplication, a magnetization-invariant signal can be derived by applying Equation (4.3). We define the initial features, $x_A = [P_r, P_z, D_r]'$, where D_r is the peak-peak separation of the radial MFL signal, and $x_B = P_z$.

$$h(d, l) = f(x_A, x_B) = \frac{g_2\{P_z\}}{g_1\{P_r, P_z, D_r\}} \quad (4.5)$$

The function h is chosen to be a linear function of depth. g_2 is taken as an identity function, i.e. $g_2\{P_z\} = P_z$. An alternative interpretation can be given for the permeability invariance process described above. The variations of P_z and P_r with changes in B-H curves are dissimilar, and consequently, a ratio of these two quantities will not yield a value that is constant with t . The function g_1 , in effect, “warps” the permeability response of P_r to that of P_z , such that a ratio $P_z/g_1\{P_r, P_z, D_r\}$ is invariant to t . A warping transformation of this nature is called a Geometric Transformation and is extensively applied in image processing [25]. The

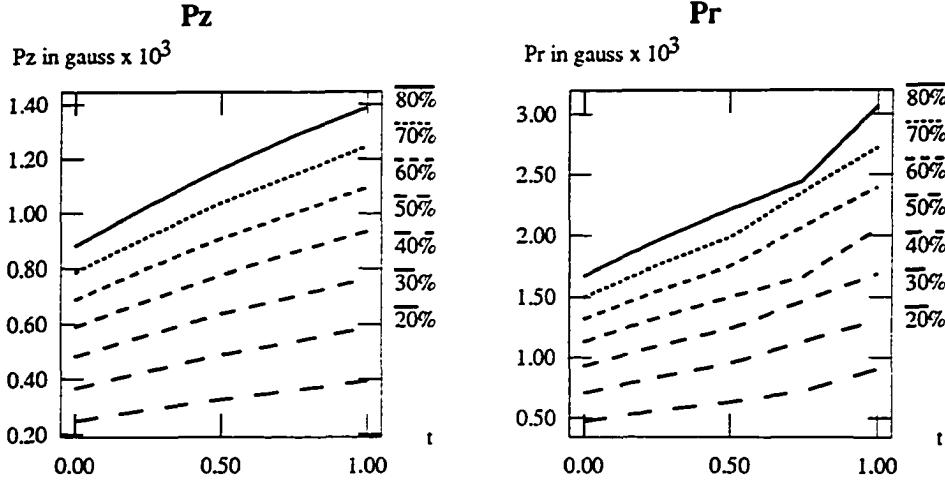


Figure 4.1: Variation of P_z and P_r with magnetization characteristic

entire axial flux density can be scaled by the t -invariant feature $P_z/g_1\{P_r, P_z, D_r\}$, so that a magnetization-invariant signal,

$$B_{zinv} = \frac{B_z - B_{zmin}}{B_{zmax} - B_{zmin}} \cdot h \quad (4.6)$$

is obtained. Figure 4.2 shows a flow-chart of the entire procedure.

Various choices of the “warping function”, g_1 lead to different methods of data interpolation. In general, interpolation techniques can be divided into phenomenological and non-phenomenological methods. Phenomenological methods derive their interpolating function from analytical models of the physical processes from which the interpolation data evolves. Non-phenomenological methods, on the other hand rely on mathematical functions not necessarily based on the underlying physics. A discussion of both these methods follows, with specific application to obtaining invariance to magnetization characteristic in MFL NDE signals.

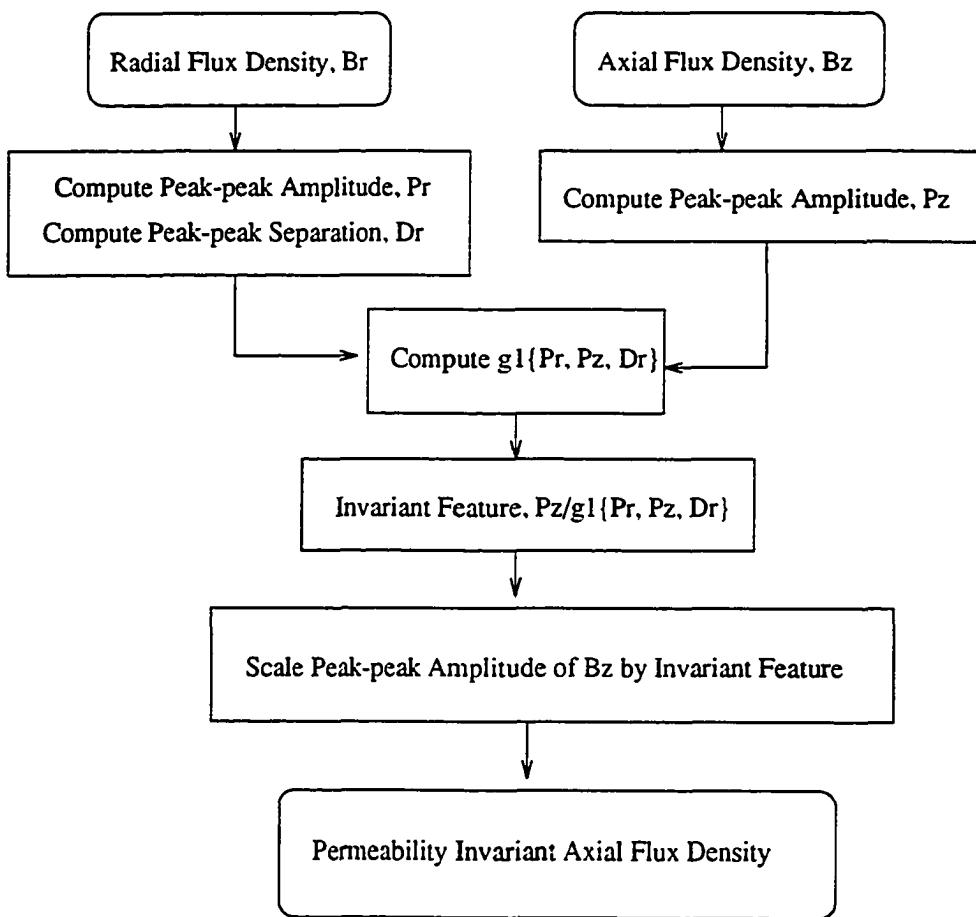


Figure 4.2: Flow-chart of the permeability invariance scheme

4.4 Phenomenological Methods

If the components of the leakage field can be expressed as a function of the defect dimensions, magnetization characteristic of the specimen and the applied field, the warping function can be exactly determined. However, the processes responsible for the origin of leakage fields are very complex and only approximate analytical expressions exist for simple geometries.

4.4.1 The Zatsepin and Shcherbinin model

The Zatsepin and Shcherbinin model approximates the magnetic leakage fields arising from rectangular defects using strip dipoles [34]. Fields are computed for a 2-D rectangular defect of length $2b$ and depth h , as shown in Figure 4.3. The x and y components of the leakage field, H are given by

$$\begin{aligned} H_x(x, y) &= 2\sigma_s \left\{ \arctan \left[\frac{h(x+b)}{(x+b)^2 + y(y+h)} \right] + \arctan \left[\frac{h(x-b)}{(x-b)^2 + y(y+h)} \right] \right\} \\ H_y(x, y) &= \sigma_s \log \left\{ \frac{[(x+b)^2 + (y+h)^2][(x-b)^2 + y^2]}{[(x-b)^2 + (y+h)^2][(x-b)^2 + y^2]} \right\} \end{aligned} \quad (4.7)$$

where σ_s is the surface density of magnetic “charge”. Using Equations (4.7), the warping function, g_1 , is derived as

$$g_1\{x_A\} = a \arctan[b \exp(cx_A)] \quad (4.8)$$

where the coefficients a , b and c are determined using Equation (4.4).

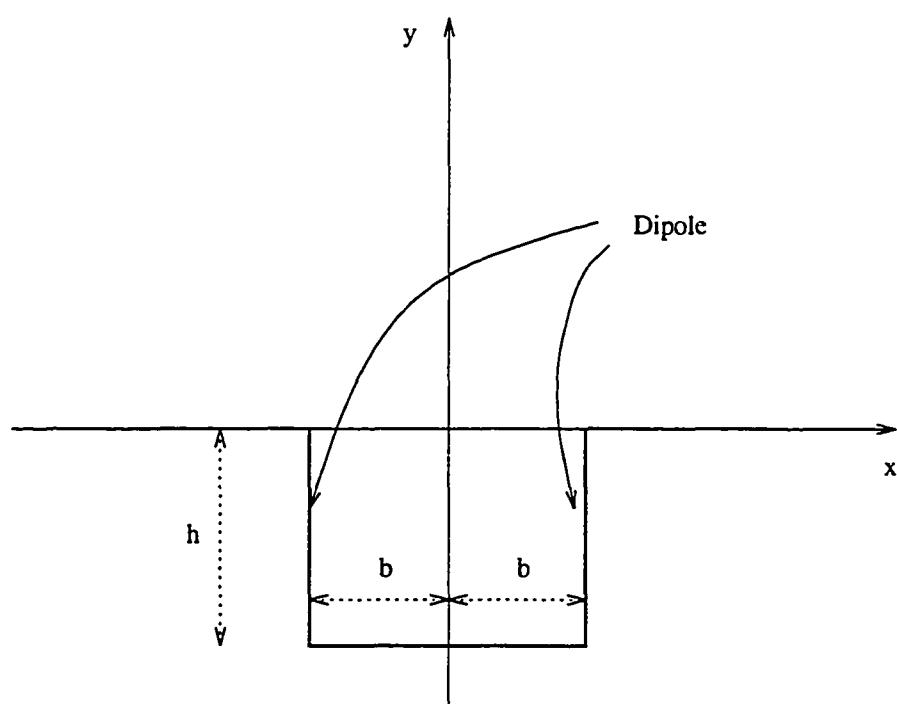


Figure 4.3: Strip dipole configuration

4.4.1.1 Implementation An implementation of the invariance transformation derived from the Zatsepin model is presented below. As an initial study, the various defect classes were all rectangular with varying depth and constant length. The parameters of Equation (4.8) are determined by solving the set of simultaneous equations (4.4). Data points of a single defect depth and varying pipe-wall permeability were used in Equation (4.4), since the Zatsepin warping transformation, which is not a universal approximation scheme, does not guarantee a solution for any arbitrary set of distinct input points. The transformation parameters a , b and c were derived for a 2 inch long, 60% deep (of wall thickness) defect with varying permeability. The transformation function was tested with MFL signals from 2 inch long, 10%, 30% and 40% deep defects of varying pipe-wall permeabilities. The permeability invariant features, h , are shown in Figure 4.4.

The invariance transformation can be quantified by the following measure:

$$\text{Figure of Merit} = \frac{\text{SNR after transformation}}{\text{SNR prior to transformation}}$$

where

$$\text{SNR} = \frac{\text{Mean of the function}}{\text{Standard deviation of the function}}$$

The Figure of Merit for the transformation defined by Equation (4.8) varies from 10.93 to 28.62. However, the discrimination between different depths is greatly reduced, as can be seen in Figure 4.4.

4.5 Non-phenomenological Methods

The phenomenological model described above is semi-empirical and makes gross approximations. The model cannot handle complicated geometries encountered in

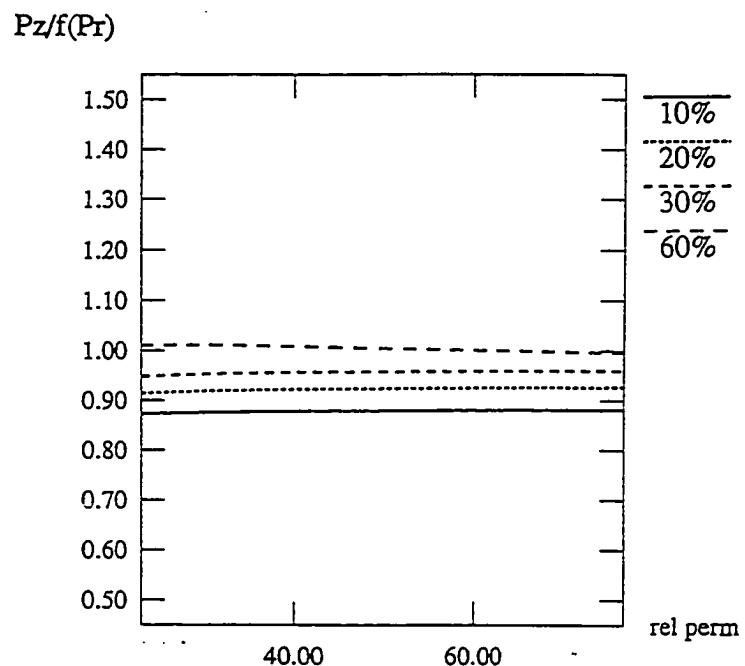


Figure 4.4: Permeability invariant feature derived using Zatsepin model

pipeline inspection applications. Non-phenomenological systems are concerned primarily with the analysis of the data. These techniques have been shown to be powerful in interpreting diverse data types and are extensively used in engineering applications. A variety of methods belonging to this class have been investigated and these are discussed below.

4.5.1 Simple functions

The category of simple functions encompasses combinations of quadratics, exponentials etc. Some of the functions, generated by commercial curve fitting software such as *TableCurve*[©] [35], that have been studied are:

$$\begin{aligned} g_1\{x_A\} &= ax_A^2 + bx_A + c \\ g_1\{x_A\} &= a + bx_A + cx_A^2 + \frac{d}{x_A} \\ g_1\{x_A\} &= a + \frac{b}{x_A} + c \exp(x_A) \\ g_1\{x_A\} &= a + bx_A \log(x_A) \end{aligned}$$

It should be noted that these functions are not universal approximators, since they do not satisfy the denseness criteria discussed in the previous section.

4.5.1.1 Implementation The parameters a , b , c and d of the warping functions shown above were found by applying Equation (4.4) to one class of defects (50 % deep defects) using the *TableCurve*[©] software. The “invariant” feature was then estimated using these coefficients for defects of other depths (20% and 80% deep defects). These coefficients could not be determined for several classes simultaneously

because — (a) Matrices resulting from the implementation of Equation (4.4) turn out to be near-singular, and (b) The commercial curve fitting software is not adequate for handling multidimensional surface fits. Consequently, the algorithm exhibited poor invariance outside the defect class for which the coefficients were determined. In some cases the invariant feature exhibited non-monotonic behavior, as shown in Figures 4.5 (a) through (d). Invariance schemes using such transformations can confuse subsequent defect characterization algorithms.

4.5.2 Feedforward networks for universal approximation

Any function can serve as a universal approximator if it satisfies the denseness and/or existence criteria described in the previous section. Multilayer feedforward networks, as shown in Figure 4.6, have been proven to satisfy the denseness criteria by applying the Stone-Weierstrass theorem, and can therefore approximate any continuous function to an arbitrary degree of accuracy [33] [36] [37] [38]. Shown in Figure 4.6 is a 3-layer network with an input vector $\underline{x} = [x_1, x_2]'$, activation functions, H_i , $i = 1, 2, 3$ and weights w_i , $i = 1, 2, 3$. Typically, the activation function can take one of the following two forms [39].

If the activation function contains an inner product of the form

$$H_i = \sigma(\mathbf{w}^T \cdot \mathbf{x} + \theta_i) \quad (4.9)$$

then, the network is of the Multilayer Perceptron (MLP) type.

If the activation function is comprised of an Euclidean norm of the form

$$H_i = \sigma(\|\mathbf{x} - \mathbf{a}_i\|) \quad (4.10)$$

then, the network is of the Radial Basis Function (RBF) type.

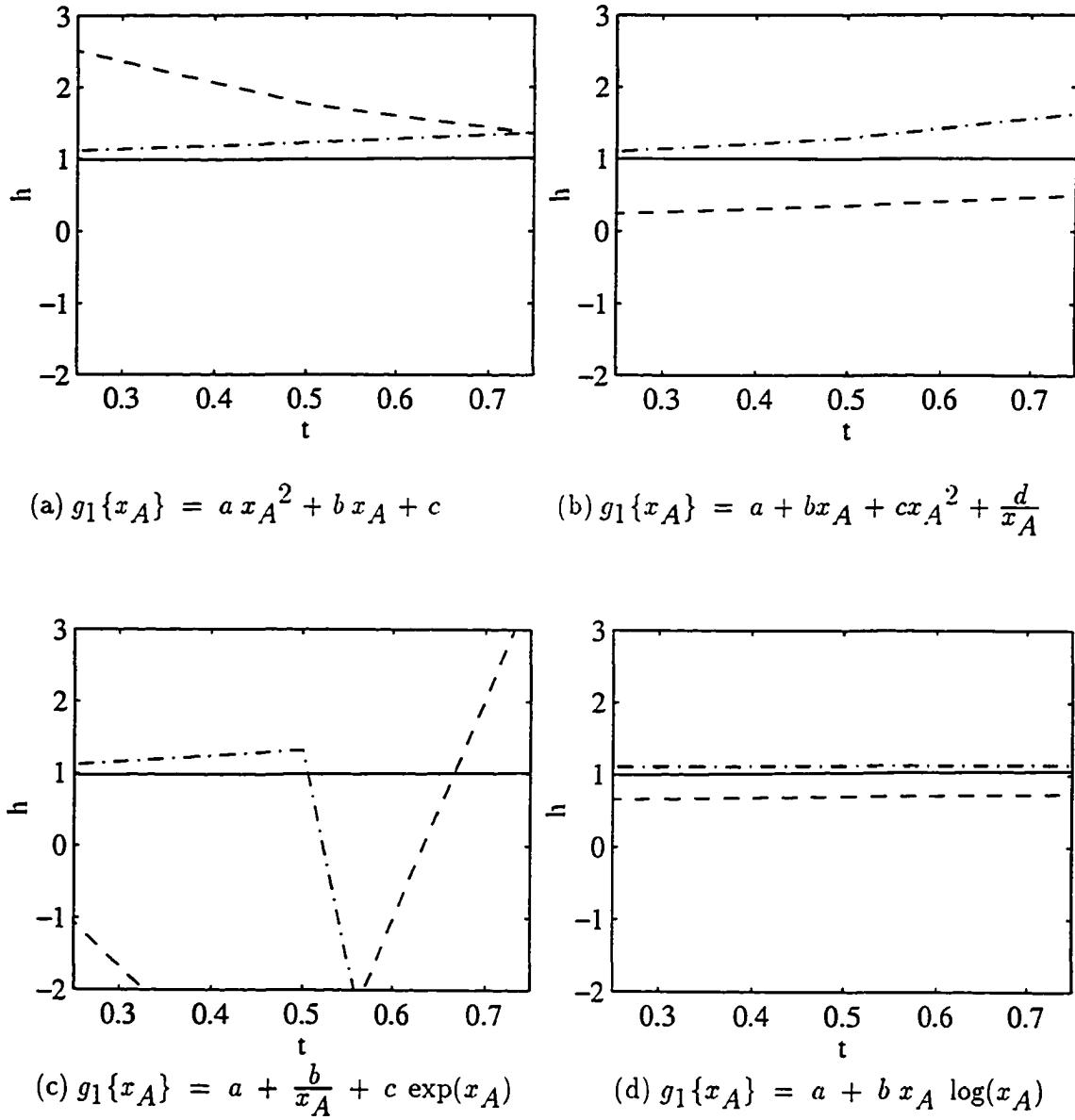


Figure 4.5: Permeability “invariant” feature derived using simple functions from the *TableCurve[©]* software. All defects are 2 inches long. - - - : 20% deep; — : 50% deep; - - - - : 80% deep

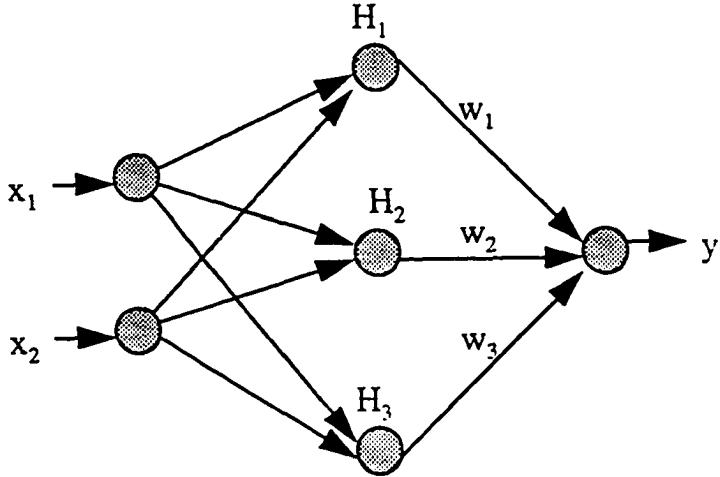


Figure 4.6: A typical multilayer feedforward neural network

Although both MLPs and RBFs have been shown to satisfy the denseness criteria, the RBFs satisfy both denseness and existence criteria [36] [31] [39]. This means, while MLPs and RBFs are arbitrarily good approximators, a network of the RBF type is the *best* approximator (See Appendices A and B). It will be shown later that invariance conditions are best met by RBF type networks. Other networks such as Fuzzy Basis Function (FBF) networks and Wavelet Basis Function (WBF) networks are variations on the RBF theme and these too possess the universal approximation property. The architecture of these networks and their region of application is discussed in later sections.

4.6 Radial Basis Functions

The permeability invariant feature is derived by performing the invariance transformation given by $h = x_B/g_1\{x_A\}$, where $x_B = P_z$ and $x_A = [P_z, P_r, D_r]'$. Suppose the warping function, g_1 is chosen such that

$$g_1 = \sum_{j=1}^m \lambda_j \phi(\|x_A - c_{A_j}\|) \quad (4.11)$$

where ϕ is a suitable “basis” function, c_A is a subset of all the points in x_A and $\|\cdot\|$ denotes a norm or distance measure. It has been shown that, given a set of distinct points, there exists a unique multiquadric surface as depicted in Equation (4.11) that interpolates the data [27]. In other words, the set of equations defined by Equations (4.4), (4.11) are always solvable at the discrete data points (d_i, l_j, t_k) , thus guaranteeing a t -invariant signal at these points. Moreover, an implementation of Equation (4.11) essentially involves the setting up of a radial basis function (RBF) neural network [40], to obtain the warping function g_1 . The architecture of the neural network is as shown in Figure 4.7. The network inputs are P_z , P_r and D_r . The network output is a function of P_r , P_z and D_r , $g_1\{P_r, P_z, D_r\}$, which divides P_z to yield a signal h that is constant in t .

The Gaussian basis function defined by

$$\phi(\|x_A - c_{A_j}\|) = \exp\left(-\frac{\|x_A - c_{A_j}\|^2}{2\sigma^2}\right) \quad (4.12)$$

is one of the most popular radial basis functions. Here, c_A 's are known as the basis centers and σ is the radius of the Gaussian kernel. The centers are usually chosen using a clustering algorithm, such as the K-means [41]. Clustering usually leads to

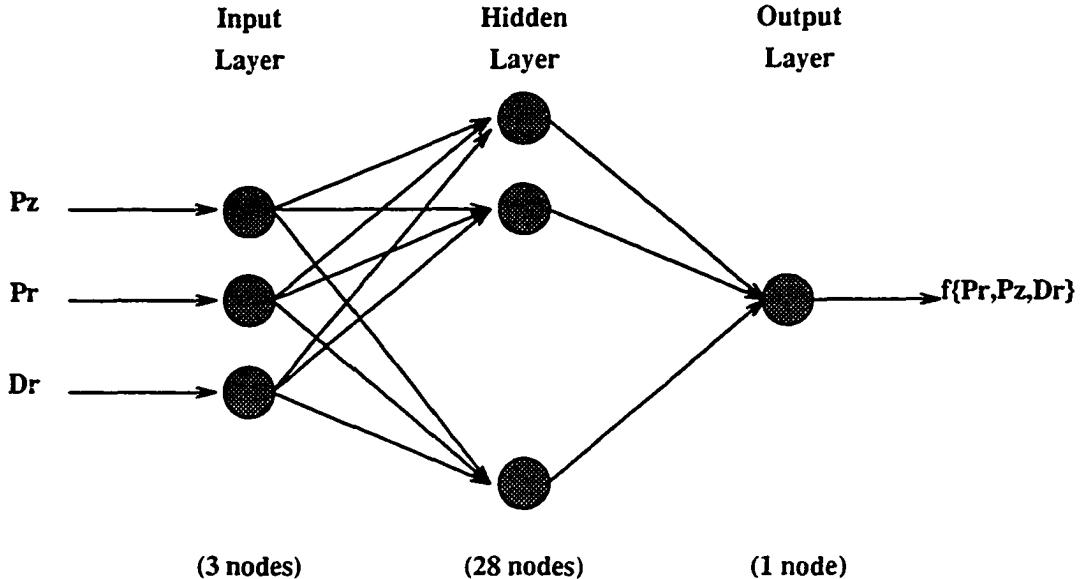


Figure 4.7: RBF network for the permeability invariance scheme

better global interpolation, i.e. the invariance transformation can operate effectively over varying defect depths (d). However, this occurs at the cost of local interpolation; which means the performance of the invariance transformation deteriorates for variation in B-H curves (t). The radius of the basis function controls the extent of interpolation and is typically determined from analyzing the variance of the data samples [42]. Alternatively, the RBF parameters may be determined using the orthogonal least squares (OLS) algorithm [43].

4.7 Fuzzy Inference Systems

Fuzzy logic employs fuzzy sets, to which objects may partially belong by means of membership functions [44]. The fuzzy sets are the variables in a fuzzy inference system and membership functions take on a value between 0 (the object does not

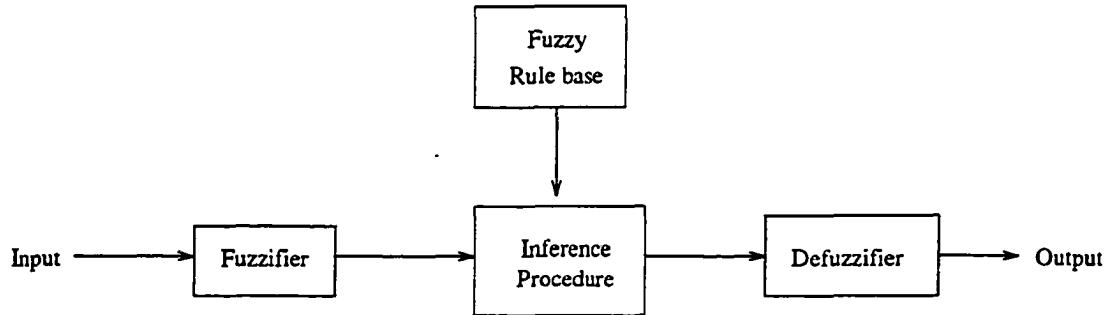


Figure 4.8: Fuzzy inference system

belong to the set) and 1 (the object completely belongs to the set). Fuzzy inference systems (FIS) are composed of the following four blocks — (a) Fuzzifier (b) Fuzzy rule base (c) Inference procedure and (d) Defuzzifier, as shown in Figure 4.8 [28]. The fuzzifier maps the input crisp sets to fuzzy sets and the defuzzifier does the opposite, to generate a crisp output given a fuzzy input. Information relating fuzzy input conditions to fuzzy output responses is contained in the rule base. The inference procedure determines the corresponding fuzzy maps.

The fuzzy rule base consists of IF-THEN conditional statements. For example, if x_1, x_2, \dots, x_m are fuzzy input variables characterized by membership functions, $A_{i1}, A_{i2}, \dots, A_{im}$ and y_1, y_2, \dots, y_n are the output variables characterized by membership functions $B_{i1}, B_{i2}, \dots, B_{in}$; the rule i is given by

IF

x_1 is A_{i1} AND x_2 is A_{i2}AND x_m is A_{im}

THEN

y_1 is B_{i1} AND y_2 is B_{i2}AND y_n is B_{in}

A variety of fuzzy inference systems can arise using different choice of operations for combining the variables. For instance, the AND operator can be chosen to be a product or *infimum* operator. The output defuzzification can be accomplished using the centroid or mode of the output fuzzy set. Also, various types of membership functions can be chosen.

Fuzzy inference systems can be shown to be capable of universal approximation since they satisfy the conditions of the Stone-Weierstrass theorem [45] [46]. In particular, if the product rule is used to specify the AND's in the IF part, a weighted average is used for the THEN part and the membership functions are Gaussian, then, the outputs of the FIS and a RBF neural network are identical (See Appendix C) [47]. The average of the product of the individual membership functions can be denoted as a fuzzy basis function (FBF) [46]. In which case, if

$$\mu_i = A_{i1} \cdot A_{i2} \cdot \dots \cdot A_{im} \quad (4.13)$$

the output of the FIS is given by

$$f(\mathbf{x}) = \frac{\sum_{i=1}^n \mu_i f_i}{\sum_{i=1}^n \mu_i} \quad (4.14)$$

where the f 's are the centroidal values of the output fuzzy sets. Commonly used membership functions are the triangular and the Gaussian functions [48]. The triangular membership function takes the form

$$A_{ij}(x_j) = 1 - 2 \frac{|x_j - c_{ij}|}{r_{ij}} \quad (4.15)$$

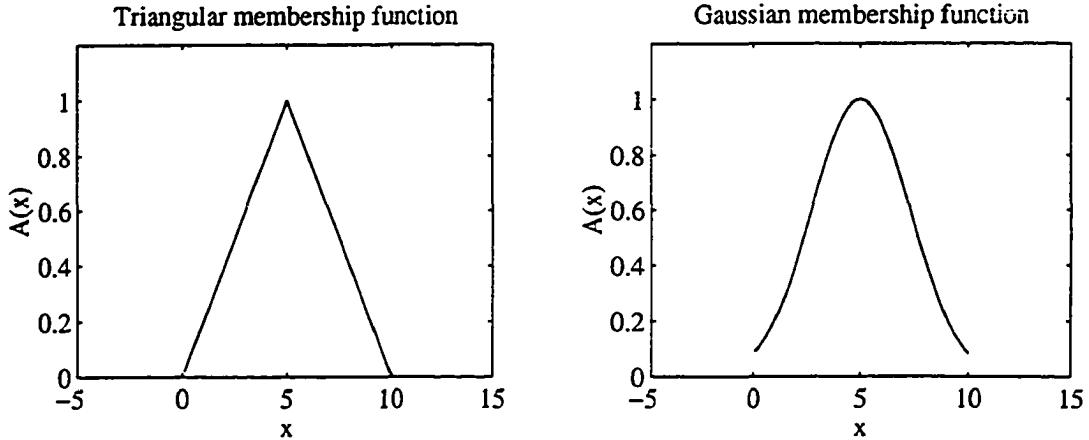


Figure 4.9: Triangular and Gaussian membership functions

where x_j is the j^{th} dimension of the input data vector and c_{ij} and r_{ij} represent the cluster center and radius in the j^{th} dimension for the i^{th} fuzzy inference rule. The Gaussian membership function is identical to the Gaussian RBF:

$$A_{ij}(x_j) = \exp \left[-\frac{(x_j - c_{ij})^2}{2r_{ij}^2} \right] \quad (4.16)$$

Figure 4.9 shows plots of the triangular and Gaussian membership functions.

The FIS is used as a universal approximator for invariant feature selection as follows. The input variables to the FIS are $x_A = [x_1, x_2, x_3] = [P_z, P_r, D_r]$. The membership functions A_{ij} , $j = 1, 2, 3$ are chosen as shown in Equation (4.16). The warping function $g_1(x_A)$ is chosen to be the output of the FIS, as shown in Equation (4.13). Then, the invariant feature $h = P_z/g_1(x_A)$.

Recent reports in neural network literature claim that FBF's combine the best characteristics of RBF (superior local interpolation) and multi-layer perceptron (superior global interpolation) networks [46]. However, since RBF's and FBF's have

been shown to be functionally equivalent, a mere fuzzification of the variables will not yield better results in invariance transformations. Improvements in the algorithm are possible by choosing the parameters defining the membership function appropriately. These parameters can be determined in several ways such as orthogonal least squares (OLS) algorithm [46], vector quantization with radius estimation using covariance ellipsoids [49] or gradient descent methods [48] [50]. Owing to the sparsity of the data vectors, the gradient descent method was deemed the most appropriate in the current application.

The iterative gradient descent algorithm operates in the following manner. Initial cluster centers are chosen using a K-means algorithm and the radii are assigned arbitrarily. The inference rules (membership functions) are tuned so as to minimize an error criterion, E , defined by

$$E = \frac{1}{2}(y - y^r)^2 \quad (4.17)$$

where y^r is the desired output and $y = f(x)$ is the output predicted by the FIS. The gradient of the error is found with respect to the centers $\left(\frac{\partial E}{\partial c_{ij}}\right)$, the radii $\left(\frac{\partial E}{\partial r_{ij}}\right)$ and the weights $\left(\frac{\partial E}{\partial f_i}\right)$. The error gradients for rules employing the Gaussian membership function are:

$$\begin{aligned} \left(\frac{\partial E}{\partial c_{ij}}\right) &= \frac{\mu_i}{\sum_{i=1}^n \mu_i} (y - y^r)(f_i - y) \frac{(x_j - c_{ij})}{r_{ij}^2} \\ \left(\frac{\partial E}{\partial r_{ij}}\right) &= \frac{\mu_i}{\sum_{i=1}^n \mu_i} (y - y^r)(f_i - y) \frac{(x_j - c_{ij})^2}{r_{ij}^3} \end{aligned}$$

$$\left(\frac{\partial E}{\partial f_i} \right) = \frac{\mu_i}{\sum_{i=1}^n \mu_i} (y - y^r) \quad (4.18)$$

The rule update equations are:

$$\begin{aligned} c_{ij}(k+1) &= c_{ij}(k) + K_a \left(\frac{\partial E}{\partial c_{ij}} \right) \\ r_{ij}(k+1) &= r_{ij}(k) + K_r \left(\frac{\partial E}{\partial r_{ij}} \right) \\ f_i(k+1) &= f_i(k) + K_f \left(\frac{\partial E}{\partial f_i} \right) \end{aligned} \quad (4.19)$$

where K_a , K_r and K_f are step lengths of the gradient descent algorithm and k denotes the iteration number. The iterations are allowed to proceed until the mean squared error falls below a certain minimum value. The step length can be made variable and the update mechanism modified so as to achieve faster convergence [51].

4.8 Multiresolution Networks: Wavelet Basis Functions

The adaptive FBF algorithm is a useful tool for universal approximation, which minimizes the trade-off required between local and global interpolation. However, in certain situations, it is required that the neural networks learn and operate at different resolutions of the input data. For example, in the case of the permeability invariance scheme, the network must be able to interpolate both for different B-H curves (t) as well as for different defect depths (d). The input-output mapping required of the network is shown in Figure 4.10. It can be seen that the signal variation in t and d occur at different resolutions. RBF and FBF networks operate at single resolutions and are, therefore, unable to simultaneously map defect depth and permeability. If

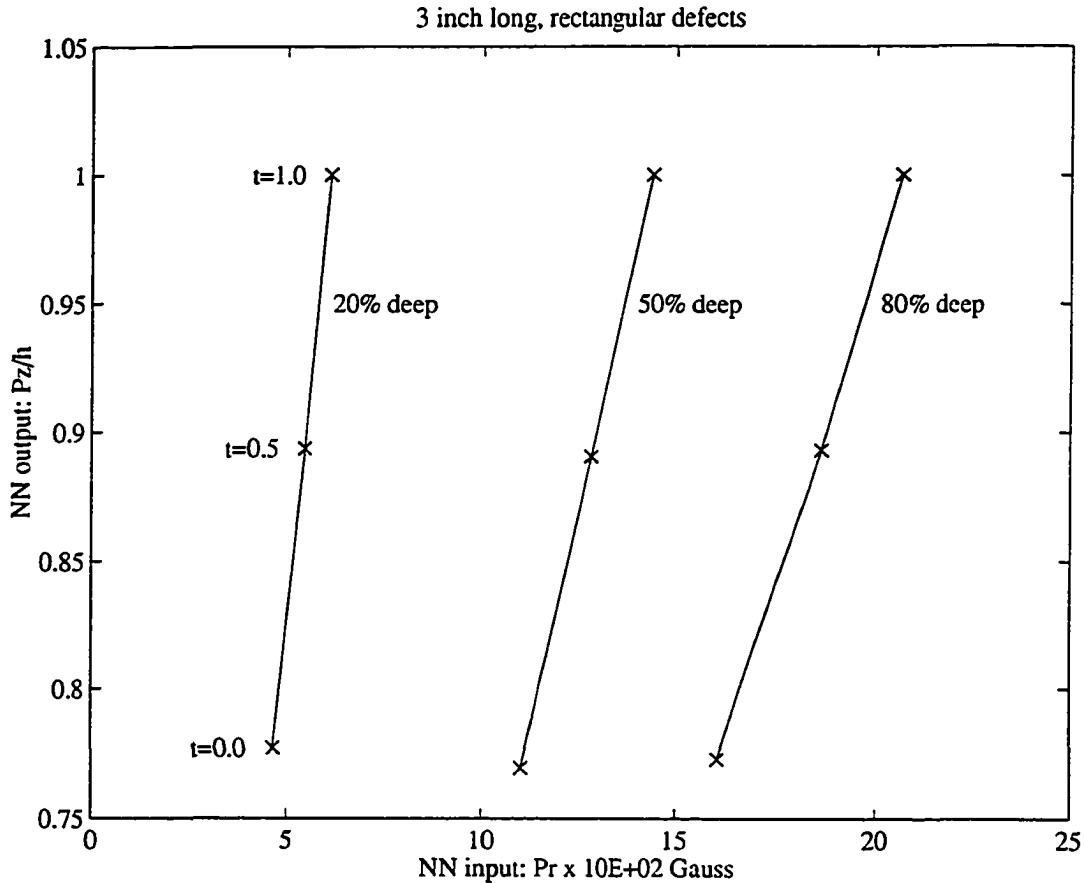


Figure 4.10: The input-output map of the permeability invariance scheme

an RBF or FBF is trained to interpolate in t , interpolation in d will be poor, and vice versa. Multi-resolution basis functions such as wavelets are potential candidates for such interpolation-based invariance transformations [29].

The wavelet basis network is based on the theory of multiresolution representation of a function by wavelet transforms [52] [53] [54]. It is known that functions can be represented as a weighted sum of orthogonal basis functions. Typical orthogonal functions are sinusoids, Legendre polynomials etc. However such basis functions do

not have a finite length of support and are therefore global approximators. A special class of functions, called wavelets, can be constructed as a basis for all square-integrable functions. Such a set of basis functions are both local and orthogonal. The wavelet transform represents functions in terms of fixed building blocks, the wavelets, at different scales and positions.

The wavelets are generated by the translation (positional shift) and dilation (scaling) of a single fixed function called the “mother wavelet” [55]. The mother wavelet is the function, $\psi(t)$, whose Fourier transform, $\Psi(\omega)$, satisfies the property

$$\int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{\omega} d\omega = C_\psi < \infty \quad (4.20)$$

This is known as the “admissibility condition” and implies that

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (4.21)$$

Translation and dilation operations on the mother wavelet creates a family of wavelets,

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi \left(\frac{t-b}{a} \right) \quad (4.22)$$

where a and b are the translation and dilation parameters respectively. The factor $\frac{1}{\sqrt{|a|}}$ equalizes the energy content of the family of wavelets with that of the mother wavelet.

A flurry of research activity in recent years, especially in the field of data compression, has led to the construction of different types of wavelets, with varying properties. Wavelets, such as the Mexican hat, Battle-Lemarie, Gaussian etc. can be explicitly defined. Others, including the Daubechies, Chui-Wang etc., do not lend themselves to explicit representation and are defined recursively in terms of low-pass

filter coefficients. Wavelets come in all flavors — orthogonal, semi-orthogonal¹ or nonorthogonal. The Mexican hat mother wavelet is defined as

$$\psi(t) = \frac{2}{\sqrt{3}}\pi^{-\frac{1}{4}}(1-t^2)\exp\left(-\frac{t^2}{2}\right) \quad (4.23)$$

Figure 4.11 shows the mother wavelet, and its translated and dilated versions. Wavelet decomposition of a function captures its time and frequency information among the translated and dilated wavelets.

Multiresolution representation of a function essentially involves describing it as a limit of successive approximations, each of which is a smoother version of the function. A multiresolution analysis of the Hilbert space, $L^2(R)$ is defined as a sequence of closed subspaces V_j of $L^2(R)$, $j \in Z$, with the following properties [52] [55]:

1. $V_j \subset V_{j-1}$. This condition states that the coarser approximations can be computed from the finer resolutions.
2. $f(t) \in V_j \Leftrightarrow f(2t) \in V_{j-1}$. This means that since the approximations at the various resolutions are similar, appropriate dilation of $f(t)$ results in the function in the next larger scale.
3. $f(t) \in V_0 \Leftrightarrow f(t+1) \in V_0$. Integer translation of $f(t)$ keeps the function at the same scale.
4. $\cap_{j \in Z} V_j = \{0\}$, and $\cup_{j \in Z} V_j = L^2(R)$. As the resolution decreases, the resulting approximation contains less and less information and converges to

¹Semi-orthogonal wavelets are orthogonal to their scaled and translated versions at different scales, but not within the same scale.

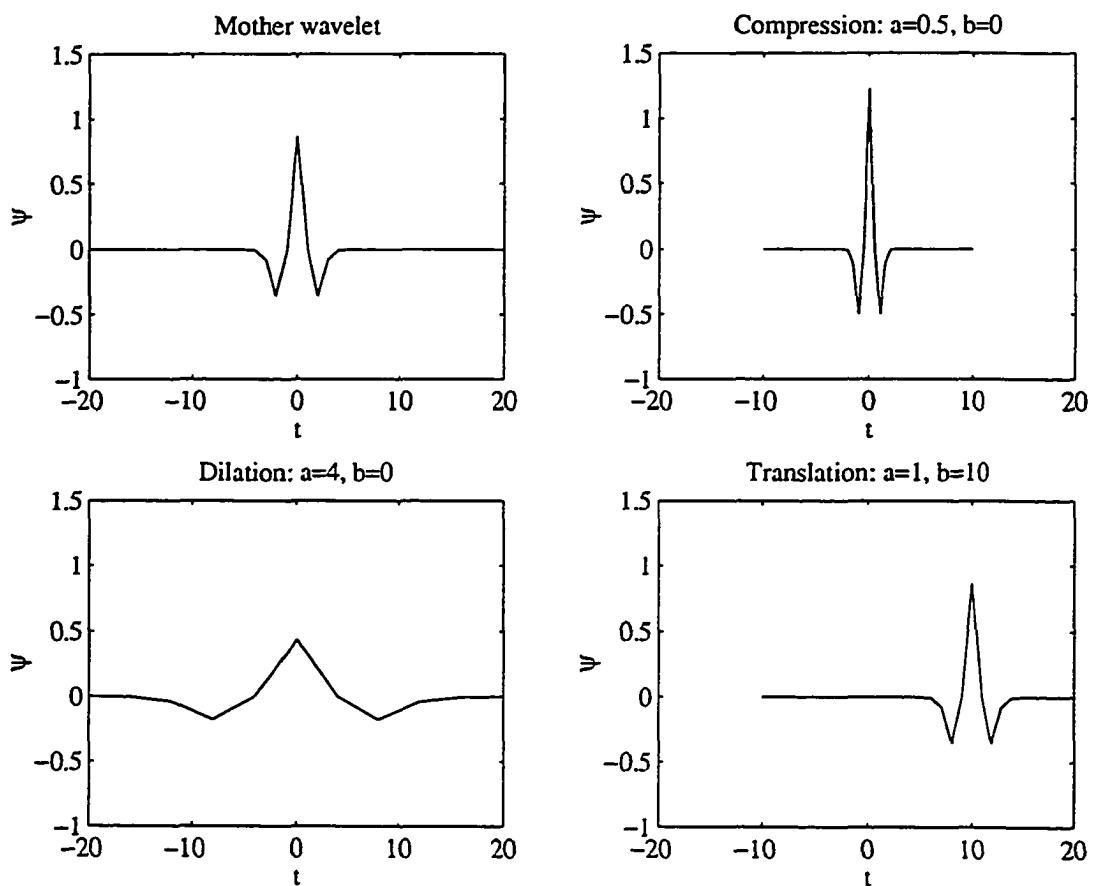


Figure 4.11: The Mexican hat wavelet and its dilated and translated versions

zero. As the resolution increases the approximation converges to the actual function. This is a denseness criterion and guarantees the completeness of the subspaces.

5. There exists a $\phi \in V_0$ such that $\{\phi(t-n)\}_{n \in \mathbb{Z}}$ is an unconditional orthonormal basis for V_0 .

Here, ϕ is known as the scaling function. A collection of scaling functions can be obtained by dilation and translation of $\phi(t)$ as follows:

$$\phi_{jn}(t) = \sqrt{2^{-j}} \phi(2^{-j}t - n) \quad (4.24)$$

If W_j is an orthogonal complement of V_j in V_{j-1} , it can be similarly shown that there exists a unique function, $\psi(t)$, the wavelet, which constitutes an orthonormal basis for W_j . A family of wavelet functions can be derived from translation and scaling as before.

$$\psi_{jn}(t) = \sqrt{2^{-j}} \psi(2^{-j}t - n) \quad (4.25)$$

The scaling functions, ϕ and the wavelets, ψ are related via the following set of equations. There exists $\{c_n\}$ such that

$$\begin{aligned} \phi(t) &= \sum_n c_n \phi(2t - n) \\ \psi(t) &= \sum_n (-1)^n c_{n+1} \phi(2t + n) \end{aligned} \quad (4.26)$$

We now construct a hierarchical, multiresolution representation for a function $f(t)$, starting from the coarsest resolution $f_L(t)$.

$$f_L(t) = \sum_{k=1}^{n_L} a_{Lk} \phi_{Lk}(t)$$

$$\begin{aligned}
f_{L-1}(t) &= f_L(t) + \sum_{k=1}^{n_L} d_{Lk} \psi_{Lk}(t) \\
f_{L-2}(t) &= f_{L-1}(t) + \sum_{k=1}^{2n_L} d_{L-1,k} \phi_{L-1,k}(t) \\
&\vdots \\
f(t) \approx f_0(t) &= \sum_{k=1}^{n_L} a_{Lk} \phi_{Lk}(t) + \sum_{m=1}^L \sum_{k=1}^{2^{L-m} n_L} d_{mk} \psi_{mk}(t) \quad (4.27)
\end{aligned}$$

Here the a 's and the d 's represent the projections of $f(t)$ on to the ϕ and ψ basis functions respectively. Equation (4.27) forms the foundation for development of the Wavelet Basis Function (WBF) network or Wave-Net. WBF networks mimic the wavelet transform by using different basis functions to map different frequency components of the input signal. The architecture of the network is as shown in Figure 4.8 [29]. It consists of two sets of nodes: the ϕ -nodes which capture information at a coarse resolution and the ψ -nodes, which model the function at finer resolutions. Network parameters are determined by training the two sets of nodes separately. The network is first trained with the ϕ -nodes to obtain a coarse fit. This procedure isolates the dc level of the input-output map. Subsequently, the network is trained with the ψ -nodes to compensate for the errors that have occurred as part of the first training. It can be shown that such networks can be used for universal approximation [56].

In the present application to the permeability invariance scheme, the ϕ -nodes can be used to interpolate along the depth and the ψ -nodes to resolve the function amongst the various B-H curves (see Figure 4.10). Recall that the permeabil-

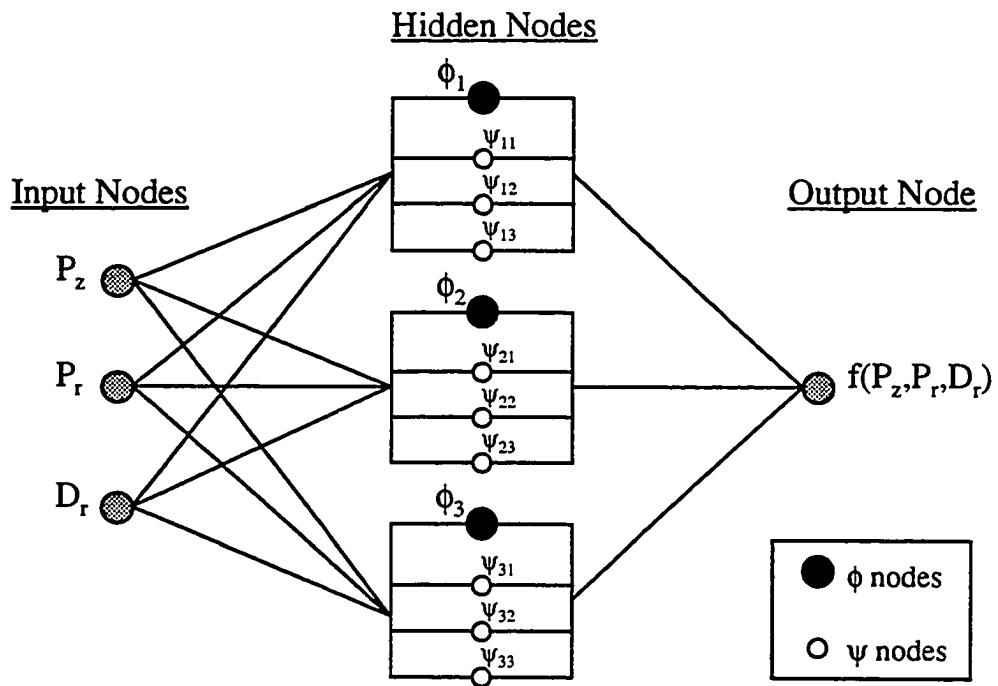


Figure 4.12: Wavelet basis function network

ity invariance scheme derives the invariant feature by performing a transformation $h = x_B/g_1(\mathbf{x}_A)$, where $x_B = P_z$, $\mathbf{x}_A = [P_z, P_r, D_r]$ and g_1 is the universal approximator; in this case, the WBF network. The warping function g_1 takes the form

$$g_1(\mathbf{x}_A) = \sum_{i=1}^n a_i \phi(\|\mathbf{x}_A - \mathbf{c}_{\phi_i}\|) + \sum_{i=1}^n \sum_{j=1}^m d_{ij} \psi(\|\mathbf{x}_A - \mathbf{c}_{\psi_{ij}}\|) \quad (4.28)$$

where c_ϕ and c_ψ are the centers of the WBF network for the ϕ and ψ nodes respectively, and are chosen to be subsets of all the data points in $\mathbf{x}_A = [P_z(d, l, t), P_r(d, l, t), D_r(d, l, t)]$. The centers c_ϕ are chosen from data points constant B-H curve (t) and varying depths (d). The centers c_ψ are chosen from data points wherein both the B-H curve (t) and depths (d) vary. First, the network is trained i.e. the weights a_i determined, using the c_ϕ centers alone. This procedure models the transformation function, g_1 , for depth variations of the input feature. The network is further trained by incorporating the c_ψ centers. The weights d_{ij} model the transformation function for input feature variations in permeability. In this manner, the interpolation at two different resolutions — depth and B-H curve is explicitly built into the invariance transformation.

While orthonormality and compact support are requirements for data compression problems, the current application requires basis functions that possess good approximation properties. Gaussian functions, used in RBF and FBF networks are good approximators and are used to solve regularization problems. Therefore scaling functions, ϕ , can be chosen to be Gaussian and the wavelet functions, ψ , to be its first derivative, as shown below.

$$\phi(x) = \exp\left(-\frac{(x - c)^2}{2\sigma^2}\right) \quad (4.29)$$

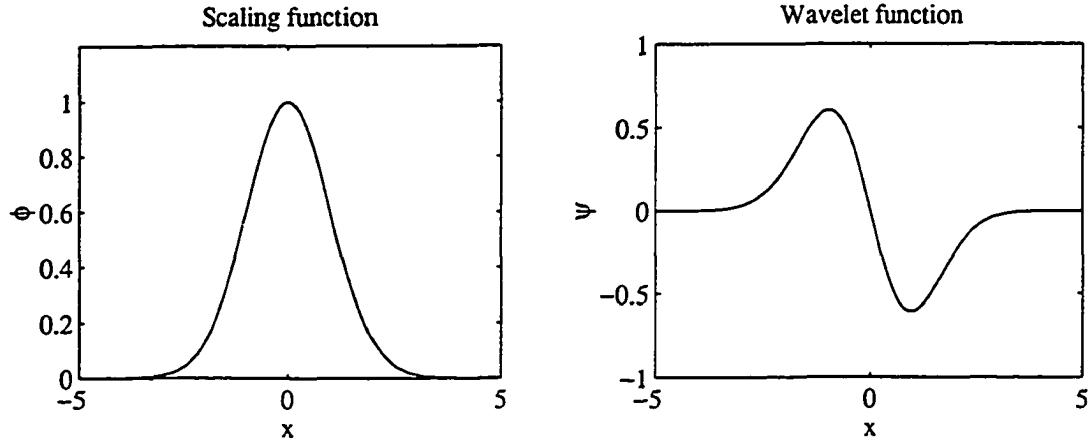


Figure 4.13: Gaussian scaling and wavelet functions

$$\psi(x) = -\frac{(x - c)}{\sigma^2} \exp\left(-\frac{(x - c)^2}{2\sigma^2}\right) \quad (4.30)$$

The Gaussian wavelet function, ψ , can be shown to satisfy the admissibility conditions as stated in Equations (4.20) and (4.21) (see Appendix D for proof). The Gaussian scaling function and wavelet is shown in Figure 4.13.

Wavelet networks can also be developed using multilayer perceptron type architectures [57] [58]. These, however, are unsuitable for the current application since MLP type networks do not possess the best approximation property (see Appendix B).

4.9 3-D Permeability Invariance Scheme

The permeability compensation scheme described earlier was developed for interpreting one-dimensional axial scans of the pipe, that yields 2-D information such as defect axial length and radial depth. However if a complete 3-D defect characterization is required, two-dimensional scan data along the axis and circumference of the

pipe must be analyzed. Such a scheme would require the circumferential component of the MFL signals, along with the axial and radial components, as input. Three dimensional finite element model simulations of the pipeline inspection system were conducted to study the variation of the leakage flux density arising from a 3-D defect.² Figures 4.14, 4.15 and 4.16 show the variation of the axial, radial and circumferential components of the MFL signal for three different B-H curves, respectively.

3-D permeability compensation is achieved by a simple modification of the existing scheme. The input vector $P_z/g_1\{P_r, P_z, D_r\}$ is augmented with another parameter, D_c , that specifies signal extent along the circumferential direction. Here, D_c equals the peak-peak separation of the circumferential component of the MFL signal. To compensate for permeability variations in a 2-D signal, the invariance transformation is specified as $P_z/g_1\{P_r, P_z, D_r, D_c\}$, or, $P_z/g_1\{P_c, P_z, D_r, D_c\}$, where P_c is the peak-peak amplitude of the circumferential leakage flux density.

4.10 Issues in Permeability Invariance

A brief discussion of some key issues related to the design methodology of the invariance transformation and the robustness of the mapping scheme is presented.

4.10.1 The necessity of using two signals (features)

The invariance technique developed here maps the variation of one feature, x_A , to another, x_B , by means of a warping transformation, g_1 . The invariant feature, h is obtained by dividing the two quantities: $h = x_B/g_1(x_A)$. Since this is a function

²The author thanks Mr. Jiatun Si for providing the 3-D finite element simulation results.

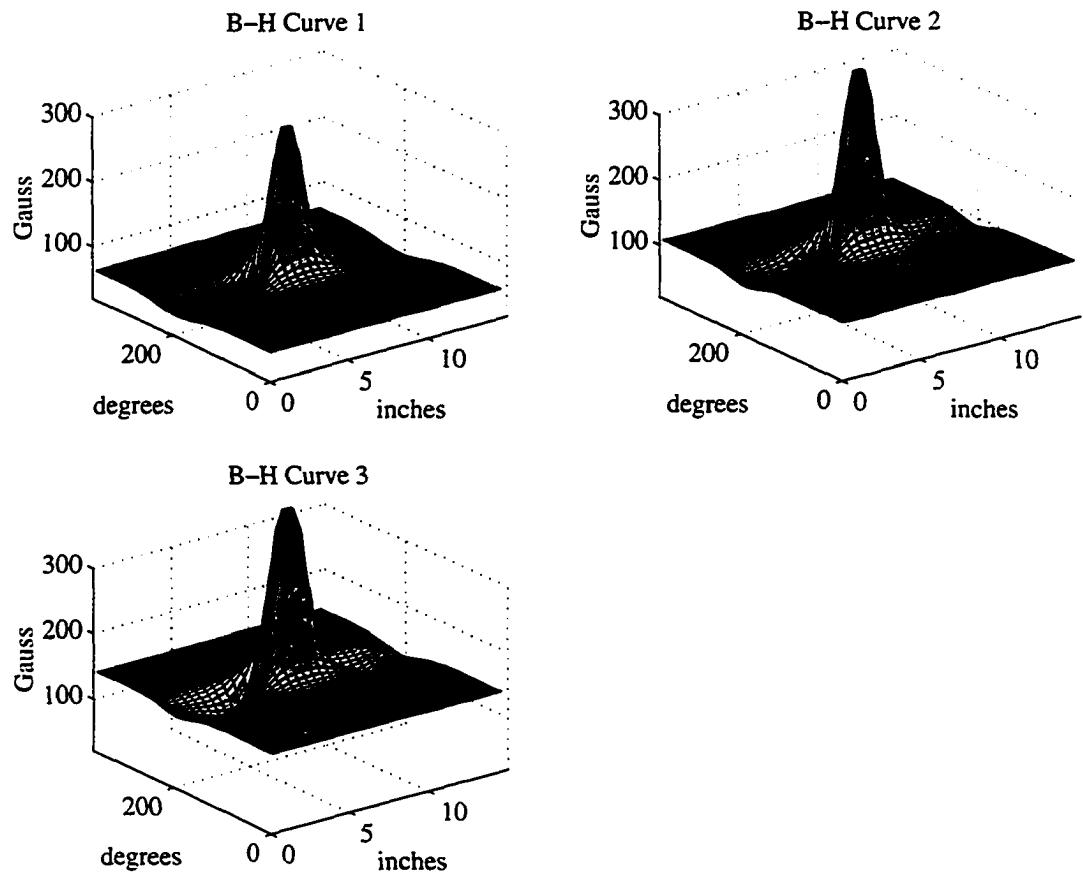


Figure 4.14: Variation of the axial component of the flux density for different B-H curves for a 3-D defect

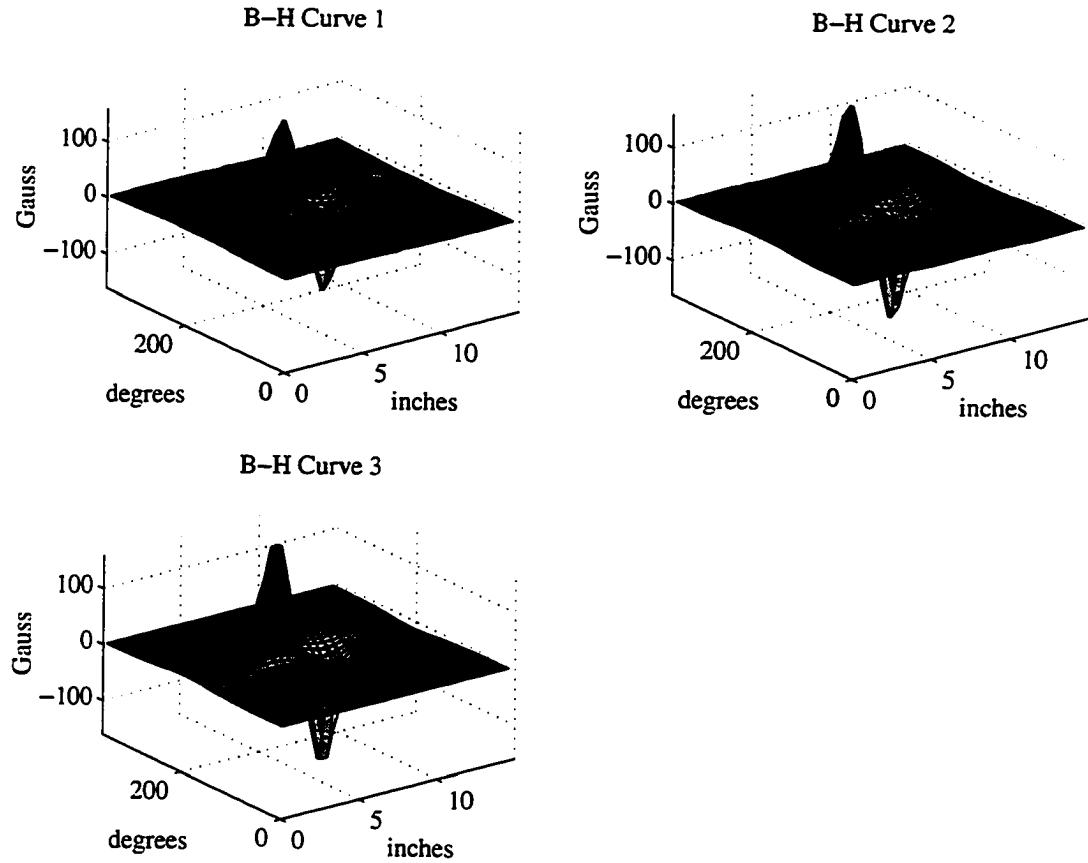


Figure 4.15: Variation of the radial component of the flux density for different B-H curves for a 3-D defect

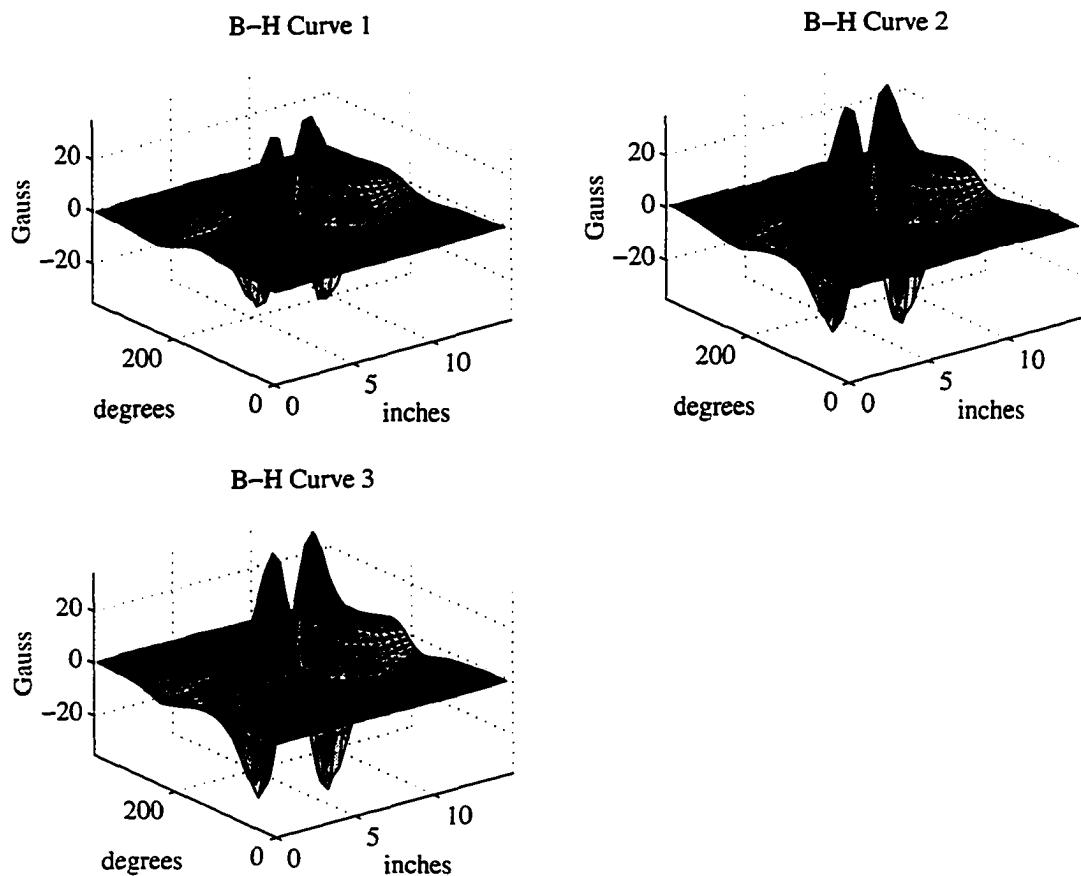


Figure 4.16: Variation of the circumferential component of the flux density for different B-H curves for a 3-D defect

mapping scheme, it is conceivable that only one varying feature, say x_A , could be used and mapped on to the invariant feature, h , i.e. $h = g_1(x_A)$. This type of functional mapping is possible, using the non-phenomenological methods described earlier. However, a mapping of this nature, wherein $g_1\{x_A(d, l, t)\}$ is made to map onto constant values, h , for different d , l and t , is *many-one*. A function mapping scheme cannot discern between the feature variation due to the d and t parameters. Like the conventional methods for invariant pattern recognition, such a scheme destroys any defect related information. To preserve the defect signature, a *one-to-one* mapping is desired and this can be accomplished by using two features. Each x_A can be mapped distinctly and separately via the transformation $g_1\{x_A(d, l, t)\} = x_B(d, l, t)/h(d, l)$. However, if only a single feature is available, a second feature can be derived from it, as shown in Chapter 6.

4.10.2 Robustness of the transformation

The robustness of the transformation scheme can be discussed under the following headings:

4.10.2.1 Training data The initial features selected must ensure that the transformation remains *one-to-one*. This is a requirement of the feedforward network wherein distinct input vectors are required to guarantee non-singular matrices for inversion. If the features chosen are not distinct, additional signal features; must be selected so that the transformation scheme does not degenerate to a *one-to-many* mapping. For example, in case of the 3-D permeability invariance scheme, the features, P_z , P_r and D_r are not adequate for ensuring distinct input vectors for signals

of different defect widths. Therefore a parameter, D_C , that varies with the width of the defect must be included for ensuring a *one-to-one* transformation.

4.10.2.2 Test data Whereas invariant mapping is guaranteed for the training data; similar theoretical proofs, that predict the output of neural network based schemes for testing data, do not exist. However, proper design of the network architecture and parameters; adequate diversity of training data; prudent choice of network type — all of these provide generalization capabilities with regard to local and global interpolation.

CHAPTER 5. INVARIANCE BY SIGNAL COMPENSATION

As described earlier, compensation for effects of sensor velocity is achieved using a restoration filter. This class of invariance schemes operate by a “learning” process similar to a neural network. The exact manner in which invariant signal parameters are isolated may be transparent to the designer. In defect characterization applications, the filter must operate consistently over a wide variety of defect classes. Depending on the application, the filter design procedure can be analytical, deterministic or stochastic.

5.1 Velocity Effects Compensation

5.1.1 Analytical methods

Analytical methods for designing restoration filters, as shown in Chapter 3, have simple formulations only when motion of the imaging device (sensor) does not alter the physical origins of the image (signal). However, this is not true of the distortion of the MFL signal induced by the velocity of the probe. The exact manner in which velocity affects the electromagnetic fields in the gas pipeline inspection geometry is still a subject of intense debate and forms part of contemporary dissertations. A completely analytical scheme for velocity compensation can be formulated only after the forward models have been completely understood and validated. In this study,

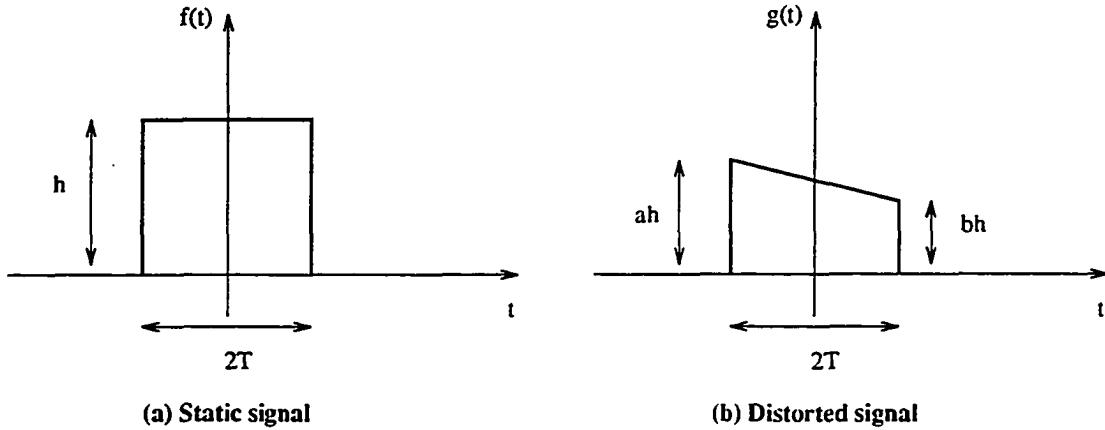


Figure 5.1: A simple velocity compensation filter model

a simple semi-empirical model of the effect of velocity on the MFL signal has been constructed. The undistorted (zero velocity) signal, $f(t)$, is modeled as a rectangular pulse of height h and width $2T$ as shown in Figure 5.1 (a). The distorted (velocity affected) signal, $g(t)$ is modeled as a trapezoid with sides ah and bh ; $a, b \leq 1$, as shown in Figure 5.1 (b)

The degradation filter, $H(\omega)$ is then constructed in the frequency domain by taking the ratio of the Fourier transforms of $g(t)$ and $f(t)$.

$$H(\omega) = \frac{G(\omega)}{F(\omega)} = \frac{(a+b)}{2} - \frac{(a+b)}{2\omega T} + j \frac{(a-b)}{2} \cos(\omega T) \quad (5.1)$$

The parameters a and b are functions of the probe velocity; their variations are determined either from experiment or numerical models.

5.1.1.1 Implementation Figure 5.2 shows the variation with velocity of the parameters a and b , of the analytical velocity compensation filter defined in Equation (5.1). These have been obtained from finite element studies. The performance of the

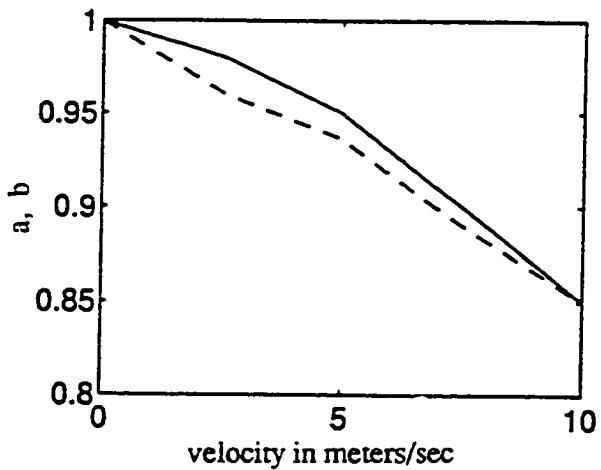


Figure 5.2: Variation of parameters a (—) and b (- -) with velocity

compensation filter can be seen in Figure 5.3.

5.1.2 Deterministic methods

In this approach, compensation filters are derived directly from data that have been procured from experiment or numerical models. The physical origins of the data are not modeled. In the case of MFL NDE, the situation is complicated by the fact that the underlying physical processes are themselves affected by the velocity. Hence, compensating for probe velocity is not straightforward, and deterministic methods are applied. The compensation/restoration filter $H_{\underline{\theta}}$ where the filter parameter, $\underline{\theta}$ depends on velocity, is constructed as follows. The static (less than 2mph) and velocity distorted signals for an ensemble of defect profiles are obtained, either from experiment or finite element studies. The compensation filter is derived by taking an

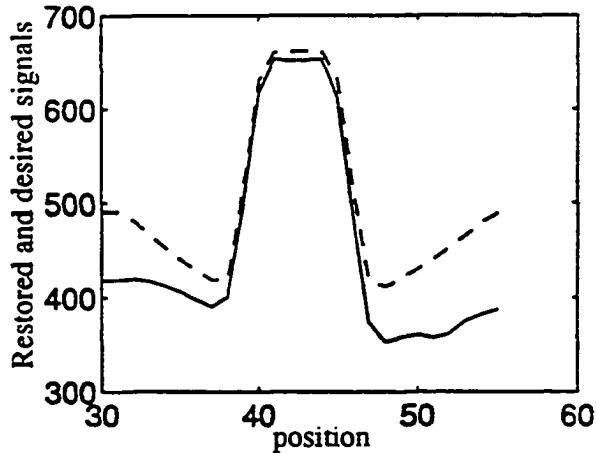


Figure 5.3: Velocity effects compensation using analytical filter. — restored, - - desired

average of the ratio of the Fourier transforms of the distorted and static signals, i.e.,

$$H(\omega) = \frac{1}{N} \sum_{i=1}^N \frac{G_i(\omega)}{F_i(\omega)} \quad (5.2)$$

where $F(\omega)$ and $G(\omega)$ are the Fourier transforms of the static and velocity distorted MFL signals respectively, and i indexes an ensemble of N defect profiles that have been used in the filter construction.

The significant advantage of the deterministic method lies in its simplicity. The technique simply involves the computation of the FFT of defect signatures at low and high velocities from a suitable range of defect profiles. However, while performing a point-to-point division of the frequency components of the low and high velocity MFL signals, zeros in the denominator of Equation (5.2) make the filter coefficients indeterminate at those points. To overcome this disadvantage, one can assign arbitrarily high values to the corresponding filter coefficients. However, this decreases

the accuracy of signal reconstruction.

5.1.3 Stochastic methods

Stochastic methods for filter design are based on the underlying signal statistics. When used in the context of velocity effects compensation, the velocity-affected signal is modeled as a waveform that is corrupted by noise and a deconvolution procedure is used for estimating the original “transmitted” signal. Signal statistics are defined by the auto-correlation matrix of the input velocity-affected signal and the cross-correlation vector between the filter input and the desired filter output (zero-velocity signal) [59]. Such a scheme is relevant only if the process is stationary and exhibits well behaved statistics. One criterion often used for determining the statistical behavior involves the computation of the eigenvalue spread of the auto-correlation matrix. Consider a discrete time stochastic process with an observation vector \mathbf{u} of the form

$$\mathbf{u}^T(n) = [u(n), u(n-1), \dots, u(n-M+1)] \quad (5.3)$$

A process is called wide-sense stationary if the the first and second order moments of the observation vector do not change with time (The theory of moments has been reviewed in Chapter 3). Most of the processes encountered in science and engineering belong to this category. The auto-correlation function, $r(k)$, for such a process is given by

$$r(k) = E\{u(n) u^*(n-k)\} \quad (5.4)$$

where E is the expectation operator and $*$ denotes a complex conjugation. In the case of an ergodic¹ process, the expectation operator takes on the following form.

$$r(k) = \lim_{N \rightarrow \infty} \sum_{n=1}^N u(n) u^*(n-k) \quad (5.5)$$

The auto-correlation matrix, \mathbf{R} , is defined as

$$\mathbf{R} = E\{\mathbf{u}(n) \mathbf{u}^H(n)\}. \quad (5.6)$$

Here, H indicates Hermitian transposition, i.e. transposition coupled with complex conjugation. For a wide-sense stationary process, the auto-correlation matrix can be written as

$$\begin{bmatrix} r(0) & r(1) & \cdots & r(M-1) \\ r(-1) & r(0) & \cdots & r(M-2) \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ r(-M+1) & r(-M+2) & \cdots & r(0) \end{bmatrix} \quad (5.7)$$

The eigenvalue spread, which is defined as the ratio of the maximum to the minimum eigenvalue of the auto-correlation matrix, is often used to gain an insight into the statistical behavior of the discrete time process. It can be shown that the largest eigenvalue of the correlation matrix, \mathbf{R} , is bounded by [59]

$$\lambda_{max} \leq \sum_{k=0}^{M-1} |r(k)| \quad (5.8)$$

A large eigenvalue spread, i.e., $\lambda_{max}/\lambda_{min}$, therefore indicates a highly correlated input signal. It will be shown later that the inverse of the auto-correlation matrix

¹For an *ergodic* process, the ensemble averages can be replaced by time (position) averages. An ensemble is a collection of realizations of the same process.

is employed to compute the filter coefficients. The degree of ill condition of the auto-correlation matrix is described by defining a condition number as follows.

$$\chi(\mathbf{R}) = \|\mathbf{R}\| \|\mathbf{R}^{-1}\| \quad (5.9)$$

where $\|\mathbf{R}\|$ is called the norm of the matrix \mathbf{R} and is a measure of the magnitude of the matrix. The spectral norm of \mathbf{R} is defined as

$$\|\mathbf{R}\| = (\text{largest eigenvalue of } \mathbf{R}^H \mathbf{R})^{(1/2)} \quad (5.10)$$

Based on this definition of the norm, it can be shown that the condition number of \mathbf{R} is given by

$$\chi(\mathbf{R}) = \frac{\lambda_{max}}{\lambda_{min}} \quad (5.11)$$

Therefore, an auto-correlation matrix with a large eigenvalue spread is ill-conditioned. In summary, higher the eigenvalue spread, stronger is the input correlation: and more ill-posed the filter construction problem (and vice versa).

The MFL signal is the result of a non-linear process as can be seen from Equation (2.1) and in theory, a non-linear filter should offer superior performance. Linear filters are used for a wide variety of signal processing applications and they take the form

$$y(n) = \sum_{i=0}^{N-1} w_i u(n-i) \quad (5.12)$$

where $u(n)$ is the filter input, $y(n)$ is the filter output and w_i are the weights or coefficients of the linear filter.

A commonly used approach in the design of a non-linear filter is to use the Volterra family of non-linear filters [61]. The name arises because the input-output relationship can be expressed in the form of a discrete Volterra series (a Taylor series with memory). A quadratic Volterra filter is described by the equation

$$y(n) = \sum_{i=0}^{N-1} a_i u(n-i) + \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} b_{ij} u(n-i) u(n-j) \quad (5.13)$$

where a_i and b_{ij} are the weights corresponding to the linear and quadratic terms respectively. Higher order filters are constructed using additional terms of the Volterra series.

The optimal filter weights are determined using one of the following criteria:

1. *Cost function minimization* In this approach, the filter weights are computed by minimizing a “cost” function that is defined to be the mean-square value of the estimation error. If $d(n)$ is the desired output of the filter, and $y(n)$ is the actual filter output, the estimation error, $e(n)$, is defined as

$$e(n) = d(n) - y(n) \quad (5.14)$$

and cost function is

$$J(\mathbf{w}) = E\{e(n)e^*(n)\}. \quad (5.15)$$

where \mathbf{w} is the weight vector, $[w_1, w_2, \dots, w_N]^T$. The dependence of the mean-squared error, J , on the weight vectors, \mathbf{w} can be pictured as a bowl-shaped surface with a unique minimum. The optimal filter weights correspond to the minimum value of this error-performance surface, obtained from the solution of

$$\left[\frac{dJ(\mathbf{w})}{d\mathbf{w}} \right] = \mathbf{0} \quad (5.16)$$

2. *Orthogonality principle* The orthogonality principle, which is equivalent to the minimum mean-squared error criterion, states that, the optimum filter weights are derived when the input vector and the estimation error are orthogonal to

each other. That is,

$$E[\mathbf{u}(n) \ e_o^*(n)] = \mathbf{0} \quad (5.17)$$

where $e_o(n)$ is the estimation error associated with the optimum filter weights.

Both of the above principles can be used in either a fixed or adaptive algorithm to compute the optimal filter weights.

5.1.4 Fixed scheme

In a fixed scheme, the filter weights are directly determined using the criteria given in Equations (5.16) or (5.17). The pertinent matrices are calculated and inverted.

5.1.4.1 Linear filter The cost function criterion can be used to determine the optimal weights for a fixed linear filter. The mean-squared error for a linear filter takes the form

$$J(\mathbf{w}) = E \left[\left(d(n) - \mathbf{w}^H \mathbf{u}(n) \right) \left(d^*(n) - \mathbf{u}^H(n) \mathbf{w} \right) \right] \quad (5.18)$$

which can be rewritten as

$$J(\mathbf{w}) = \sigma_d^2 - \mathbf{p}^H \mathbf{w} - \mathbf{w}^H \mathbf{p} + \mathbf{w}^H \mathbf{R} \mathbf{w} \quad (5.19)$$

where σ_d^2 is the variance of the desired response

$$\sigma_d^2 = E[d(n) \ d^*(n)] \quad (5.20)$$

and \mathbf{p} is the cross-correlation vector between the filter input and the desired output

$$\mathbf{p} = E\{\mathbf{u}(n) \ d^*(n)\}. \quad (5.21)$$

which can be expanded for the M^{th} order filter

$$\mathbf{p} = [p(0), p(-1), \dots, p(1 - M)]^T. \quad (5.22)$$

The gradient of the mean squared error is given by

$$\nabla = -2\mathbf{p} + 2\mathbf{R}\mathbf{w} \quad (5.23)$$

When filter weight vector is optimum, i.e. $\mathbf{w} = \mathbf{w}_o$, the error gradient is equal to the null vector, as given in Equation (5.16). Therefore, from Equation (5.23), the optimal linear filter weights are computed as

$$\mathbf{w}_o = \mathbf{R}^{-1}\mathbf{p} \quad (5.24)$$

Linear filters of this type are also called Wiener filters.

5.1.4.2 Quadratic filter The optimal weights for a second order Volterra filter can be determined using the orthogonality principle. By an extension of the orthogonality principle, the estimation error of the optimal Volterra filter is orthogonal not only to the filter input, but also all possible products of the input samples. Therefore, for a quadratic filter of the form

$$y(n) = \sum_{i=0}^{N-1} a_i u(n-i) + \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} b_{ij} u(n-i)u(n-j) \quad (5.25)$$

an application of the orthogonality principle gives

$$E[e(n)u(n-i)] = 0 \quad (5.26)$$

$$E[e(n)u(n-i)u(n-j)] = 0 \quad (5.27)$$

For simplicity, all quantities are assumed real. Substituting Equation (5.25) in Equations (5.26) and (5.27) and accounting for the symmetry of the quadratic kernel,

$$\begin{aligned}
 & \sum_{i=0}^{N-1} a_i E[u(n-i)u(n-l)] + \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} b_{ij} E[u(n-i)u(n-j)u(n-l)] \\
 & \quad = E[y(n)u(n-l)], \\
 & l = 0, 1, \dots, N-1 \\
 & \sum_{i=0}^{N-1} a_i E[u(n-i)u(n-l)u(n-m)] + \\
 & \quad \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} b_{ij} E[u(n-i)u(n-j)u(n-l)u(n-m)] = E[y(n)u(n-l)u(n-m)], \\
 & l = 0, 1, \dots, N-1; \quad m = l, \dots, N-1 \\
 & \tag{5.28}
 \end{aligned}$$

The solution of the system of equations (5.28) gives the coefficients, a_i and b_{ij} which define the second order Volterra filter. This requires the computation of the inverse of a matrix that is populated by the second-, third-, and fourth-order moments of $u(n)$. Also, there exists a coupling between the linear and quadratic filter coefficients. However, the problem can be simplified with the following assumptions:

1. The input vector has a zero mean.
2. The input signal is Gaussian.

If the input vector, \mathbf{u} , is assumed to have a zero mean then the filter output can be written as [62]

$$y(n) = \sum_{i=0}^{N-1} a_i u(n-i) + \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} b_{ij} [u(n-i)u(n-j) - r_u(i-j)] \quad (5.29)$$

where $r_u(k) = E[u(n)u(n - k)]$ denotes the auto-correlation of \mathbf{u} . Equation (5.29) can be rewritten as

$$y(n) = \mathbf{a}^T \mathbf{u}(n) + Tr\{\mathbf{B}[\mathbf{u}(n)\mathbf{u}^T(n) - \mathbf{R}_u]\} \quad (5.30)$$

where

$$\mathbf{a} = [a_0, a_1, \dots, a_{N-1}]^T. \quad (5.31)$$

Tr denotes the trace of the matrix,

$$\mathbf{B} = \{b_{i,j}\} = \begin{bmatrix} b_{0,0} & \dots & \dots & b_{0,N-1} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ b_{N-1,0} & \dots & \dots & b_{N-1,N-1} \end{bmatrix}, \quad (5.32)$$

and \mathbf{R}_u is the auto-correlation matrix of \mathbf{u} , as defined in Equation (5.7). If $d(n)$ is the desired output of the filter, the cross-correlation vector, \mathbf{p} is defined as in Equation (5.21) and the cross-bicorrelation function, t_{du} , is defined as

$$t_{du}(i,j) = E[d(n)u(n-i)u(n-j)] \quad (5.33)$$

The cross-bicorrelation function measures the third-order statistical dependency between $d(n)$ and $u(n)$, which is required for computing the optimum quadratic filter weights. The cross-bicorrelation matrix is defined as

$$\mathbf{T}_{du} = \{t_{du}(i,j)\} = \begin{bmatrix} t_{du}(0,0) & \dots & \dots & t_{du}(0,N-1) \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ t_{du}(N-1,0) & \dots & \dots & t_{du}(N-1,N-1) \end{bmatrix}, \quad (5.34)$$

Applying the orthogonality principle, the optimum Volterra filter should satisfy the following matrix equations.

$$E[\mathbf{u}(n)d(n)] = E[\mathbf{u}(n)\mathbf{a}^T \mathbf{u}(n) + \mathbf{u}(n)Tr\{\mathbf{B}[\mathbf{u}(n)\mathbf{u}^T(n) - \mathbf{R}_u]\}] \quad (5.35)$$

$$E[\mathbf{u}(n)\mathbf{u}^T(n)d(n)] = E[\mathbf{u}(n)\mathbf{u}^T(n)\mathbf{a}^T \mathbf{u}(n) + \mathbf{u}(n)\mathbf{u}^T(n)Tr\{\mathbf{B}(\mathbf{u}(n)\mathbf{u}^T(n) - \mathbf{R}_u)\}] \quad (5.36)$$

It can be shown that

$$\begin{aligned} E[\mathbf{u}(n)\mathbf{a}^T \mathbf{u}(n)] &= \mathbf{R}_u \mathbf{a} \\ E[\mathbf{u}(n)Tr\{\mathbf{B}[\mathbf{u}(n)\mathbf{u}^T(n) - \mathbf{R}_u]\}] &= \mathbf{0}_{(N \times 1)} \\ E[\mathbf{u}(n)\mathbf{u}^T(n)\mathbf{a}^T \mathbf{u}(n)] &= \mathbf{0}_{(N \times N)} \end{aligned}$$

Also, if x_1, x_2, x_3 and x_4 are zero-mean, jointly Gaussian processes, then

$$E[x_1x_2x_3x_4] = E[x_1x_2]E[x_3x_4] + E[x_1x_3]E[x_2x_4] + E[x_1x_4]E[x_2x_3] \quad (5.37)$$

Since \mathbf{B} is symmetric, we have

$$E[\mathbf{u}(n)\mathbf{u}^T(n)Tr\{\mathbf{B}(\mathbf{u}(n)\mathbf{u}^T(n) - \mathbf{R}_u)\}] = 2\mathbf{R}_u \mathbf{B} \mathbf{R}_u \quad (5.38)$$

Consequently, Equations (5.35) and (5.36) can be written as

$$\mathbf{p} = \mathbf{R}_u \mathbf{a}$$

$$\mathbf{T}_{du} = 2\mathbf{R}_u \mathbf{B} \mathbf{R}_u$$

Therefore, the design equations for the optimal second order Volterra filter are

$$\begin{aligned} \mathbf{a}_o &= \mathbf{R}_u^{-1} \mathbf{p} \\ \mathbf{B}_o &= \frac{1}{2} \mathbf{R}_u^{-1} \mathbf{T}_{du} \mathbf{R}_u^{-1} \end{aligned} \quad (5.39)$$

It can be seen that the optimal weights for the linear term of a second order Volterra filter are the same as those for a Wiener filter, i.e. $\mathbf{a}_o = \mathbf{w}_o$.

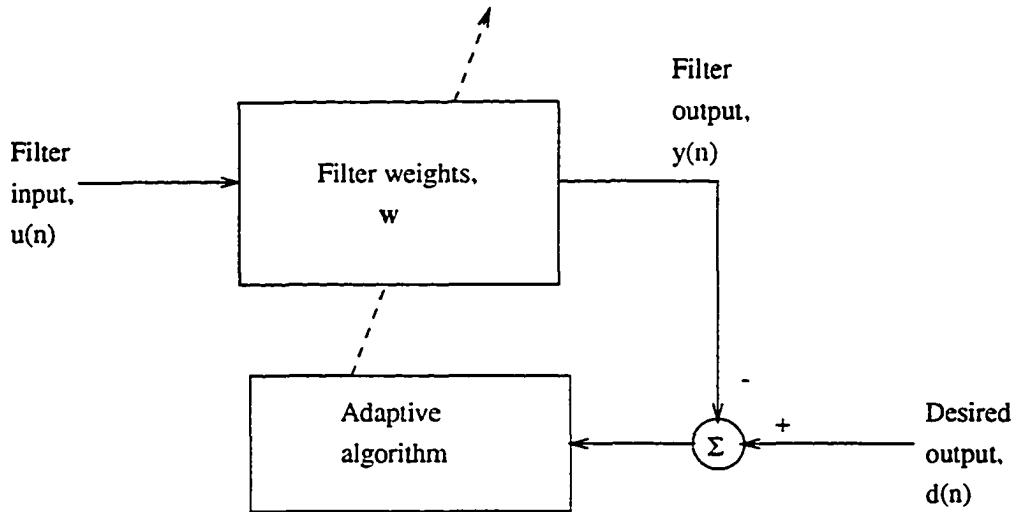


Figure 5.4: Adaptive filter

5.1.5 Adaptive scheme

Adaptive algorithms determine the weights in a iterative manner starting from an initial guess; the weights converge to the optimal value as the iterations proceed. Figure 5.4 shows the structure of a typical adaptive digital filter, where typical adaptive algorithms include the gradient descent, least mean squares and the recursive least squares techniques.

In the current application, the adaptive algorithm has the following three features:

1. Direct determination of the filter weights requires calculation and inversion of the correlation matrices. As the ensemble of defect signatures becomes large, these operations become computationally intensive. Adaptive filters eliminate the need for matrix inversion and hence reduce the computational burden.

2. The restoration filter is designed by using a comprehensive set of defect signatures. However, the input data to the filter can arise outside this “training set”. Exact determination of filter weights or reducing the error to zero during estimation of the filter weights, using the training data often results in poor performance for new data. This is usually handled in optimization problems by trading bias for variance. The resulting estimation error does not have a zero mean, but a smaller covariance matrix [60] and the mean squared error is reduced. Adaptive algorithms smooth out sharp variations in the filter weight vector which in turn helps in achieving better global interpolation. The filter is therefore more robust.
3. An adaptive algorithm can be used to adapt the filter weights to the changing defect profile. The MFL signal from each defect can be treated as a single data set, for which the filter coefficients are determined. These coefficients are then used to update an average weight vector that “learns” with time. The adaptive filter is therefore more flexible.

In the approach used here, the stochastic gradient descent algorithm updates the filter weights iteratively based on the value of the estimation error, defined in Equation (5.14), for the current iteration. The update mechanism for the Wiener filter weights using a stochastic gradient descent algorithm is given by

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{1}{2}\mu[-\nabla(J(n))] \quad (5.40)$$

where $\mathbf{w}(k)$ is the weight vector at iteration k , μ is the step size parameter and $\nabla(J(n))$ is the gradient of the cost function at time (position) n . For a Wiener filter,

substituting for $\nabla(J(n))$ from Equation (5.23), the update mechanism is written as

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu[\mathbf{p} - \mathbf{R}\mathbf{w}(k)]. \quad (5.41)$$

Other adaptive algorithms, such as the LMS, estimate the value of the gradient vector, $\nabla(J(n))$, without explicitly computing the auto-correlation matrix, \mathbf{R} , and the cross-correlation vector, \mathbf{p} .

The implementation of an adaptive algorithm for the estimation of filter coefficients consists of the following steps:

1. Initialize the elements of the weight vector to arbitrary values (usually zero).
2. Compute the filter output.
3. Determine the estimation error and the gradient of the cost function.
4. Update the filter weights.
5. Check if the cost function falls below a pre-defined minimum value, go to step 2.

As the iterations proceed, when the cost function falls below a certain arbitrarily small value, the algorithm is said to have converged. As mentioned earlier, the error surface has a unique minimum, which corresponds to the location of the optimum filter weights. This minimum error is called the global minimum. In practice, a multi-dimensional error surface has a complex shape. It is eminently possible that the convergence criterion could halt the iterations when the error, although below the specified value, does not correspond to the global minimum. Such minima are termed local minima and are often responsible for deterioration of the performance

of the adaptive filter. The problem can sometimes be overcome by choosing optimal step sizes that enable the algorithm to “jump” out when trapped in a local minimum.

5.2 Pre-processing MFL Signals for Velocity Effects Compensation

A stochastic or deterministic approach for the development of the restoration filter requires signals originating from the same defects at two different tool velocities. However, consecutive runs of the pig will not necessarily originate at identical sensor location and orientation. The effect of changes in the origin of sampling was shown in Chapter 2. Furthermore, data files obtained for the same defects at two velocities are shifted both axially and circumferentially relative to each other. The high and low velocity signals have to be registered before the restoration filter can be constructed. This problem is more evident in the case of 2-D scans, wherein 3-D compensation/characterization is to be performed. As a result of these two experimental artifacts, the MFL signal that is obtained from a pigging operation has to be pre-processed, prior to estimating the restoration filter weights. Analytical and stochastic techniques to accomplish this are discussed below.

5.2.1 Analytical method: Compensation for variation in sampling origin

The compensation for variation in the sampling origin of the MFL signal can be performed using an analytical method. The discrete-position signal obtained from the sensor array is considered to be samples of a continuous-position signal, that has been sampled respecting the Nyquist criterion. The continuous-position signal determined can then be uniquely resampled (at any desired rate) appropriately to ensure that signals obtained from different runs are invariant with respect to spatial

shift. One method of accomplishing this is to use an interpolation-decimation scheme. Interpolation and decimation are signal processing methods in which the sampling rate can be increased or decreased by a specified amount, as described below [63].

5.2.1.1 Decimation Consider a discrete time signal $x[n]$ that is obtained from sampling a continuous time signal $x_c(t)$. This operation can be written as

$$x[n] = x_c(nT) \quad (5.42)$$

where T is the sampling period and its reciprocal, $f_s = 1/T$, is the sampling frequency. The sampling rate can be reduced by a factor M by defining a new sequence

$$x_d[n] = x[nM] = x_c(nMT) \quad (5.43)$$

Such a process is called downsampling. A Fourier transform analysis yields the relationship in the frequency domain between the original signal, $x[n]$ and the resampled version $x_d[n]$; which can be written as

$$X_d(e^{j\omega}) = \frac{1}{M} \sum_{i=0}^{M-1} X\left(e^{j(\omega/M - 2\pi i/M)}\right) \quad (5.44)$$

The above equation shows that the Fourier transform $X_d(e^{j\omega})$ is given by M copies of the periodic Fourier transform $X(e^{j\omega})$, frequency scaled by M and shifted by integer multiples of 2π . In order to avoid aliasing, $X(e^{j\omega})$ must be bandlimited:

$$X(e^{j\omega}) = 0, \quad \omega_N \leq |\omega| \leq \pi \quad (5.45)$$

and $2\pi/M \geq 2\omega_N$. This condition can be ensured by pre-filtering the signal $x[n]$ with a low-pass filter with a cut-off frequency π/M before downsampling. The combination of a low-pass filter and a downsampler is called a decimator.

5.2.1.2 Interpolation The sampling rate can be increased by a factor L , by defining a sequence of the form

$$x_i[n] = x[n/L] = x_c(nT/L), \quad n = 0, \pm L, \pm 2L, \dots \quad (5.46)$$

This process is referred to as upsampling. $x_i[n]$ can be obtained from $x[n]$ in two steps. First, $x[n]$ is passed through a system called an expander whose output is

$$x_e[n] = \begin{cases} x[n/L], & n = 0, \pm L, \pm 2L, \dots \\ 0, & \text{otherwise} \end{cases} \quad (5.47)$$

The Fourier transforms of $x_e[n]$ and $x[n]$ are related by the equation

$$X_e(e^{j\omega}) = X(e^{j\omega L}) \quad (5.48)$$

Thus the Fourier transform of $x_e[n]$ is a frequency scaled version of the Fourier transform of $x[n]$.

In the second step, the signal $x_e[n]$ is passed through a low-pass filter with a gain of L and cut-off frequency of π/L . This process constructs the Fourier transform of $x_i[n]$, $X_i(e^{j\omega})$.

In the current application, the MFL signal is first passed through an interpolator to reconstruct the “underlying continuous-position” signal. By this process, all MFL signals originating from the same defect are rendered similar, regardless of the origin of sampling. The resulting signal is resampled with a convenient sampling period using a decimator. Figure 5.5 shows such an interpolator-decimator cascade.

5.2.2 Stochastic method: Signal registration

The necessity for signal registration can be shown by observing MFL signals obtained at two velocities which are shifted with respect to each other in the axial

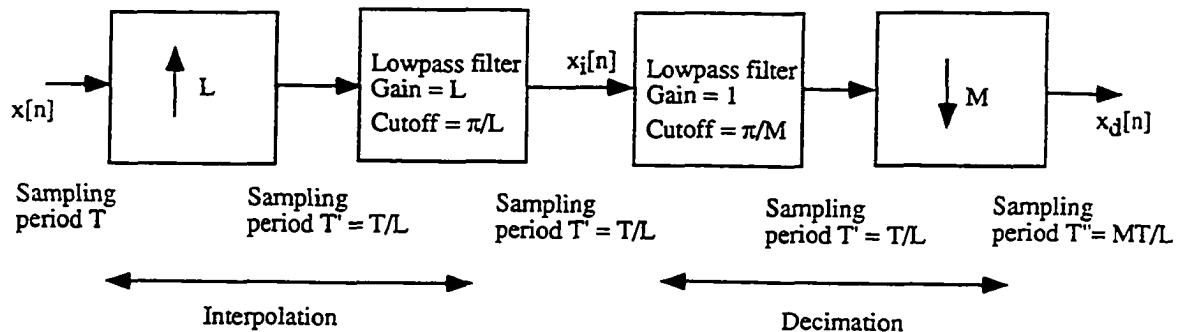


Figure 5.5: Interpolator-decimator cascade to compensate for variations in the origin of sampling

and circumferential directions. Figure 5.6 shows axial flux density signals (1-D scans) for a defect obtained at tool velocities of 2 mph and 7 mph; both perfectly aligned ($\text{shift} = 0$) and with a relative shift of -50 units. Figure 5.7 shows a 2-D axial-circumferential scan of a defect, obtained at 2 mph. Figure 5.8 shows the perfectly aligned signal for the same defect obtained at 7 mph and Figure 5.9, the same defect signature misaligned in the axial and circumferential directions.

A simple automated signal/image registration scheme is therefore included as part of the filter design process. The cross-correlation vector/matrix, p , between the low and high velocity signals/images is employed for this purpose. The position of the peak value of the cross-correlation sequence indicates the measure of misalignment. If the peak occurs at the exact center of the vector/matrix, then the two signals/images are perfectly aligned. The displacement in the peak of the cross-correlation from the

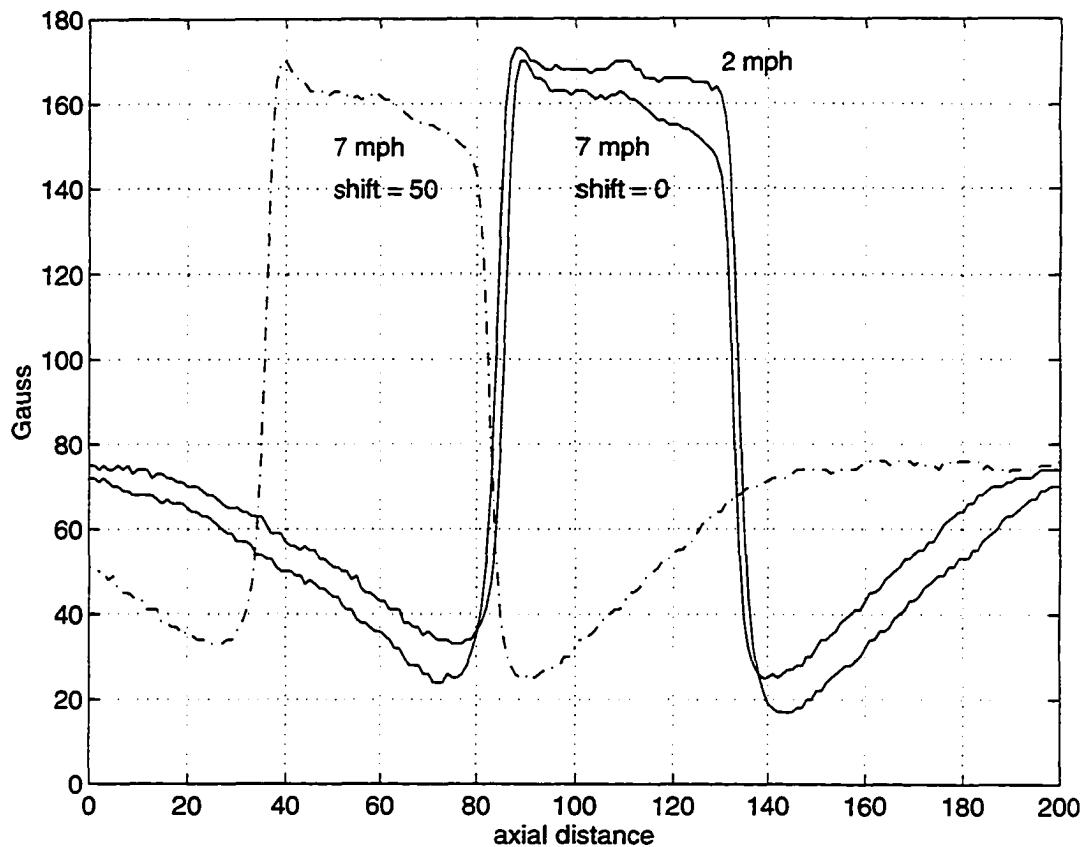


Figure 5.6: Misalignment in 1-D axial scans

MFL signal at 2 mph

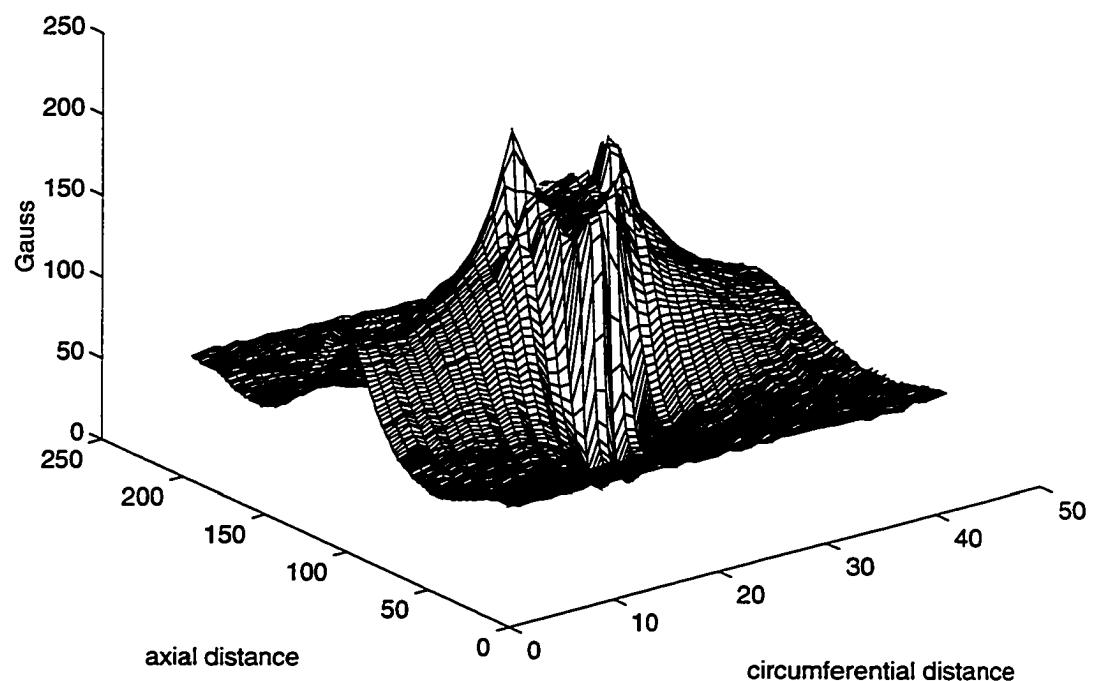


Figure 5.7: 3-D MFL signal obtained at 2 mph

Perfectly aligned MFL signal at 7 mph

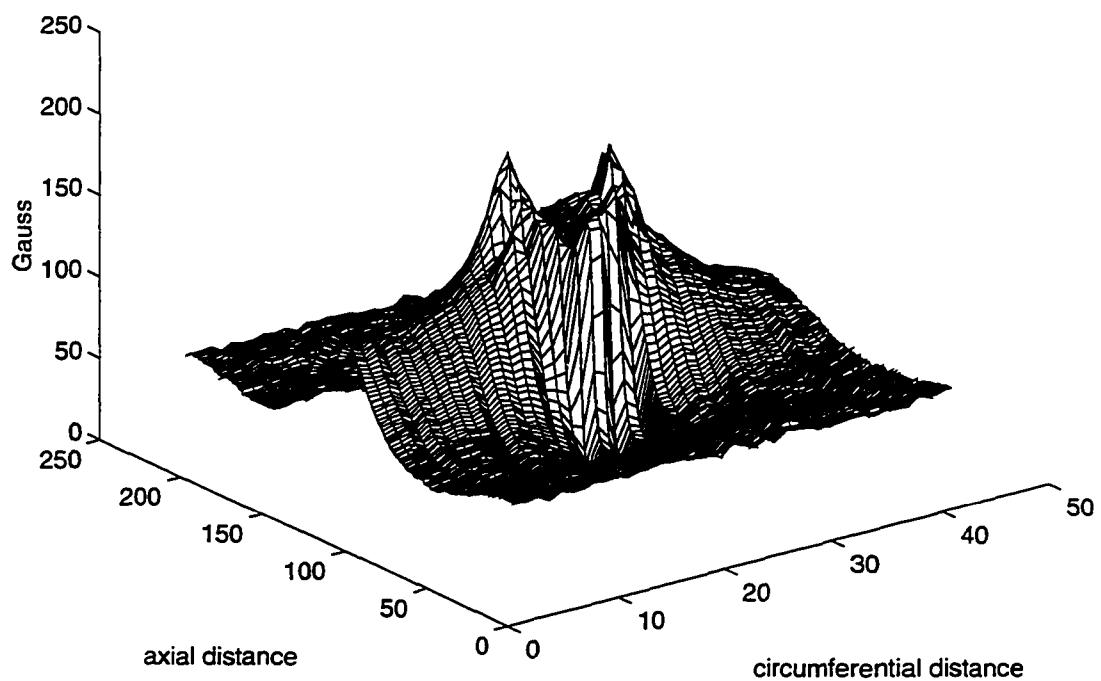


Figure 5.8: Perfectly aligned 3-D MFL signal, obtained at 7 mph

Signal at 7 mph displaced by +50 axial and -5 circumferential

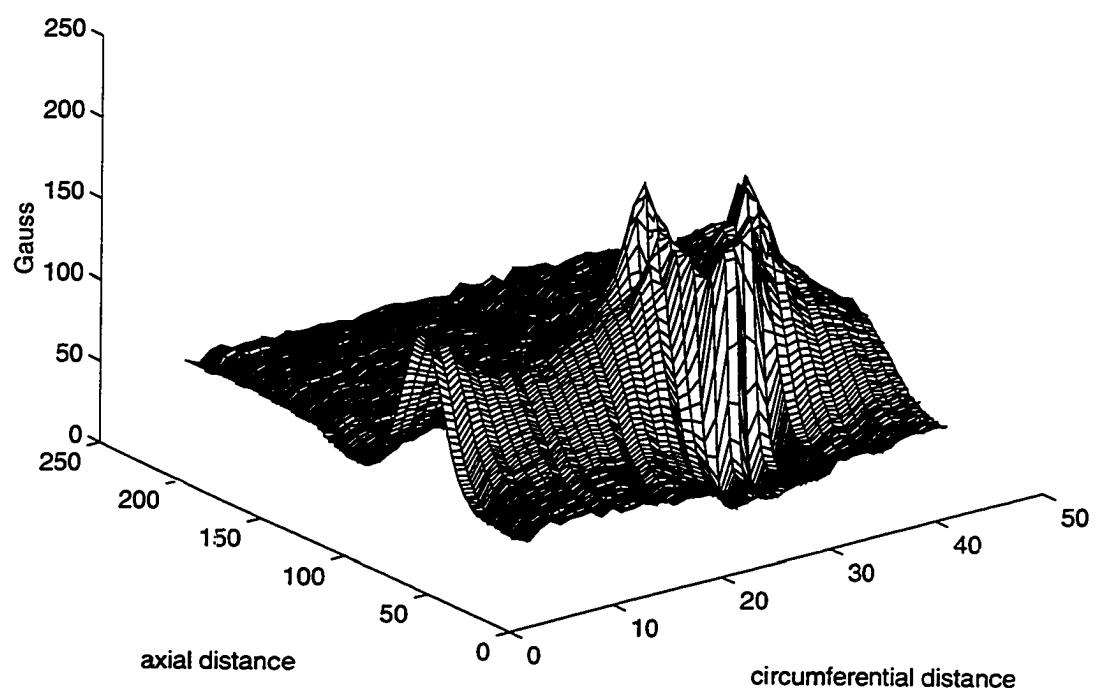


Figure 5.9: Misaligned 3-D signal, obtained at 7 mph

center of the sequence provides the translation parameters required for registration. The underlying principle of this scheme is the fact that the MFL indication at both velocities occurs in regions surrounding the defect. Since the cross-correlation vector is computed as part of the filter development process, this technique does not involve any additional computational burden.

5.3 3-D Velocity Invariance Scheme

The velocity invariance techniques described in the previous sections can be applied to deconvolve the effects of tool velocity for MFL signals obtained from axial scans of a single sensor. The restoration filter coefficients are computed by studying the statistical properties of an ensemble of static and velocity affected MFL signals, obtained from a variety of defect geometries. A natural extension of this technique can be used to perform 3-D velocity effects compensation for signals arising from multiple sensors that scan the pipe. The MFL signal that is derived from each sensor along the circumference of the pipe is fed to a circumferential array of digital filters, one for each sensor, for pre-processing prior to defect characterization. This arrangement is shown in Figure 5.10. The kernel of the digital filter connected to each sensor channel is obtained as described in earlier sections, either using fixed or adaptive algorithms.

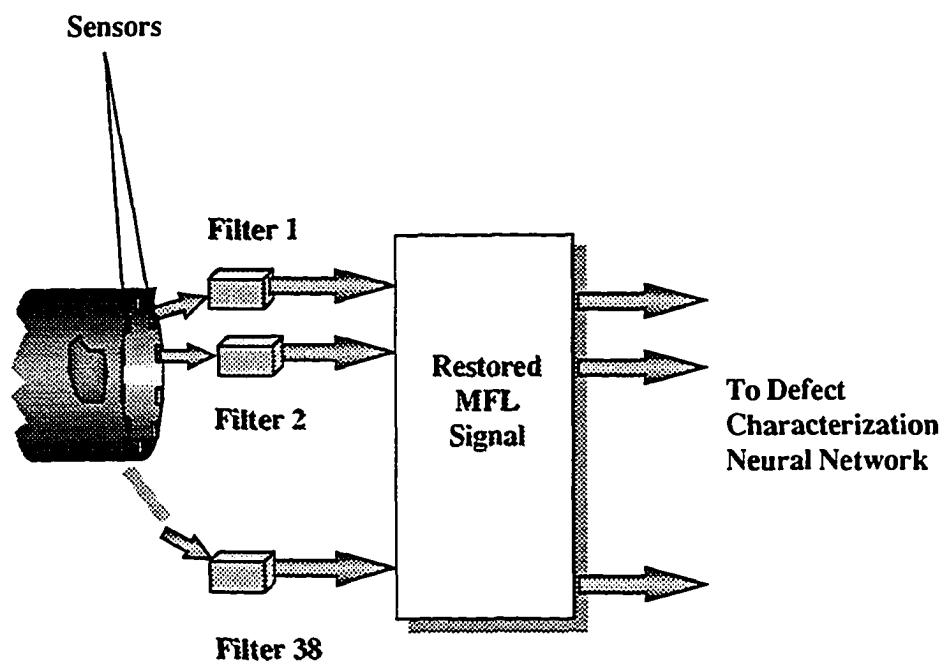


Figure 5.10: 3-D velocity effects compensation scheme

CHAPTER 6. RESULTS AND DISCUSSION

The results of implementing the novel invariance transformation algorithms that have been described in the previous chapters are presented in the following sections.

6.1 Permeability Invariance

As described in Chapter 4, the permeability invariance scheme is based on techniques for data interpolation. The results of the interpolation methods derived from phenomenological models and from simple non-phenomenological models have already been shown in Chapter 4. In this section, results of the permeability invariance scheme based on universal approximation capabilities of feed-forward neural networks are presented.

6.1.1 Radial basis function networks

Figure 6.1 summarizes the results of the permeability invariance scheme using the radial basis warping functions with Gaussian kernel functions. Figure 6.1 (a) shows the axial flux density signals prior to the invariance transformation and Figure 6.1 (b) shows the resulting permeability invariant signals. The defect data set consists of rectangular and trapezoidal defects of depths varying from 20% – 80% of wall thickness and lengths 2 inches – 4 inches. The invariance procedure described

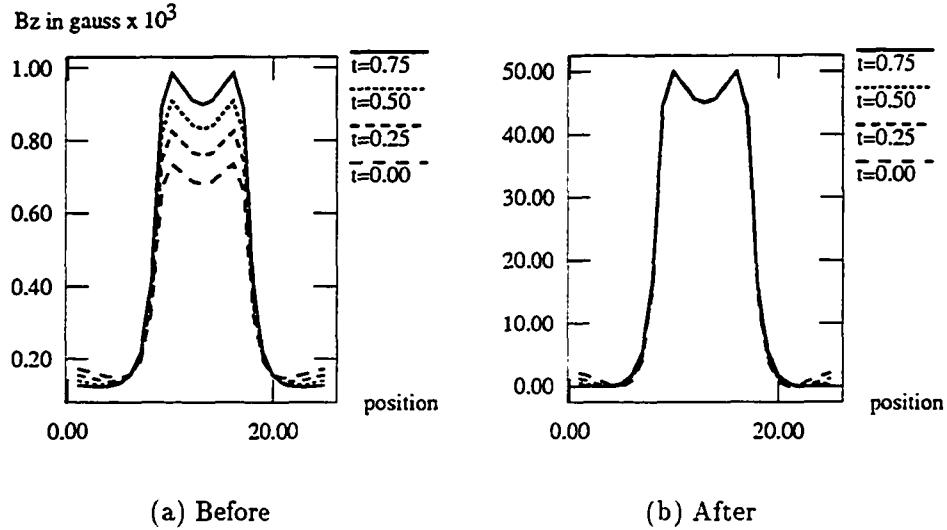


Figure 6.1: Permeability invariant signals derived using RBF network

in the flow chart of Figure 4.2 is implemented using the axial and radial components of the magnetic flux density as x_A and x_B . The Figure of Merit, as defined in Chapter 4 is of the order of 10^7 . The discrimination between defect classes can be user-specified.

The permeability invariance scheme was also implemented on the axial and circumferential components of the MFL signal. The circumferential component of the MFL signal is a better indicator of the defect width. Unlike the axial leakage flux density, the circumferential signal is not subject to “blooming” in the presence of multiple defects. Figure 6.2 (a) shows the variation in the magnetization characteristics of the pipe-wall, for which corresponding axial and circumferential components of the MFL signal are shown in Figures 6.2 (b) and (c) respectively. These signals have been derived from 3-D finite element models. The permeability invariant axial

flux density signals are seen in Figure Figure 6.2 (d).

6.1.1.1 Permeability invariance using a single component of the MFL signal Gas pipeline inspection systems currently in use retrieve only one component, namely, the axial component of the MFL signal. The permeability invariance scheme described earlier calls for the use of both axial and radial components on the premise that the two components respond differently to changes in magnetization characteristic. A practical implementation of these schemes require additional (twice the number) sensors and memory in the data acquisition system. An invariance scheme that relies on only one component of the flux density would significantly reduce the data acquisition burden. An alternate strategy that can be pursued is to identify a function f_1 such that $P\{f_1(B_z)\} \neq P\{B_z\}$, or, a function f_2 such that $P\{f_2(B_r)\} \neq P\{B_r\}$, where P represents an operator that provides the peak value. Then, $P\{f_1(B_z)\}$ can be used in place of P_r , or $P\{f_2(B_r)\}$ can be used in place of P_z . Suitable candidates for f_1 and f_2 include spatial derivatives and integrals, square root and other non-linear operators.

An application of the scheme incorporating only one component of the flux density (radial) and employing an integrating preprocessor is shown in Figure 6.3. Figures 6.3 (a) and (b) respectively show the variation of the radial flux density (B_r) and its integral ($\text{Int}\{B_r\}$) with respect to the sensor position, for a 20% deep, 2 inch long defect for various pipe-wall B-H curves (t). Figures 6.3 (c) and (d) show the variation of the peak-peak values of (B_r) and ($\text{Int}\{B_r\}$), respectively, for variation in both B-H curve ($0 \leq t \leq 1$) and defect depth ($20\% \leq d \leq 80\%$). The permeability invariant feature, h , obtained after applying the invariance transformation is seen in

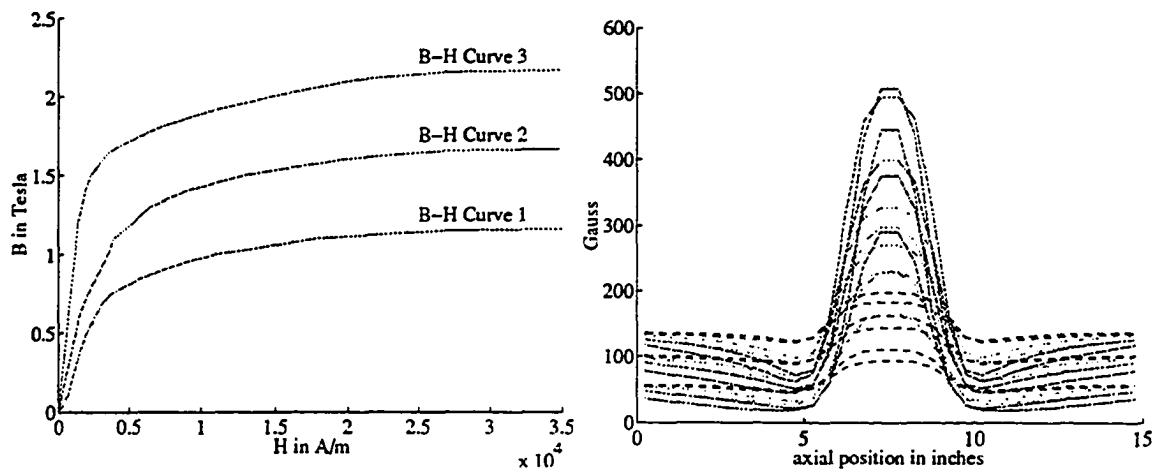
Figure 6.3 (e).

6.1.2 Fuzzy inference systems

Fuzzy inference systems were investigated as tools for universal approximation in order to improve the performance of the permeability invariance scheme for data sets that are not part of the training samples. Figure 6.4 shows the data vectors, $[P_r, P_z, D_r]'$, that are input to the fuzzy inference system that performs the permeability invariance transformation, and the cluster centers of the fuzzy membership functions determined using the gradient descent technique described in Chapter 4. The above procedure generates the fuzzy inference rules and is equivalent to the training of a neural network. Results for data interpolation amongst the magnetization characteristic or t -parameter using Gaussian membership functions of Equation (4.16) are shown in Figure 6.5. The input MFL signals to the fuzzy inference system, shown in Figure 6.5 (a), do not form part of the training or rule determination data set. The corresponding permeability invariant signals for two different defect dimensions and varying pipe-wall permeability are seen in Figure 6.5 (b). The use of triangular membership functions yields almost identical results.

6.1.3 Wavelet basis function networks

Fuzzy inference systems proved incapable of performing accurate invariance transformations over MFL data from defect depths that were not part of the training data set. Examination of the input-output mapping required of the invariance scheme, shown in Figure 4.10, leads to the conclusion that the parametric variations for defect depth and pipe-wall permeability occur at different resolutions. Multires-



(a) Pipe-wall magnetization characteristics (b) Axial component of MFL signal

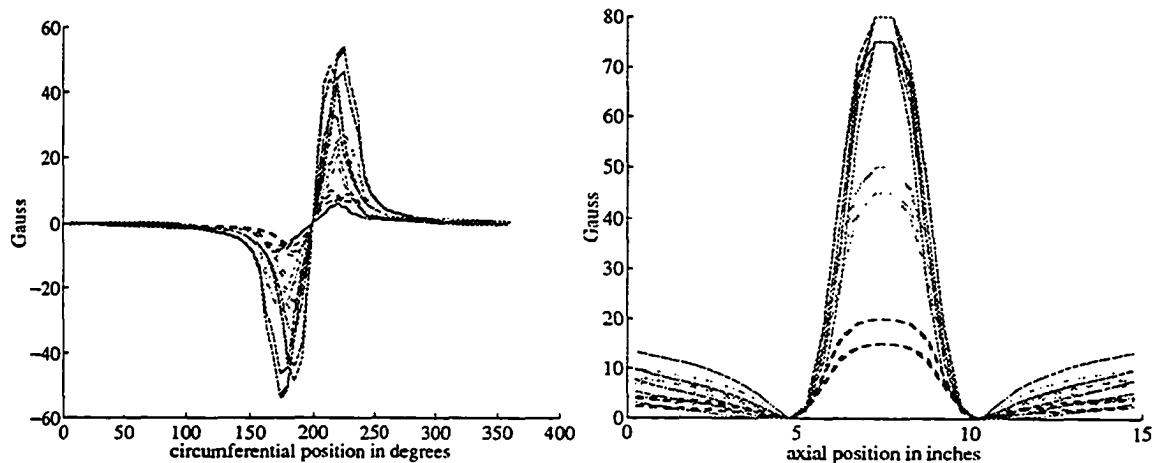


Figure 6.2: Permeability invariance using the axial and circumferential components of the MFL signal. - - - 20%, - - - 50%, — 80%

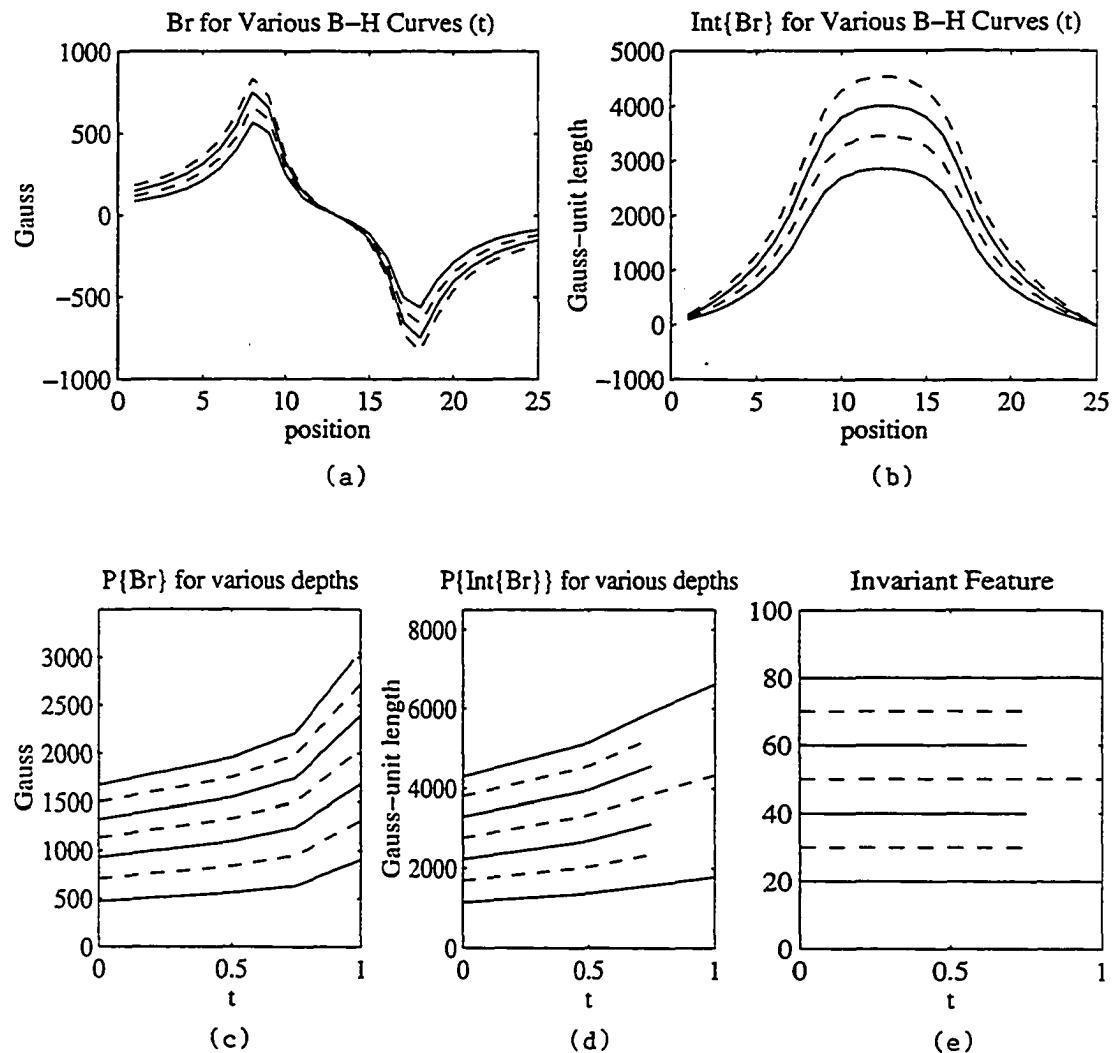


Figure 6.3: Permeability invariance scheme using only the radial component

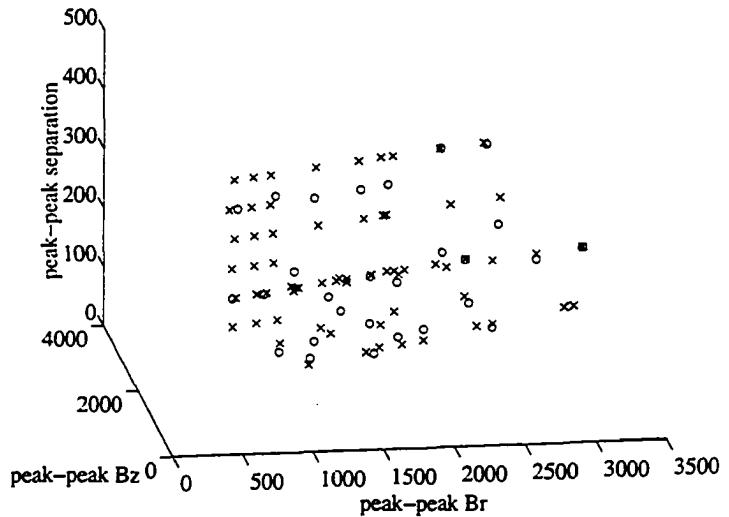


Figure 6.4: Data vectors and cluster centers determined using gradient descent in FIS: \times = data points, \circ = cluster centers

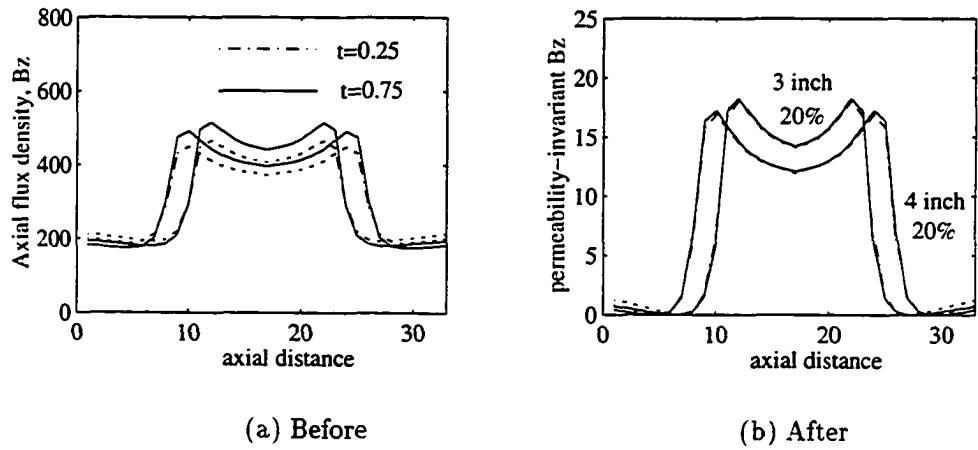


Figure 6.5: Interpolation amongst magnetization characteristic for FIS

solution interpolation techniques such as WaveNets appear to be ideal candidates for performing this sort of mapping. The WaveNet consists of two sets of nodes — the ϕ -nodes, which model the transformation function for variations in depth and the ψ -nodes, which model the function for variations in permeability.

Gaussian scaling functions and wavelets, that are shown in Figure 4.13, were used to construct the wavelet based permeability invariance scheme. The resulting input-output mapping of the WaveNet for training and testing data for different defect depths and B-H curves is shown in Figure 6.6. The network was trained only with MFL signals from 3 inch long, 20% and 80% deep defects from varying pipe-wall B-H curves (t). As explained earlier, network training is a two-stage procedure. The network contained two ϕ -nodes, with the basis function centers situated at the input features corresponding to $d = 20\%$, $t = 0.5$; and $d = 80\%$, $t = 0.5$. The weights of the network were determined using only these two nodes. The approximation error was noted. In the next stage, the network was augmented with six ψ -nodes, three corresponding to each ϕ -node. The centers for each of these nodes were derived from the input feature vectors denoting the following parametric variations — $d = 20\% : t = 0.0, 0.5, 1.0$; $d = 80\% : t = 0.0, 0.5, 1.0$. The network was retrained to reduce the approximation error determined in the first stage. For both stages, the support of the Gaussian scaling and wavelet functions were determined heuristically.

The resulting depth interpolation performance of the network is shown in Figure 6.7. Figure 6.7 (a) shows input data to the system. Recall that the network was trained with MFL signals from 20% and 80% deep defects only; the network was tested with signals from 20%, 50% and 80% deep defects. The corresponding permeability invariant signals are shown in Figure 6.7 (b).

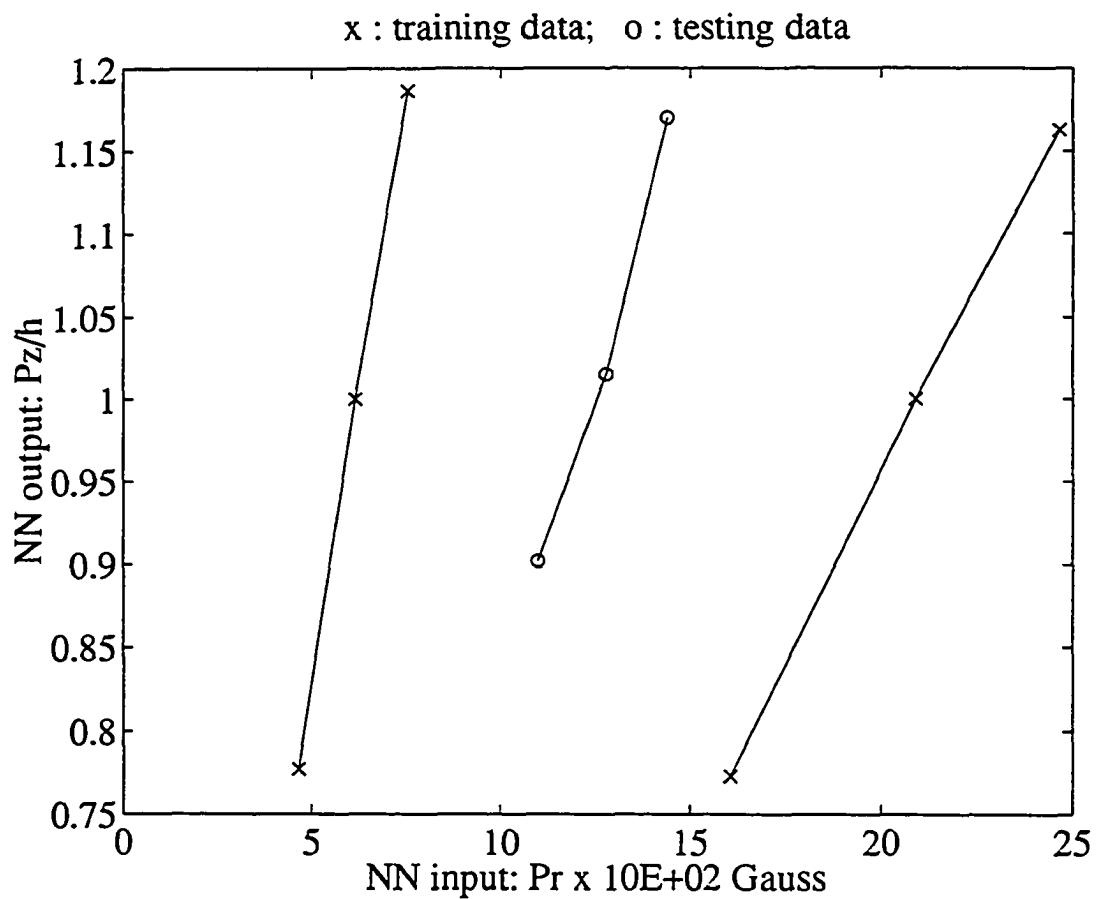


Figure 6.6: Input-output mapping of the WaveNet for training and testing data

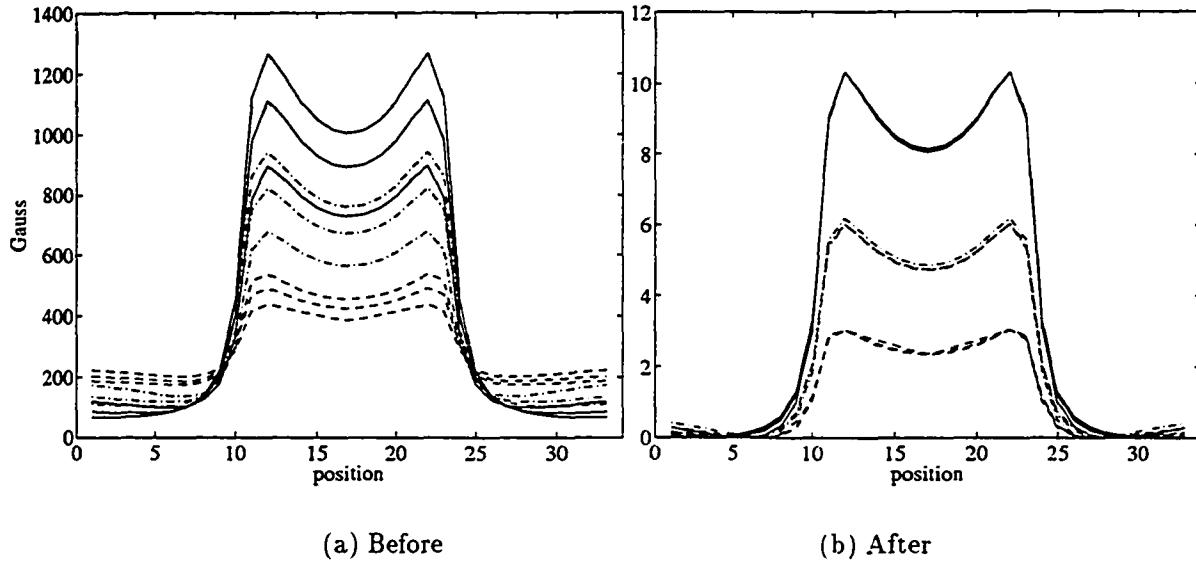


Figure 6.7: Depth interpolation results using WaveNet. - - - 20%, - - - 50%, — 80%

6.1.4 3-D permeability invariance

Finally, the 3-D permeability invariance scheme was exercised with MFL signals obtained from finite element simulation. The axial, radial and circumferential components of the flux density were obtained for the pipe-wall B-H curves shown in Figure 6.2 (a). from which the input features to the transformation network, namely, $[P_z, P_r, D_r, D_c]'$ were computed. Typical 3-D axial flux density signals for variation in the pipe-wall B-H curves are shown in Figure 6.8. The results obtained after the invariance transformation is presented in Figure 6.9.

6.2 Velocity Invariance

Results obtained using the semi-empirical velocity compensation filter were shown in Chapter 5. In this section, results demonstrating the successful application of de-

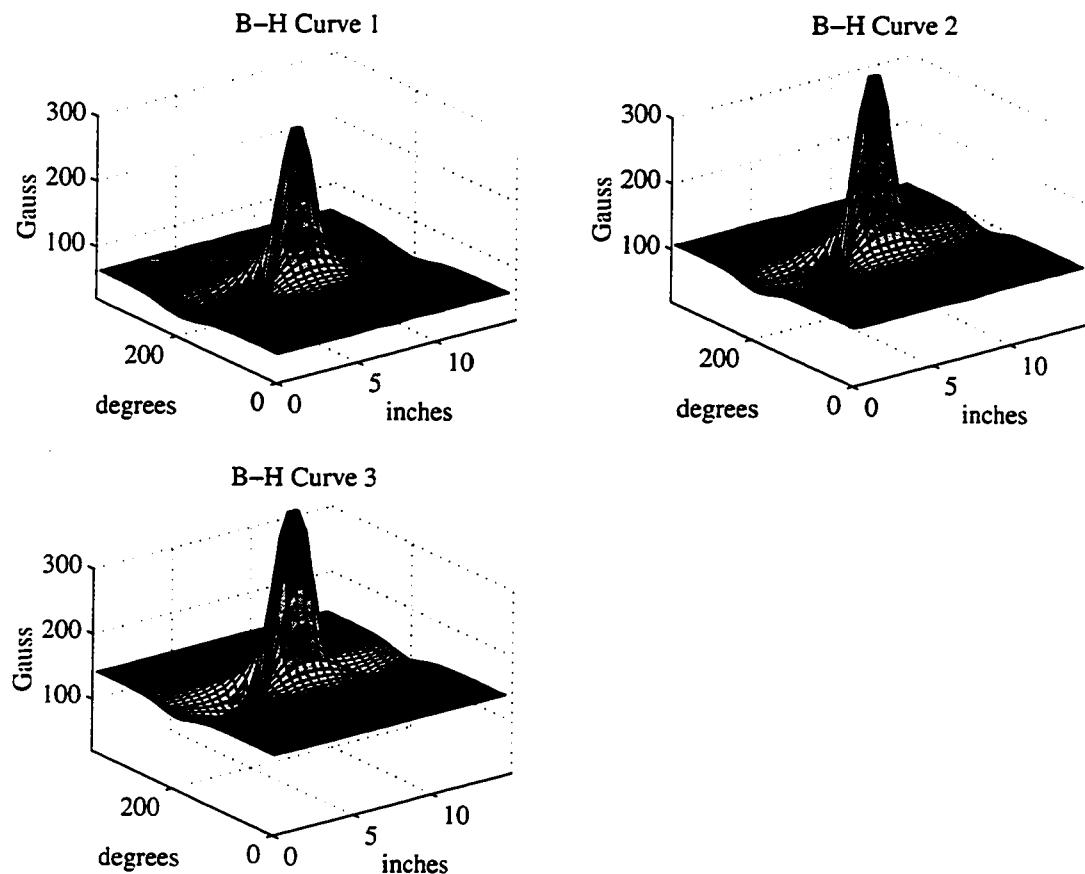


Figure 6.8: Typical input MFL signals for the 3-D permeability invariance scheme

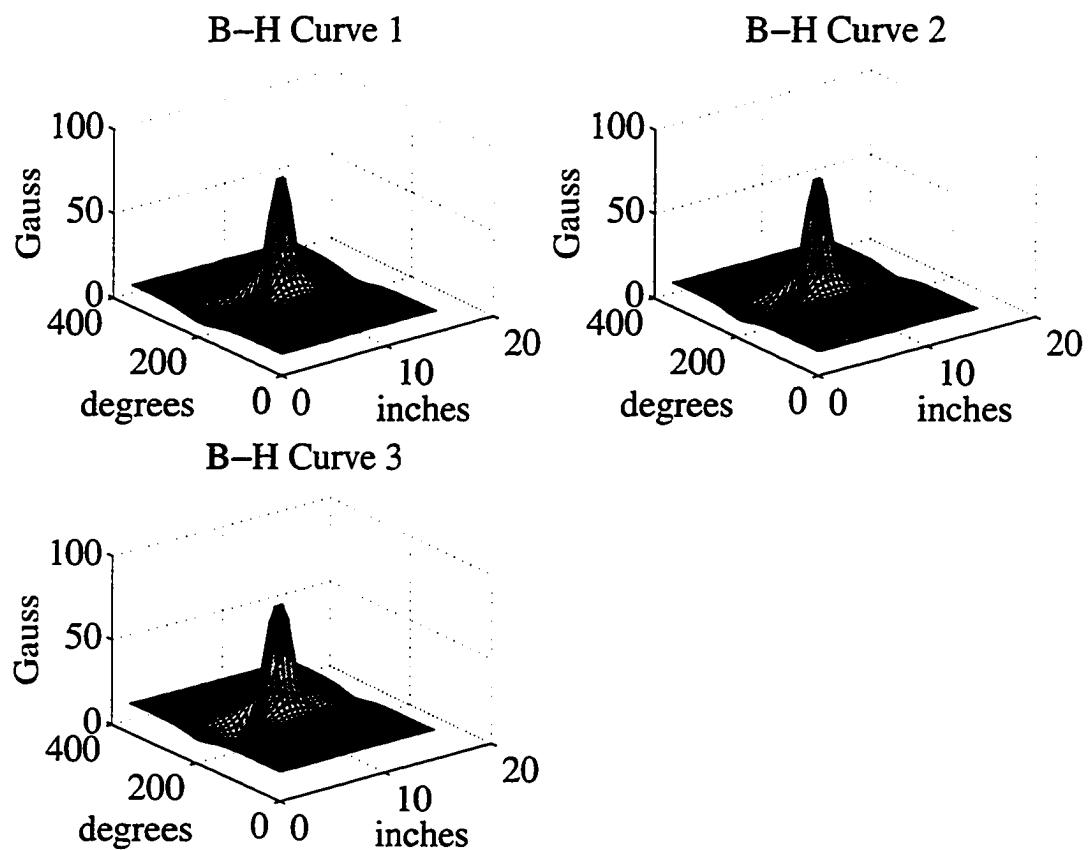


Figure 6.9: 3-D permeability invariant MFL signals

terministic and various types of stochastic filters, are presented.

6.2.1 Deterministic filter

Deterministic velocity compensation filters were derived using experimental signals¹. Figure 6.10 shows the distorted, static and restored MFL signals for a variety of rectangular defect profiles. In this experiment, the pig is moved at a velocity of 8 mph. The defect width refers to the length of the defect in the circumferential direction. In Figure 6.10, the distorted signals, after being passed through the velocity compensation filter are postprocessed by smoothing with a second order Butterworth filter [63], to remove any ripple that is introduced as an artifact of the invariance scheme.

6.2.2 Stochastic filters: Linear and quadratic

In order to compute the weights of the linear stochastic filter, the auto-correlation matrix and the cross-correlation vector of the defect signal were first determined. The MFL process was assumed to be ergodic and ensemble averages were replaced by time (position) averages. Figure 6.11 shows the distorted, restored and static signals from four defects using a linear restoration filter with fixed weights determined by Equation (5.24). A plot of the second-order term for a quadratic Volterra filter, the weights computed from Equation (5.39) is seen in Figure 6.12. A comparison of the magnitudes of the linear and non-linear components shows that the contribution from the second-order terms is minimal. A linear filter is therefore adequate, even though the velocity distortion process is non-linear. The following reasons could account for this behavior:

¹Experimental signals were provided by Battelle, Columbus, OH.

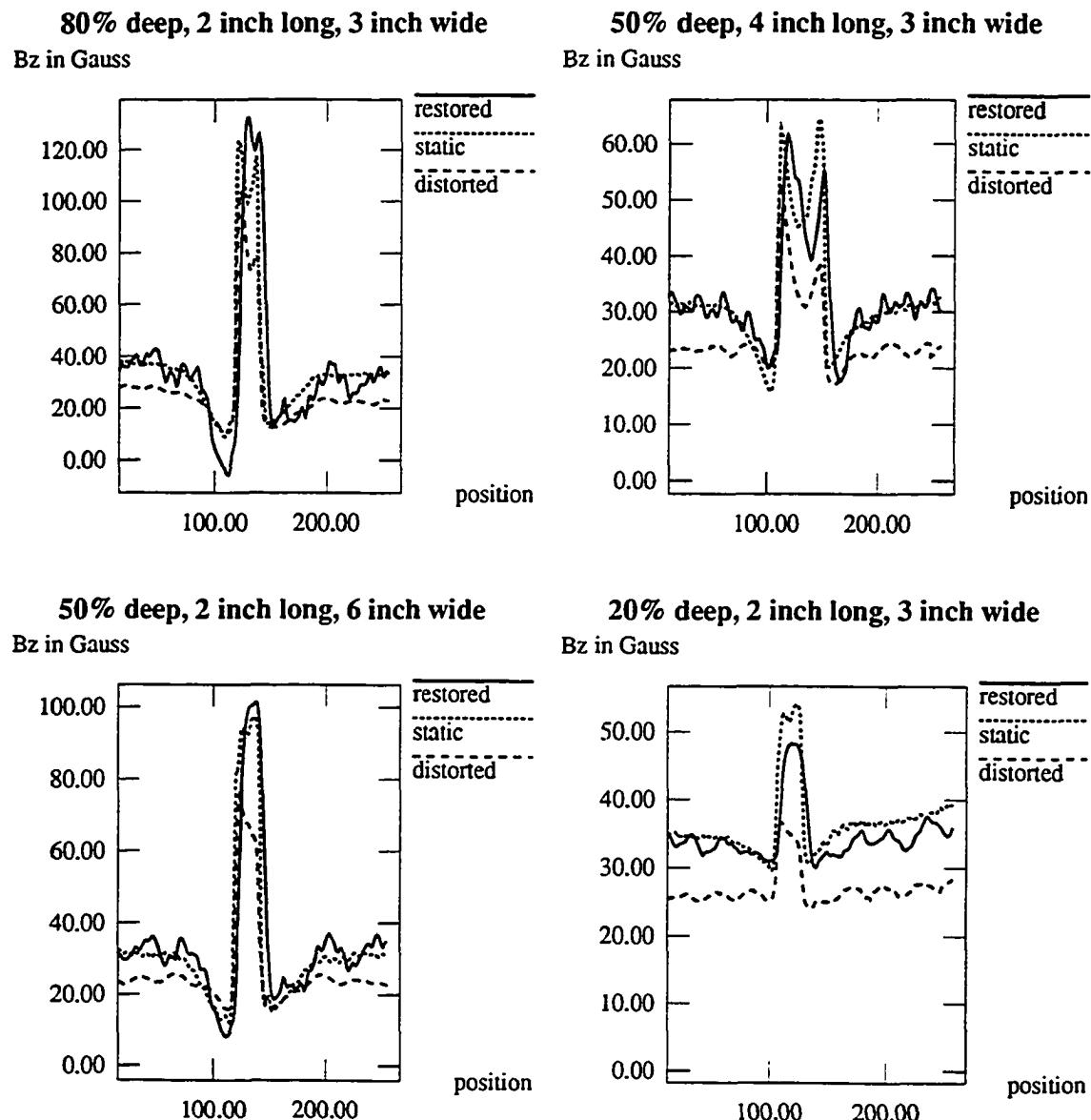


Figure 6.10: Velocity effects compensation using deterministic filter

1. A limited defect signature set was used for determining the filter coefficients. Consequently, second order effects may not be significant.
2. The equation governing the MFL process contains two velocity dependent terms, namely, $\mathbf{v} \times (\nabla \times \mathbf{A})$ and $\sigma \frac{\partial \mathbf{A}}{\partial t}$ (see Equation 2.1). These terms exert a prominent influence on the MFL signal at high tool velocities. If it is assumed that the velocity effects described by these terms occur in parallel and the filters that compensate for them are additive, it is possible that their effects could cancel each other.

Further investigation to test these hypotheses is still in progress. Since the calculation of the cross-bicorrelation matrix for constructing a quadratic filter is very time-consuming, at present, velocity compensation schemes are restricted to linear filters.

6.2.3 Adaptive filters

Adaptive algorithms estimate the optimal filter weights by iteratively minimizing a suitable cost function (usually the mean squared error). For the algorithm to converge, it is required that a well defined minimum exist for the error surface. The eigenvalue spread of the auto-correlation matrix \mathbf{R} is a measure of the correlation of the input. The larger the eigenvalue spread, more correlated is the input and the minimum mean squared error tends towards zero. The error surface then contains a sharp minimum. Figure 6.13 is a plot of the eigenvalues of the auto-correlation matrix. It can be seen that the eigenvalue spread is very large, of the order of 2.05×10^5 . This indicates that adaptive algorithms can converge to the optimal

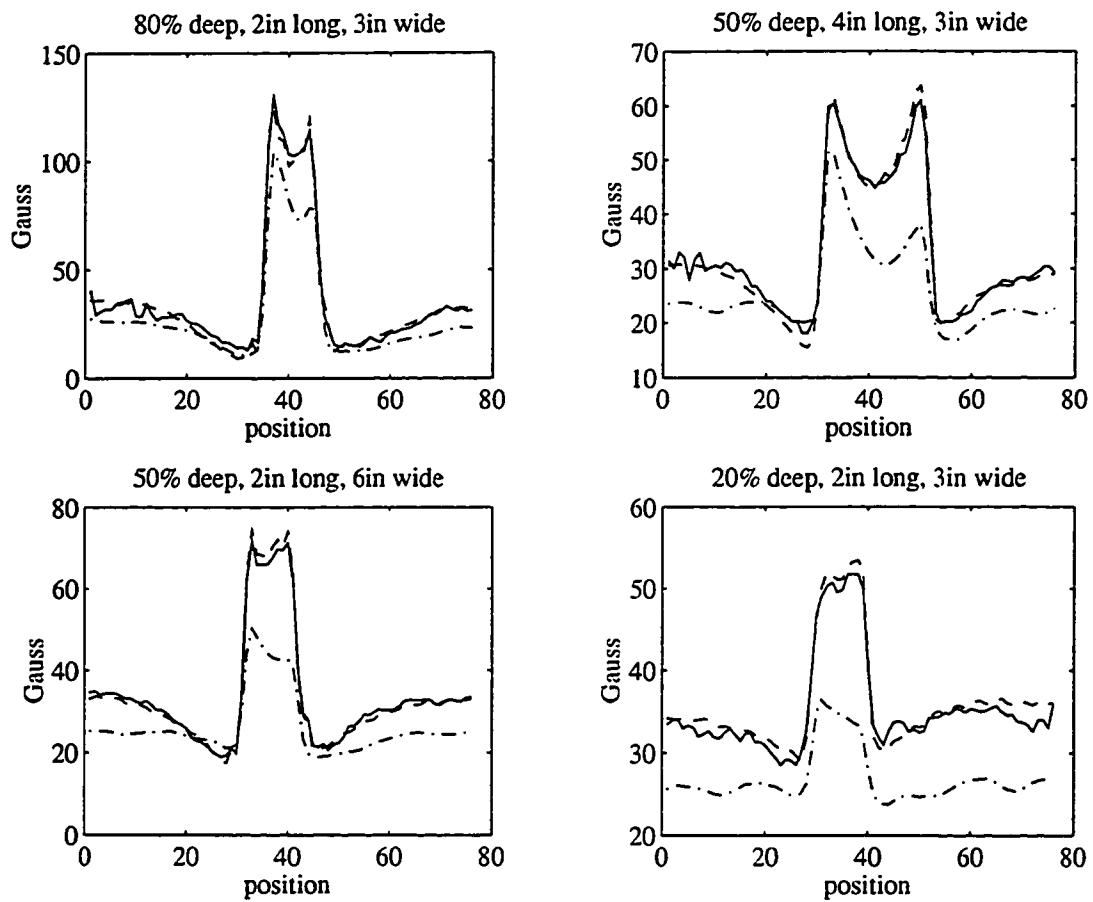


Figure 6.11: Velocity effects compensation using fixed linear filter. — restored, - - desired, -·- distorted

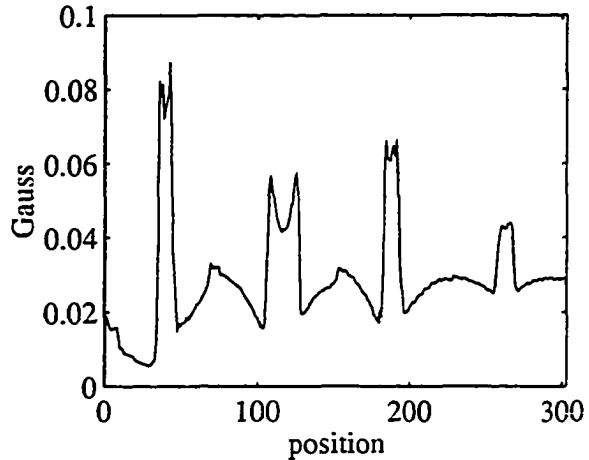


Figure 6.12: Second order terms of the quadratic Volterra filter

weights with minimum error. This also validates the filter approach for deconvolving the effects of probe velocity in the MFL signal.

The stochastic gradient descent algorithm was used to determine the filter weights for four defect signatures. The filter weights so obtained were convolved with a fifth defect signature to test the interpolation performance. Figure 6.14 shows the variation of the mean squared error with iteration for training and testing data. The performance of the velocity compensation filter for the training data can be seen in Figure 6.15 and for the testing data in Figure 6.16.

6.2.4 Preprocessing MFL signals prior to velocity effects compensation

6.2.4.1 Invariance to origin of sampling The results of the pre-processing scheme to render MFL signals invariant to the sampling origin are summarized in Figure 6.17. Two axial flux density signals, that were obtained by scanning the pipe-wall at different points along the same defect, are shown in Figure 6.17 (a).

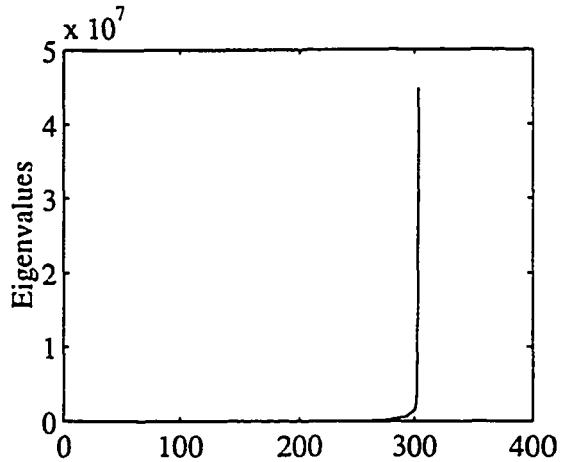


Figure 6.13: Eigenvalues of the auto-correlation matrix

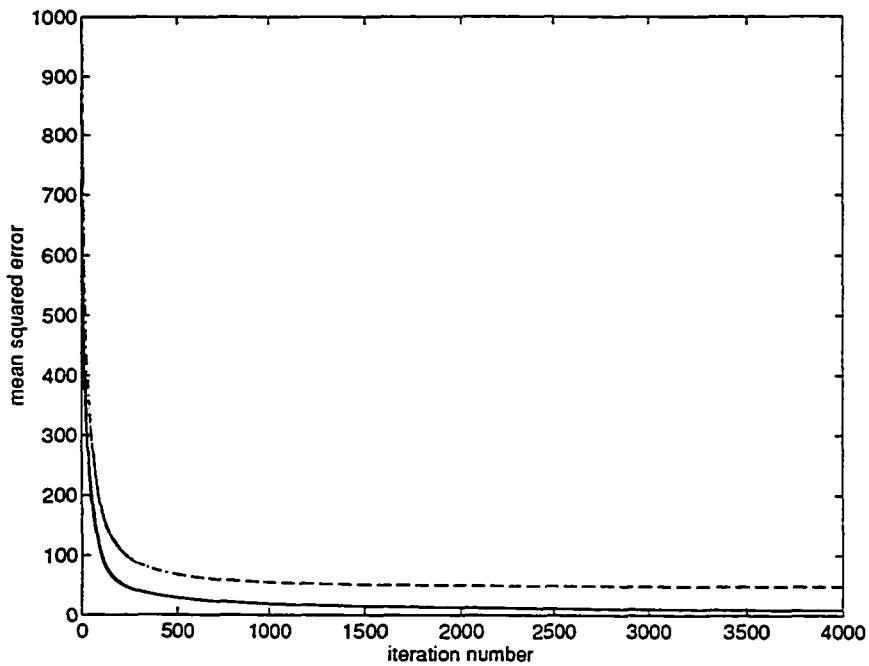


Figure 6.14: Mean squared error with iteration.— training data, - - testing data

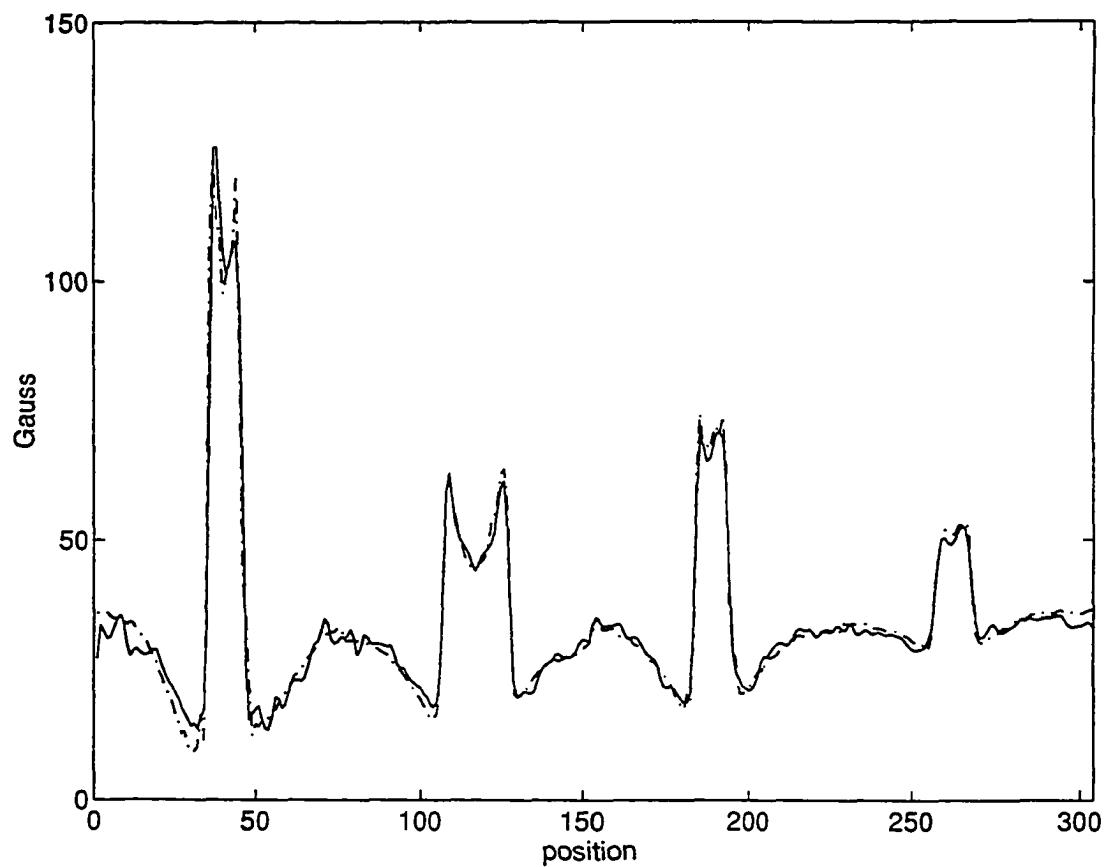


Figure 6.15: Velocity effects compensation using gradient descent algorithm: training data — restored, -·- desired

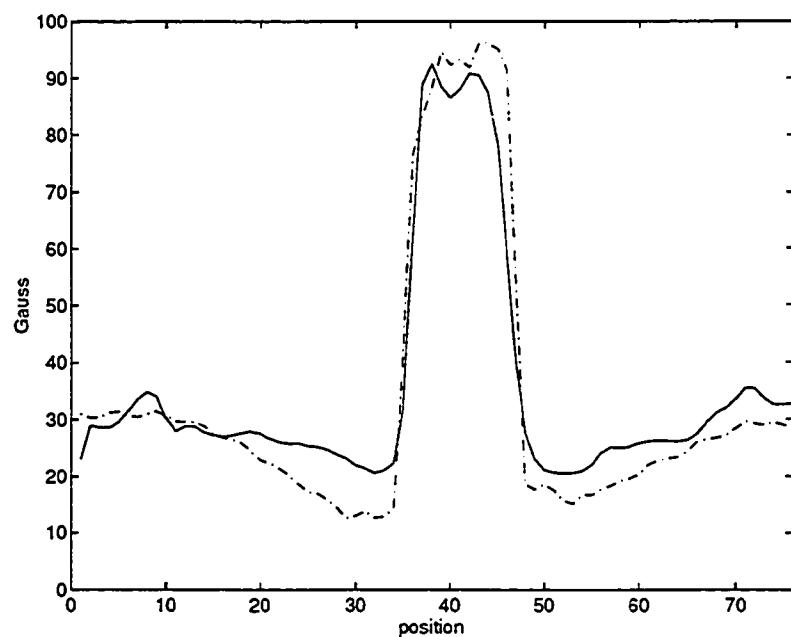


Figure 6.16: Velocity effects compensation using gradient descent algorithm: testing data — restored, -.- desired

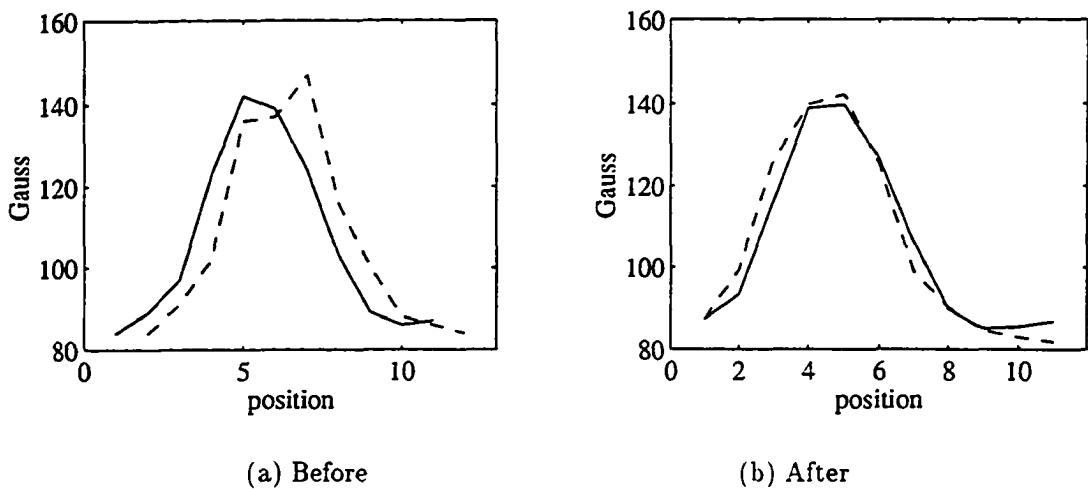


Figure 6.17: Results of the spatial shift invariance scheme

These signals were input to the interpolator-decimator cascade described in Figure 5.5. MFL signals that are invariant to the sampling origin are derived at the output of the cascade. These are shown in Figure 6.17 (b).

6.2.4.2 Signal registration The cross-correlator image registration algorithms described earlier were exercised for 2-D and 3-D MFL signals. Figure 6.18 shows the cross-correlation vector for the perfectly aligned and misaligned signals shown in Figure 5.6. It is seen that the peak of the cross-correlation vector exhibits a shift identical to the relative displacement between the two signals. Similarly, the cross-correlation matrix for perfectly aligned and misaligned 3-D signals shown in Figures 5.8 and 5.9 can be seen in Figures 6.19 and 6.20 respectively.

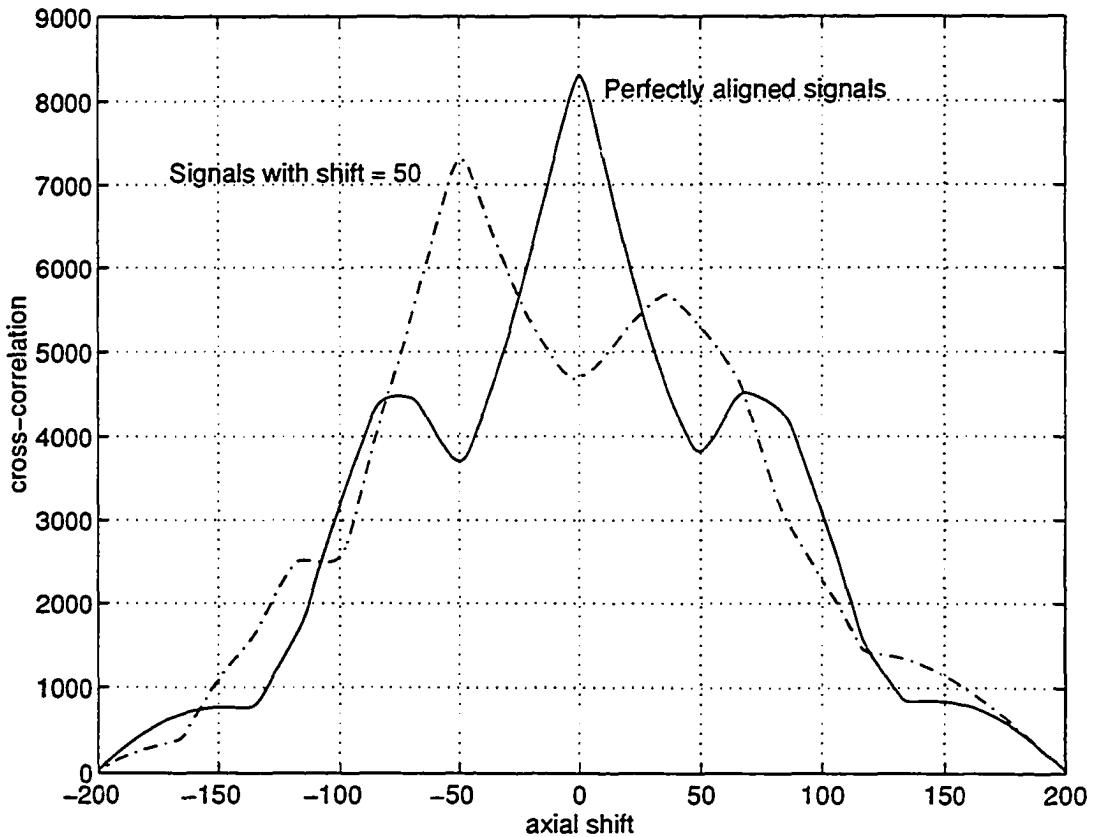


Figure 6.18: Cross-correlation vectors for perfectly aligned and misaligned 1-D MFL scans

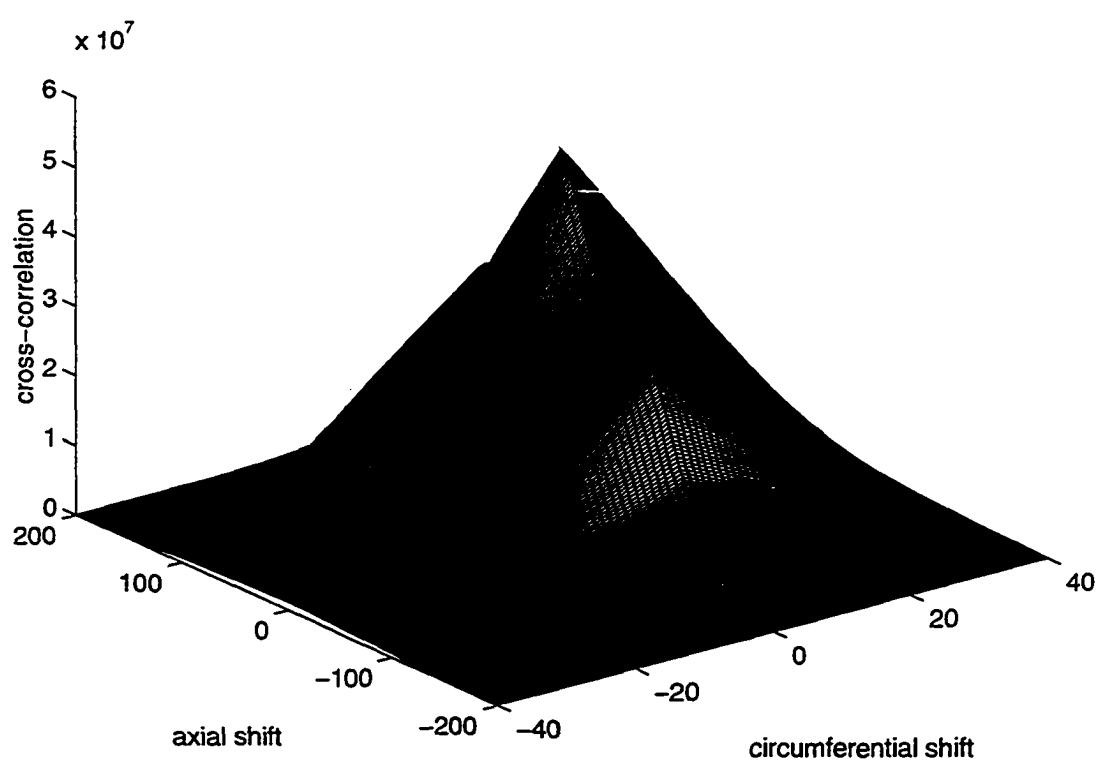


Figure 6.19: Cross-correlation matrix for perfectly aligned 2-D MFL scans

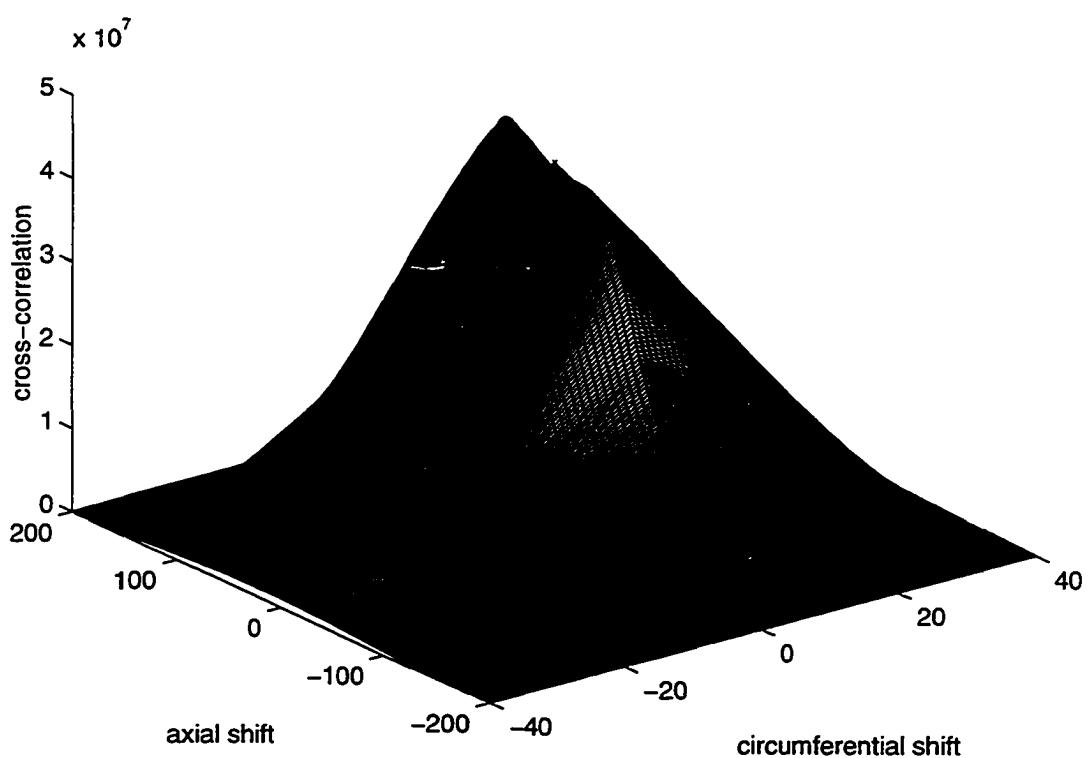


Figure 6.20: Cross-correlation matrix for misaligned 2-D MFL scans

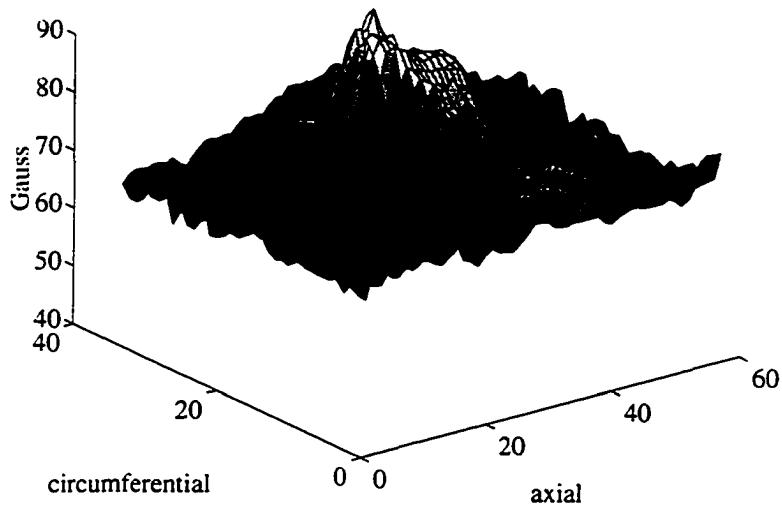


Figure 6.21: Typical input signal to the 3-D velocity effects compensation scheme

6.2.5 3-D velocity effects compensation

The 3-D velocity effects compensation scheme, as shown in Figure 5.10, was developed and tested using static (2 mph) and velocity affected (7 mph) MFL signals, obtained by experimental methods. The individual filters comprising the circumferential filter array consisted of linear filters with the coefficients computed using a fixed algorithm. Before determining the filter coefficients, the low and high velocity signals were registered using the technique described in the previous chapter. Figure 6.21 shows a typical velocity affected signal that is input to the compensation filter array. The desired signal is shown in Figure 6.22 and the restoration results are shown in Figure 6.23.

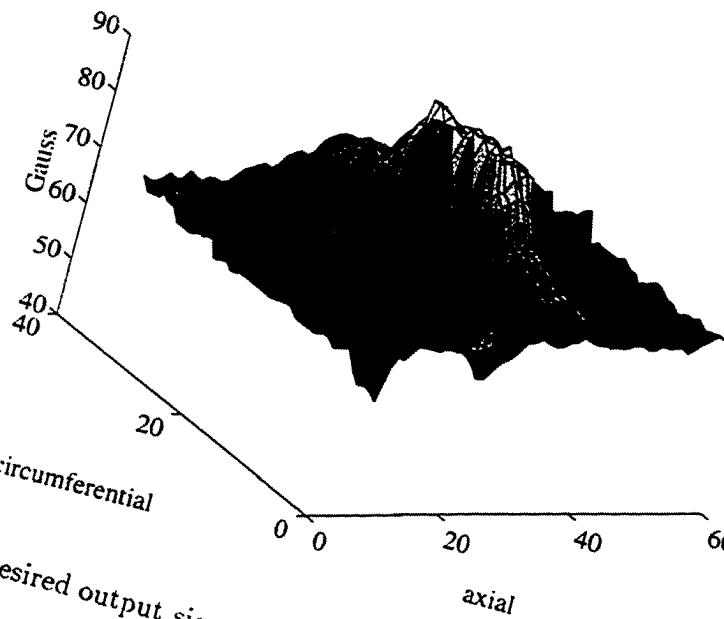


Figure 6.22: Desired output signal of the 3-D velocity effects compensation scheme

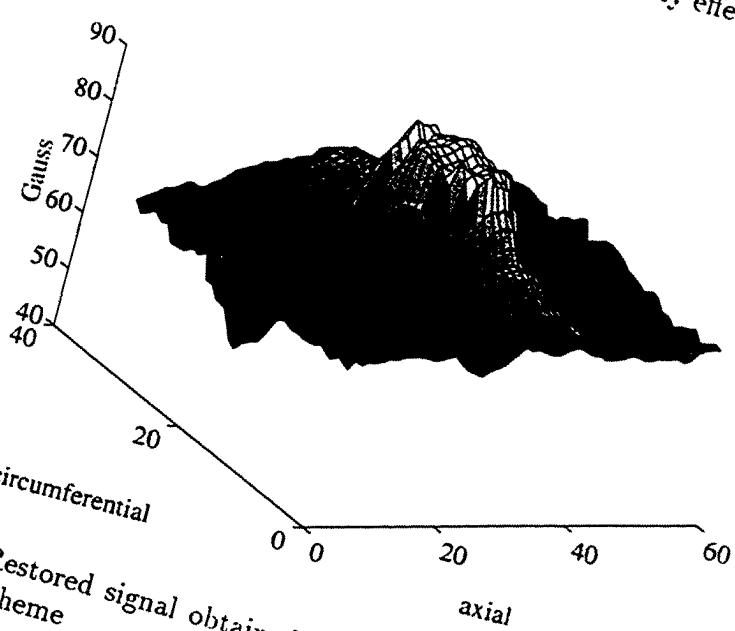


Figure 6.23: Restored signal obtained from the 3-D velocity effects compensation scheme

CHAPTER 7. CONCLUSIONS

7.1 Summary

This research work was motivated by the desire to design and develop signal processing techniques that will enhance the accuracy of interpreting magnetostatic flux leakage data. In general, the solution of the inverse problem in MFL nondestructive evaluation is complex. From an information theory point of view, it can be shown that the solution of the inverse problem becomes more and more difficult as we go from wave to diffusion and to static phenomena. The work presented here is an attempt to solve some of the problems that occur during the interpretation of the MFL signal. A typical defect characterization paradigm employs an artificial neural network to perform a multidimensional mapping from the MFL signal to the defect profile. However, the effects of operational variables inherent in the signal measurement process, can deter the performance of such neural network schemes. This dissertation presents novel techniques to overcome the obstacles posed by the operational parameters for accurate defect characterization. The research work carried out also has direct and immediate applications to problems currently faced by the gas pipeline industry and tends to advance the state-of-the-art in pipeline inspection systems.

The problem of rendering a signal invariant to changes in experimental param-

eters has been addressed. This problem is of fundamental importance in almost all data interpretation applications. In the case of nondestructive evaluation of gas transmission pipelines by MFL methods, the operational parameters that have been studied are the magnetization characteristic (permeability) of the pipe and the velocity of the scanning device. These parameters are shown to cause significant variation in the resulting MFL signals, that could result in inaccurate defect characterization.

There are two principal reasons for variation in the B-H curve of the pipe-wall, namely, grade of the pipe material and residual stresses in the pipe-wall. Local variations in the pipe-wall permeability are responsible for the alteration of the MFL signal, in a manner very similar to changes in the signal due to varying defect depth. To complicate matters further, in practice, the magnetization characteristic of each section of the pipe that is being scanned is not known. In this study, methods have been developed for rendering the MFL signal insensitive to permeability changes in the pipe wall. At the same time, its sensitivity to defect depth is preserved.

The second operational parameter considered is tool velocity. As the magnetizer-sensor assembly that forms the scanning tool is conveyed along the pipe under the pressure of natural gas, motionally induced currents in the pipe-wall distort the signal received by the Hall-effect sensors. Signal processing algorithms have been developed to compensate for the effects of this tool velocity. Velocity affected MFL signals from all manner of defects are converted to equivalent reference zero-velocity signals, that are pre-determined for a representative subset of defects from experimental or finite element studies. In order to do this using experimental signals, it has been recognized that the inspection process is not exactly repeatable. Issues in signal registration and variations of the origin of sampling have been addressed.

The compensation algorithms have been exercised with both finite element (permeability invariance) and experimental (velocity invariance) data. Although it is not shown in this dissertation, these invariance algorithms improve subsequent defect characterization results [64].

7.2 Conclusions

In order to perform characterization of MFL NDE signals in terms of the underlying defect parameters and in spite of the presence of operational variables, two methods to achieve invariance transformations, namely, feature selection and signal compensation have been investigated. It is imperative that subjecting a signal to these transformations must not in any way reduce or obscure defect related information.

7.2.1 Permeability invariance

The feature selection scheme is an explicit method to construct the invariance transformation and is employed when the value of the undesirable parameter is unknown. Moreover, it is very useful when the operational and defect related parameters affect the signal in an identical manner. This type of invariance scheme is required to compensate MFL signals for variations in pipe-wall permeability. An entirely new feature selection scheme is developed wherein the invariance transformation is reduced to a problem in multidimensional data interpolation. The underlying philosophy for recasting the invariance problem in this manner is the premise that one can obtain two signals characterizing the same NDE phenomenon such that they respond differently to changes in the undesirable test parameter. It is this dissimilar parametric

response that is exploited in formulating the invariance transformation. In this sense, the feature selection method is nothing but a method for data fusion.

Reformulating the invariance transformation as a data interpolation problem has led to investigations into phenomenological and non-phenomenological interpolation methods. Phenomenological interpolation schemes are derived from analytical models of the MFL process, which due to their inherent limitations lend themselves to the construction of only approximate models. Initial forays into non-phenomenological methods included commercial curve fitting packages that provide simple interpolating functions such as quadratics, exponentials and logarithmic functions. The disadvantages associated with these two methods are:

1. These schemes are not adequate for estimating the transformation parameters for multiple defect classes, since the resulting matrices in the equation defining the invariance transformation are not always invertible.
2. Consequently, the performance of the transformation, outside of the class for which the parameters are determined, is poor.
3. Furthermore, the “invariant” feature so derived does not vary monotonically with change in defect dimensions.

As a result these methods, although simple to implement, are not suitable for ensuring accurate defect characterization.

7.2.1.1 Universal approximation methods: Neural network schemes
 Explorations in the arena of data interpolation yield a surfeit of universal approximation techniques, the most widely used being feed-forward neural networks. Studies

into the approximation capabilities of the various types of neural networks showed that for this particular application, networks with Euclidean norm type activation functions, such as radial basis function networks, are superior to multilayer perceptron type networks.

First, the radial basis function network was employed as a tool to construct the invariance transformation. The parameters defining the network can be determined simultaneously for various defect classes by direct inversion of the basis function matrix. The limitations of the earlier schemes are thus overcome. Although the RBF network performs exceedingly well as an invariant feature estimator for training data, its performance for testing data is woefully inadequate. A careful examination of the input data feature vectors in multidimensional space reveals that variations in permeability require a local approximator and variations in defect dimension require a global approximation scheme. The RBF network proved incapable of providing both these capabilities simultaneously.

Next, fuzzy inference systems, were studied to see if they could offer greater control over local and global interpolation performance. This study resulted from recent claims in neural network literature that fuzzy systems can simultaneously provide local and global interpolation. Further investigations revealed that RBF networks and FBF networks are functionally equivalent; improved interpolation is possible only by using optimal training methods to determine the basis / membership function parameters. The gradient descent scheme was employed to find the parameters of an FBF network with Gaussian basis functions. While this method yields relatively better interpolation results for testing data arising from variations in pipe-wall permeability; interpolation performance for variation in defect depth is still inadequate.

Finally, it was recognized that multi-resolution approximation capabilities were desired, since feature variations for defect depth and permeability occur at differing resolutions (frequencies) of the input feature space. Wavelet basis networks were determined to be the ideal candidates to perform the permeability invariance transformation. This method is a simple extension of the Gaussian RBF network and provides sufficient control of the local and global interpolation requirements. The WBF network, therefore, forms the culmination of this research for finding an appropriate universal approximation scheme to perform the invariance transformation.

It can be noticed that the invariant feature selection scheme developed here is sufficiently general and can be used in any signal characterization scheme where invariance to test parameters are desired. The only requirement of this technique is the availability of two signals that respond differently to the parametric variation.

7.2.2 Velocity invariance

The signal compensation scheme is an implicit method for invariant pattern recognition and is used when prior knowledge exists about the value of the distortion parameter. Such schemes are therefore appropriate for velocity effects compensation — an odometer onboard the scanning tool provides a continuous readout of the scan velocity. Signal compensation techniques are essentially seen as filtering schemes, wherein a velocity affected signal is input to a restoration filter that reconstructs the equivalent zero-velocity signal. Developing a signal restoration scheme essentially involves choosing an appropriate filter in the spatial (time) or frequency domain and then determining its weights.

Different types of filters were investigated to test whether they could adequately

perform the task of signal restoration over a suitable range of defect dimensions. Since the velocity of the tool alters the physical situation responsible for the MFL signal's origin, analytical filter models are difficult to compute. A semi-empirical analytical filter model was developed. However, the filter offers limited performance due to the gross nature of the approximations made during the derivation.

Deterministic Wiener filters derived using experimental data are significantly noisy and subsequent smoothing filters are required to suppress ripples introduced as artifacts in the restored signal. Finally, the statistical behavior of the signal was used as a basis to compute stochastic restoration filters and these offer the best performance to date, in terms of signal reconstruction.

The order of the optimal filter was studied; nonlinear second order Volterra filters were implemented and evaluated for superior restoration capabilities in compensating signals resulting from the nonlinear MFL process. However, the second order terms of the restoration filter were negligibly small, thus indicating that a linear filter is adequate. The apparent lack of second order effects could either be due to the size of the defect set used in building the filter kernel, or due to some manner of self-compensation in the MFL signal due to the velocity induced terms.

Filter coefficients were also determined using the stochastic gradient descent algorithm. The advantages of this method are that it does not require the inversion of the auto-correlation matrix and the performance of the filter is more robust, especially for outlier data. The well known concept of trading bias for variance has been exploited to assure consistent performance over a larger defect set.

As a necessary step in the filter design process, simple methods for signal/image registration and schemes for compensating for the change in the origin of sampling

were developed. Signal registration is achieved very simply by studying the cross-correlation function between the filter input and the desired output; this procedure occurs automatically when determining the filter coefficients. The effect of changes in the origin of sampling is removed by using interpolation-decimation cascade algorithms.

At this point, it is pertinent to mention that both the feature selection and signal compensation methods suffer from one disadvantage. Both of these techniques require a dense set of training data in order to determine the invariance transformation parameters. The performance of the transformation deteriorates outside of the upper and lower bounds specified by the training data set. While experimentally obtained data provides the best form of training data, it is not always convenient to procure and use for the following reasons:

1. Pigging experiments are very expensive. A single run of the pig at a pull-rig facility costs about \$20,000.
2. Experimental data is almost always corrupted by noise and image registration is not easy. While these may signify actual operating conditions, they are mere hindrances at an algorithm development stage.
3. Some manner of experiments cannot be performed in a controlled manner. For example, MFL signals are required from a family of B-H curves for designing the permeability invariance scheme. It is not possible to experimentally ensure that the B-H characteristics of an actual pipe lie in the range specified by the functional interpolation mechanism with the required degree of accuracy.

Therefore, availability of numerical models allow the generation of training data that

cover a significant range of defects, in an inexpensive and convenient manner. The other important function of such a numerical model is to lend insight into the physical operation of the devices upon which all signal processing algorithms must be based.

Although the invariance transformation methods developed in this research work can be employed to compensate signals arising from a certain range of defect profiles and magnetization levels, they offer the most pragmatic approach for solving a particularly difficult inverse problem. In any case, it is unlikely that any engineering approach for solving inverse problems can operate effectively without appropriate constraints, regardless of the region of application. In conclusion, the invariance transformation methods that have been developed in this dissertation show considerable promise of arriving at a comprehensive method for interpreting MFL signals.

7.3 Future Directions for Research

Future research in this area can focus on the following two fronts:

7.3.1 Improvements in the invariance algorithms

The invariance algorithms hitherto developed for compensating the MFL signal for variations in pipe-wall permeability and scan velocity can be improved in the following manner:

1. The wavelet basis function network that has been developed for permeability invariance can be optimally trained using adaptive techniques, to increase its robustness.

2. For the invariance transformation algorithms to be useful in the field, both the neural network and restoration filter schemes should also predict the reliability of the transformation, by using measures of confidence. These measures must be encoded into the algorithm [65].
3. Further study determining sensor-defect interaction in the presence of velocity could lead to improved restoration filter design, perhaps by purely analytical methods.

7.3.2 MFL signal characterization

There remain other significant parameters that need to be tackled before an MFL signal that depends only on the defect profile can be generated. Remanent magnetization is believed to result in signals that are not repeatable over successive runs of the pig; methods to compensate for this effect need to be investigated. The position of the Hall-effect sensor relative to the poles plays a significant part in determining the resulting MFL signal. Contemporary studies by research colleagues show that the effect of velocity is diminished by judicious sensor placement. Similar hardware based compensation schemes must be studied. Furthermore, the algorithms that have been developed here need to be thoroughly tested with more complicated defect shapes. For these techniques to be useful to the gas pipeline industry, real time implementation of the software is required. Much of this work, however, is already in progress.

The methods developed in this dissertation represent a beginning in addressing the larger problem of accurately characterizing magnetostatic flux leakage NDE signals. However, the invariance transformations have been developed using a general

purpose framework and are applicable for addressing a number of other operational variables, not only in MFL inspection, but also in other NDE methods.

BIBLIOGRAPHY

- [1] W. Lord, "Partial differential equations and the NDE inverse problem," *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 7A, pp. 425-430, Plenum Press, New York, 1985.
- [2] W. Lord and L. Udpa, "Imaging of electromagnetic NDT phenomena." *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 5A, pp. 465-471, Plenum Press, New York, 1986.
- [3] W. Lord and D. J. Oswald, "Leakage methods of field detection," *International Journal of Nondestructive Testing*, Vol. 4, pp. 249-274, Gordon and Breach, Great Britain, 1972.
- [4] W. Lord, J. M. Bridges and R. Palanisamy, "Residual and active leakage fields around defects in ferromagnetic materials," *Materials Evaluation*, Vol. 36, No. 8, pp. 47-54, July 1978.
- [5] Y. Shin, *Numerical Modeling of Probe-velocity Effects for Eddy Current NDT*, Ph. D. Dissertation, Iowa State University, Ames, IA, 1992.
- [6] R. Palanisamy and W. Lord, "Finite element modeling of electromagnetic NDT phenomena," *IEEE Transactions on Magnetics*, Vol. MAG-15, No. 6, pp. 1479-1481, November 1979.
- [7] D. C. Jiles and D. L. Atherton, "Theory of the magnetization process in ferromagnets and its application to the magnetomechanical effect," *Journal of Physics D: Applied Physics*, Vol. 17, pp. 1265-1281, 1984.

- [8] R. Langman, "The effect of stress on the magnetization of mild steel at moderate field strengths," *IEEE Transactions on Magnetics*, Vol. MAG-21 No. 4, 1985.
- [9] T. A. Bubenik, J. B. Nestleroth, M. J. Koenig, S. Udpa, A. E. Crouch, P. C. Porter and A. Teitsma, *Characterization of Magnetic Flux Leakage (MFL) Indications found during In-line Inspection of Natural Gas Transmission Pipelines: Monthly Report*, p. 33, GRI, Chicago, IL, March 1994.
- [10] R. P. Lippmann, "An introduction to computing with neural nets," *IEEE ASSP Magazine*, Vol. 4, pp. 4–22. April 1987.
- [11] S. Haykin, *Neural Networks. A Comprehensive Foundation*, Macmillan, New York, 1994.
- [12] E. Barnard and D. Casasent, "Invariance and neural nets," *IEEE Transactions on Neural Networks*, Vol. 2, No. 5, pp. 498–508, September 1991.
- [13] Y. S. Abu-Mostafa and D. Psaltis, "Image normalization by complex moments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-7, pp. 46–55, January 1985.
- [14] S. S. Udpa and W. Lord, "A Fourier descriptor classification scheme for differential probe signals." *Materials Evaluation*, 42(9), pp. 1136–1141, 1984.
- [15] G. B. Gurevich, *Foundations of the Theory of Algebraic Invariants*, P. Noordhoff Ltd., The Netherlands, 1964.
- [16] T. Y. Thomas, *The Concept of Invariance in Mathematics*, University of California Press, Berkeley and Los Angeles, CA, 1944.
- [17] *Encyclopaedia of Mathematics: An Updated and Annotated Translation of the Soviet Mathematical Encyclopaedia*, Kluwer Academic Publishers, Dordrecht, Holland, 1988.
- [18] H. W. Turnbull, *The Theory of Determinants, Matrices, and Invariants*, Dover, New York, 1960.

- [19] R. J. Holt and A. Netravali, "Using affine invariants on perspective projections of plane curves," *Computer Vision and Image Understanding*, Vol. 61, No. 1, pp. 112–121, January 1995.
- [20] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York, 1973.
- [21] E. B. Barret, P. M. Payton, N. N. Haag and M. H. Brill, "General methods for determining projective invariants in imagery," *Computer Vision, Graphics, and Image Processing: Image Understanding*, Vol. 53, No. 1, pp. 46–65, January 1991.
- [22] M. K. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, Vol. IT-8, pp. 179–187, 1962.
- [23] A. Khotanzad and Y. H. Hong, "Rotation invariant pattern recognition using Zernicke moments," *Proceedings of the 9th International Conference on Pattern Recognition*, Rome, Italy, pp. 326–328, November 1988.
- [24] G. H. Granlund, "Fourier preprocessing for hand print character recognition," *IEEE Transactions on Computers*, Vol. C-21, pp. 195–201, February 1971.
- [25] R. C. Gonzalez and P. Wintz, *Digital Image Processing*, Addison-Wesley, Reading, MA, 1987.
- [26] D. C. Ghiglia and C. V. Jakowatz, "Some practical aspects of moving object deblurring in a perspective plane," *Applied Optics*, Vol. 24, No. 22, pp. 3830–3837, November 1985.
- [27] C. A. Micchelli, "Interpolation of scattered data: distance matrices and conditionally positive definite functions," *Constructive Approximation*, Vol. 2, pp. 11–22, 1986.
- [28] T. Sudkamp and R. J. Hammel II, "Interpolation, completion and learning fuzzy rules," *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 24, No. 2, pp. 332–342, February 1994.

- [29] B. R. Bakshi and G. Stephanopoulos, "Wave-Net: a multiresolution, hierarchical neural network with localized learning," *AICHE Journal*, Vol. 39, pp. 57–81, January 1993.
- [30] E. W. Cheney, *Introduction to Approximation Theory*. McGraw-Hill Book Company, New York, 1966.
- [31] R. C. Williamson and U. Helmke, "Existence and uniqueness results for neural network approximations," *IEEE Transactions on Neural Networks*, Vol. 6, No. 1, pp. 2–13, January 1995.
- [32] D. Braess, *Nonlinear Approximation Theory*, Springer-Verlag, New York, 1986.
- [33] N. E. Cotter, "The Stone-Weierstrass theorem and its application to neural networks," *IEEE Transactions on Neural Networks*, Vol. 1, No. 4, pp. 290–295, December 1990.
- [34] N. N. Zatsepin and V. E. Shcherbinin, "Calculation of the magnetostatic field of surface defects, 1. Field topography of defect models," *Defektoskopiiya*, No. 5, pp. 50-59, September-October 1966.
- [35] *TableCurve for Windows*, Jandel Scientific Software, San Rafael, CA, c1993.
- [36] K. Hornik, M. Stinchcombe and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, Vol. 2, pp. 359–366, 1989.
- [37] T. Chen and H. Chen, "Approximation capability to functions of several variables, nonlinear functionals, and operators by radial basis function neural networks," *IEEE Transactions on Neural Networks*, Vol. 6, No. 4, pp. 904–910, July 1995.
- [38] T. Chen and H. Chen, "Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems," *IEEE Transactions on Neural Networks*, Vol. 6, No. 4, pp. 911–917, July 1995.
- [39] F. Girosi and T. Poggio, "Networks and the best approximation property," *Biological Cybernetics*, Vol. 63, pp. 169–176, 1990.

- [40] D. S. Broomhead and D. Lowe, "Multivariate functional interpolation and adaptive networks," *Complex Systems*, Vol. 2, pp. 321–355, 1988.
- [41] J. T. Tou and R. C. Gonzalez, *Pattern Recognition Principles*, Addison-Wesley, Reading, MA, 1974.
- [42] M. T. Musavi, W. Ahmed, K. H. Chan, K. B. Faris and D. M. Hummels, "On the training of radial basis function classifiers," *Neural Networks*, Vol. 5, pp. 595–603, July-August 1992.
- [43] S. Chen, P. M. Grant and C. F. N. Cowan, "Orthogonal least-squares algorithm for training multioutput radial basis function networks," *IEE Proceedings. Part F, Radar and Signal Processing*, Vol. 139, No. 6, pp. 378–384, December 1992.
- [44] B. Kosko and S. Isaka, "Fuzzy logic," *Scientific American*, pp. 76–81, July 1993.
- [45] B. Kosko, "Fuzzy systems as universal approximators," *Proceedings of the 1st IEEE International Conference on Fuzzy Systems*, pp. 1153–1162, March 1992.
- [46] L. Wang and J. M. Mendel, "Fuzzy basis functions, universal approximation, and orthogonal least squares learning," *IEEE Transactions on Neural Networks*, Vol. 3, No. 5, pp. 807–814, September 1992.
- [47] J. S. Jang and C. T. Sun, "Functional equivalence between radial basis function networks and fuzzy inference systems," *IEEE Transactions on Neural Networks*, Vol. 4, No. 1, pp. 156–159, January 1993.
- [48] J. Nie and D. A. Linkens, "Learning control using fuzzified self-organizing radial basis function network," *IEEE Transactions on Fuzzy Systems*, Vol. 1, No. 4, pp. 280–287, November 1993.
- [49] J. A. Dickerson and B. Kosko, "Fuzzy function learning with covariance ellipsoids," *IEEE International Conference on Neural Networks*, pp. 1162–1167, March 1993.

- [50] H. Nomura, I. Hayashi and N. Wakami, "A learning method for fuzzy inference rules by descent method," *IEEE International Conference on Fuzzy Systems*, pp. 203–210, March 1992.
- [51] I. Elshafiey, L. Udpa and S. S. Udpa, "Solution of inverse problems in electromagnetics using Hopfield neural networks," *IEEE Transactions on Magnetics*, Vol. 31, No. 1, pp. 852–861, January 1995.
- [52] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, pp. 674–693, July 1989.
- [53] G. Strang, "Wavelets and dilation equations: A brief introduction," *SIAM Review*, Vol. 31, No. 4, pp. 614–627, December, 1989.
- [54] G. Strang, "Wavelets," *American Scientist*. Vol. 82, pp. 250–255, May-June, 1994.
- [55] X. Dai, B. Joseph and R. L. Motard, "Introduction to wavelet transformation and time frequency analysis," *Wavelet Applications in Chemical Engineering*, R. L. Motard and B. Joseph, Editors, Kluver Academic Publishers, Boston, MA, 1994.
- [56] A. P. Pentland, "Interpolation using wavelet bases," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 4, pp. 410–414, April 1994.
- [57] Q. Zhang and A. Beneviste, "Wavelet Networks," *IEEE Transactions on Neural Networks*, Vol. 3, No. 6, pp. 889–898, November 1992.
- [58] Y. C. Pati and P. S. Krishnaprasad, "Analysis and synthesis of feedforward neural networks using discrete affine wavelet transformations," *IEEE Transactions on Neural Networks*, Vol. 4, No. 1, pp. 73–85, January 1993.
- [59] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Englewood Cliffs, NJ, 1991.
- [60] L. L. Scharf, *Statistical Signal Processing, Detection, Estimation, and Time Series Analysis*, Addison-Wesley, Reading, MA, 1991.

- [61] G. L. Sicuranza, "Quadratic filters for signal processing," *Proceedings of the IEEE*, Vol. 80, No. 8, pp. 1263–1285, August 1992.
- [62] T. Koh and E. J. Powers, "Second-order Volterra filtering and its application to nonlinear system identification," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-33, No. 6, pp. 1445–1455, December 1985.
- [63] A. V. Oppenheim and R. W. Shafer, *Discrete-time Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1989.
- [64] M. Chao, *Neural Networks for Characterizing Magnetic Flux Leakage Signals*, M. S. Thesis, Iowa State University, Ames, IA, 1995.
- [65] J. A. Leonard, M. A. Kramer and L. H. Ungar, "A neural network architecture that computes its own reliability," *Computers and Chemical Engineering*, Vol. 16, No. 9, pp. 819–835, September 1992.
- [66] I. S. Gradshteyn, *Table of Integrals, Series and Products*, 5th ed, Academic Press, Boston, MA, 1994.

APPENDIX A. UNIVERSAL APPROXIMATION BY MULTILAYER FEEDFORWARD NETWORKS

Formal proofs are now given that multilayer feedforward networks of both the multilayer perceptron type and the radial basis function type are capable of universal approximation [36] [39]. Universal approximation is the property that the network can model any continuous function to an arbitrary degree of accuracy. This means that the approximating functional must possess the denseness property. Central to the proofs is the use of the Stone-Weierstrass theorem, which was stated in Chapter 4. In order to facilitate the proof for each of the two types of networks, the theorem is restated in alternative, albeit equivalent forms.

A.1 Multilayer Perceptron Networks

The proof given here, taken from Hornik et. al. [36] establishes that MLP networks with as few as one hidden layer are capable of universal approximation.

A.1.1 Definitions

A.1.1.1 Affine function A^r is the set of all affine functions of the form $A(x) = w \cdot x + b$ where w and x are vectors in \mathcal{R}^r , b is a scalar in \mathcal{R} and $r \in N \equiv \{1, 2, \dots\}$.

As applied to an MLP network, x , w and b represent the network inputs, weights and bias, respectively.

A.1.1.2 $\sum \prod$ MLP network For any measurable function $G(\cdot)$ mapping from \mathcal{R} to \mathcal{R} and $r \in \mathcal{N}$, $\sum \prod^r(G)$ is the class of functions of the form

$$\begin{aligned} \{f : \mathcal{R}^r \rightarrow \mathcal{R} : f(x) = & \sum_{j=1}^q \beta_j \cdot \prod_{k=1}^{l_j} G(A_{jk}(x)), \\ & x \in \mathcal{R}^r, \beta_j \in \mathcal{R}, A_{jk} \in \mathcal{A}^r, l_j \in \mathcal{N}, \\ & q = 1, 2, \dots \} \end{aligned}$$

The MLP network is represented by the $\sum \prod$ configuration.

A.1.1.3 Algebra A family of real functions, \mathcal{A} , defined on a set, \mathcal{E} , together with a scalar field \mathcal{F} is called an algebra [39] if \mathcal{A} is closed under addition (+), multiplication (\times) and scalar multiplication (\cdot), such that

1. \mathcal{A} , along with \mathcal{F} , (+) and (\times) form a linear space.

2. If $f, g, h \in \mathcal{A}$ and $\alpha \in \mathcal{F}$, then

(a) $f \times g \in \mathcal{A}$

(b) $f \times (g \times h) = (f \times g) \times h$

(c) $f \times (g + h) = f \times g + f \times h$

(d) $(f + g) \times h = f \times h + g \times h$

(e) $\alpha(f \times g) = (\alpha f) \times g = f \times (\alpha g)$

A.1.2 Stone-Weierstrass Theorem

If \mathcal{A} is an algebra of real continuous functions on a compact set K , and,

1. \mathcal{A} separates points on K , i.e., for every $x, y \in K$, $x \neq y$, there exists a function,

$f \in \mathcal{A}$ such that $f(x) \neq f(y)$,

2. \mathcal{A} vanishes at no point of K , i.e., if for each $x \in K$, there exists a function,

$f \in \mathcal{A}$ such that $f(x) \neq 0$.

then, \mathcal{A} is dense in the space of real continuous functions on K .

A.1.3 Theorem

If G is a continuous nonconstant function from \mathcal{R} to \mathcal{R} , then $\sum \Pi^r(G)$ is uniformly dense in, \mathcal{C}^r , the space of continuous functions from \mathcal{R}^r to \mathcal{R} .

A.1.3.1 Proof

The Stone-Weierstrass theorem is applied.

Let $K \subset \mathcal{R}^r$ be any compact set.

For any G , $\sum \Pi^r(G)$ is obviously an algebra on K .

If $x, y \in K$, $x \neq y$, then there exists an $A \in \mathcal{A}^r$ such that $G(A(x)) \neq G(A(y))$.

This can be shown by choosing $a, b \in \mathcal{R}$, $a \neq b$ such that $G(a) \neq G(b)$. Then A is chosen so that $A(x) = a$, $A(y) = b$. Then $G(A(x)) \neq G(A(y))$, or, $\sum \Pi^r(G)$ separates points on K .

There exist $G(A(\cdot))$'s that are constant and not equal to zero. This can be shown by choosing $b \in \mathcal{R}$ such that $G(b) \neq 0$ and setting $A(x) = 0 \cdot x + b$. For all $x \in K$, $G(A(x)) \equiv G(b)$. This ensures that $\sum \Pi^r(G)$ does not vanish at any point of K .

Since the conditions of the Stone-Weierstrass theorem are satisfied, $\sum \Pi^r(G)$ is dense in the space of real continuous functions on any arbitrary K ; hence the result.

□

A.2 Radial Basis Function Networks

The proof given here is taken from Girosi and Poggio [39] and shows that radial basis function networks with Gaussian basis functions can approximate any function to an arbitrary degree of accuracy.

A.2.1 Definition

A.2.1.1 Subalgebra A set \mathcal{S} is a subalgebra of the algebra \mathcal{A} if

1. \mathcal{S} is a linear subspace of \mathcal{A} .
2. \mathcal{S} is closed under the operation \times ; i.e., if f and g are in \mathcal{S} , then $f \times g$ is also in \mathcal{S} .

A.2.2 Stone-Weierstrass Theorem

If X is a compact metric space, $\mathcal{C}[X]$ is the set of continuous functions defined on X and \mathcal{A} is a subalgebra of $\mathcal{C}[X]$ with the following properties:

1. The function $f(x) = 1$ belongs to \mathcal{A} .
2. For any two distinct points $x, y \in X$, there exists $f \in \mathcal{A}$ such that $f(x) \neq f(y)$.

Then \mathcal{A} is dense in $\mathcal{C}[X]$.

A.2.3 Theorem

The Gaussian superposition set \mathcal{G}_x , defined by

$$\begin{aligned}\mathcal{G}_x \equiv \{f \in \mathcal{C}[X] : f(x) = \sum_{i=1}^m c_i \exp\left(-\frac{(x-t_i)^2}{\sigma_i^2}\right), \\ X \subset \mathbb{R}^d, t_i \in \mathbb{R}^d, c_i, \sigma_i \in \mathbb{R}, m \in \mathbb{N}\}\end{aligned}$$

is dense in $\mathcal{C}[X]$, where X is a compact subset of \mathbb{R}^d .

A.2.3.1 Proof As in the case of the MLP network, the Stone-Weierstrass theorem is applied.

The set \mathcal{G}_x will be a subalgebra of $\mathcal{C}[X]$ if the product of two of its elements yields another element of \mathcal{G}_x . Since \mathcal{G}_x is a linear superposition of Gaussians of different variances and means, it is sufficient to deal with the product of two Gaussians.

Consider

$$\exp\left(-\frac{(x-t_1)^2}{\sigma_1^2}\right) \cdot \exp\left(-\frac{(x-t_2)^2}{\sigma_2^2}\right) = \exp\left(-\frac{(x-t_3)^2}{\sigma_3^2}\right)$$

where

$$t_3 = \frac{\sigma_2^2 t_1 + \sigma_1^2 t_2}{\sigma_1^2 + \sigma_2^2}, \quad \sigma_3^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}, \quad c = \exp\left(-\frac{(t_1 - t_2)^2}{\sigma_3^2}\right)$$

This shows that the product of two Gaussians with means t_1 and t_2 is proportional to a Gaussian with mean t_3 that is a convex linear combination of t_1 and t_2 . Therefore, \mathcal{G}_x is a subalgebra of $\mathcal{C}[X]$.

The function, $f(x) = 1$ can be considered to be a Gaussian of infinite variance and therefore belongs to \mathcal{G}_x .

Obviously, for any two distinct points x, y , we can find a function in \mathcal{G}_x such that $f(x) \neq f(y)$.

Since all the conditions of the Stone-Weierstrass theorem are satisfied, \mathcal{G}_x is dense in $\mathcal{C}[X]$.

□

APPENDIX B. BEST APPROXIMATION PROPERTY OF MULTILAYER FEEDFORWARD NETWORKS

An approximating scheme is known as the best approximation if in its set of approximating functions, there exists one that has a minimum distance from any function of a larger set of functions. Function sets that possess the best approximation property are called existence sets. The following proof, which demonstrates that RBF networks are the best approximators and MLP networks are not, is reproduced from Girosi and Poggio [39]. Essentially, the test for best approximation is a test for closure or compactness. Some of the fundamental definitions and theorems related to linear metric spaces are revisited below.

B.1 Fundamental Definitions and Theorems

B.1.1 Definitions

B.1.1.1 Closure A set F in a metric space is said to be *closed* if the limit of every convergent sequence in F is also in F .

B.1.1.2 Compactness If X is a metric space, $K \subset X$, is said to be *compact* if every sequence of points in K has a subsequence which converges to a point of K .

B.1.2 Theorems

B.1.2.1 Theorem 1 Every compact set is an existence set.

B.1.2.2 Theorem 2 Every existence set is a closed set.

B.1.2.3 Theorem 3 Every closed, bounded, finite dimensional set is a compact set.

B.2 Multilayer Perceptron Networks

B.2.1 Proposition

Backpropagation type multilayer networks, defined by functions of the set σ^m ,

$$\begin{aligned}\sigma^m \equiv \{f \in \mathcal{C}[U] : f(x) = \sum_{i=1}^m c_i \sigma(\mathbf{x} \cdot \mathbf{w}_i + \theta_i), \\ \mathbf{w}_i \in \mathcal{R}^d, c_i, \theta_i \in \mathcal{R}\end{aligned}$$

where $\sigma(x)$ is usually a sigmoidal function, do not have the best approximation property because the set σ^m is not an existence set for $m \geq 2$.

B.2.1.1 Proof The result is proved for the case of one hidden layer with σ being a sigmoidal function of the form $\sigma(x) = (1 + e^{-x})^{-1}$.

Since every existence set must be closed (by Theorem 2), it is sufficient to show that σ^m is not closed. This can be done by showing that the limit of a convergent sequence in the set does not belong to it. Consider the function

$$f_\delta(\mathbf{x}) = \frac{1}{\delta} \left(\frac{1}{1 + e^{-(\mathbf{w} \cdot \mathbf{x} + \theta)}} - \frac{1}{1 + e^{-(\mathbf{w} \cdot \mathbf{x} + (\theta + \delta))}} \right)$$

Whereas $f_\delta \in \sigma^m$, $\forall m \geq 2$, it can be seen that

$$\lim_{\delta \rightarrow 0} f_\delta(\mathbf{x}) \equiv g(\mathbf{x}) = \frac{1}{2(1 + \cosh[\mathbf{w} \cdot \mathbf{x} + \theta])} \quad (\text{B.1})$$

and $g \notin \sigma^m$, $\forall m \geq 2$. Therefore σ^m is not closed. This concludes the proof.

□

B.3 Radial Basis Function Networks

B.3.1 Proposition

The set of radial basis function networks, T^m

$$T^m \equiv \{f \in \mathcal{C}[U] : f(\mathbf{x}) = \sum_{i=1}^m c_i \phi_i(\mathbf{x}), \quad c_i \in \mathcal{R}\}$$

possesses the best approximation property because T^m is an existence set for $m \geq 1$.

B.3.1.1 Proof Let f be a prescribed element of $\mathcal{C}[U]$, and a_0 be an arbitrary point of T^m . The closest point to f from T^m has to lie in the set

$$\{s \in T^m : \|a - f\| \leq \|a_0 - f\|\}$$

This set is obviously closed and bounded. Therefore, by Theorem 3 it is compact. Consequently, by Theorem 1 it is an existence set. This completes the proof.

□

APPENDIX C. FUNCTIONAL EQUIVALENCE BETWEEN RBF NETWORKS AND FUZZY INFERENCE SYSTEMS

The following proof, taken from Jang and Sun [47] shows that under certain conditions, a fuzzy inference system with IF-THEN rules produces the same output as a radial basis function network. In that respect, fuzzy inference systems and RBF networks are functionally equivalent.

Consider a fuzzy inference system as shown in Figure 4.8 in Chapter 4 with multiple inputs and a single output. The fuzzy rule base consists of the rules

IF

x_1 is A_{i1} AND x_2 is A_{i2}AND x_m is A_{im}

THEN

y is f_i

where x_1, x_2, \dots, x_m are fuzzy input variables characterized by membership functions, $A_{i1}, A_{i2}, \dots, A_{im}$ and y is the output variable.

Let the AND's in the above rule base be specified by a product rule, i.e.,

$$\mu_i = A_{i1} \cdot A_{i2} \cdot \dots \cdot A_{im} \quad (C.1)$$

The inference procedure then maps the input variables \mathbf{x} to the output variable y by

the rule

$$y = f(\mathbf{x}) = \frac{\sum_{i=1}^n \mu_i f_i}{\sum_{i=1}^n \mu_i} \quad (\text{C.2})$$

If the membership functions A_{ij} are Gaussian,

$$A_{ij}(x_j) = \exp \left[-\frac{(x_j - c_{ij})^2}{2r_i^2} \right] \quad (\text{C.3})$$

then,

$$\begin{aligned} \mu_i &= A_{i1} \cdot A_{i2} \cdot \dots \cdot A_{im} \\ &= \exp \left[-\frac{(x_1 - c_{i1})^2}{2r_i^2} \right] \cdot \exp \left[-\frac{(x_2 - c_{i2})^2}{2r_i^2} \right] \cdot \dots \cdot \exp \left[-\frac{(x_m - c_{im})^2}{2r_i^2} \right] \\ &= \exp \left[-\frac{\|\mathbf{x} - \mathbf{c}_i\|^2}{2r_i^2} \right] \end{aligned} \quad (\text{C.4})$$

Therefore the output of the FIS is given by,

$$y = f(\mathbf{x}) = \frac{\sum_{i=1}^n \exp \left[-\frac{\|\mathbf{x} - \mathbf{c}_i\|^2}{2r_i^2} \right] f_i}{\sum_{i=1}^n \exp \left[-\frac{\|\mathbf{x} - \mathbf{c}_i\|^2}{2r_i^2} \right]} \quad (\text{C.5})$$

Equation (C.5) however represents the output of a normalized RBF network with Gaussian basis functions and n hidden nodes. This completes the proof.

□

APPENDIX D. GAUSSIAN WAVELETS AND THE ADMISSIBILITY CONDITION

The scaling and wavelet functions, ϕ and ψ respectively, of the Gaussian wavelets take the form

$$\phi(x) = \exp\left(-\frac{(x-c)^2}{2\sigma^2}\right) \quad (\text{D.1})$$

$$\psi(x) = \frac{\partial}{\partial x}(\phi(x)) = -\frac{(x-c)}{\sigma^2} \exp\left(-\frac{(x-c)^2}{2\sigma^2}\right) \quad (\text{D.2})$$

In order for these functions to be legitimate wavelets, $\psi(x)$ should satisfy the admissibility condition, that its Fourier transform, $\Psi(\omega)$, have the property

$$\int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{\omega} d\omega = C_{\psi} < \infty \quad (\text{D.3})$$

which implies that

$$\int_{-\infty}^{\infty} \psi(x) dx = 0 \quad (\text{D.4})$$

We now show that the Gaussian wavelet (D.2) satisfies the condition (D.3). First, the Fourier transform of $\psi(x)$ is computed. We use the property

$$\mathcal{F}\left\{\frac{df(x)}{dx}\right\} = (-j\omega)\mathcal{F}(f) \quad (\text{D.5})$$

where \mathcal{F} is the Fourier transform operator. Therefore,

$$\Psi(\omega) = (-j\omega)\Phi(\omega) \quad (\text{D.6})$$

For simplicity, choose a Gaussian scaling function with zero mean ($c = 0$) and unit variance ($\sigma = 1$). Then [66],

$$\begin{aligned}\Phi(\omega) &= \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2}\right) \exp(j\omega x) dx \\ &= \exp\left(-\frac{\omega^2}{2}\right)\end{aligned}\quad (\text{D.7})$$

Therefore,

$$\begin{aligned}\Psi(\omega) &= (-j\omega) \exp\left(-\frac{\omega^2}{2}\right) \\ |\Psi(\omega)|^2 &= \omega^2 \exp(-\omega^2) \\ \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{\omega} d\omega &= \int_{-\infty}^{\infty} \omega \exp(-\omega^2) d\omega \\ &= -\frac{1}{2} [\exp(-\omega^2)]_{-\infty}^{\infty} \\ &= 0 < \infty\end{aligned}\quad (\text{D.8})$$

The integral in (D.3) is finite; consequently the admissibility condition is satisfied. Equivalently, we can show that $\psi(x)$ satisfies the condition (D.4)

$$\begin{aligned}\int_{-\infty}^{\infty} \psi(x) dx &= \int_{-\infty}^{\infty} \frac{d}{dx} \phi(x) dx \\ &= \int_{-\infty}^{\infty} d(\phi(x)) dx \\ &= [\phi(x)]_{-\infty}^{\infty} \\ &= \left[\exp\left(-\frac{x^2}{2}\right) \right]_{-\infty}^{\infty} \\ &= 0\end{aligned}\quad (\text{D.9})$$

□