Ultrasonic flaw classification: an approach using modeling, signal processing, and adaptive learning

Lat Sang Koo
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd
Part of the Applied Mechanics Commons

Recommended Citation
https://lib.dr.iastate.edu/rtd/8667
INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the original text directly from the copy submitted. Thus, some dissertation copies are in typewriter face, while others may be from a computer printer.

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 copyrighted 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 oversize page is available as one exposure on a standard 35 mm slide or as a 17" × 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. 35 mm slides or 6" × 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.
Ultrasonic flaw classification: An approach using modeling, signal processing, and adaptive learning

Koo, Lat Sang, Ph.D.
Iowa State University, 1987
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark √.

1. Glossy photographs or pages ______
2. Colored illustrations, paper or print ______
3. Photographs with dark background ______
4. Illustrations are poor copy ______
5. Pages with black marks, not original copy ______
6. Print shows through as there is text on both sides of page ______
7. Indistinct, broken or small print on several pages ✓
8. Print exceeds margin requirements ______
9. Tightly bound copy with print lost in spine ______
10. Computer printout pages with indistinct print ______
11. Page(s) ______ lacking when material received, and not available from school or author.
12. Page(s) ______ seem to be missing in numbering only as text follows.
13. Two pages numbered ______. Text follows.
14. Curling and wrinkled pages ______
15. Dissertation contains pages with print at a slant, filmed as received ✓
16. Other ________________________________

______________________________
______________________________

UMI
Ultrasonic flaw classification: an approach using modeling, signal processing, and adaptive learning

by

Lat Sang Koo

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

Department: Engineering Science and Mechanics
Major: Engineering Mechanics

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

Iowa State University
Ames, Iowa

1991
<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. INTRODUCTION</strong></td>
</tr>
<tr>
<td>1.1 Statement of the Problem</td>
</tr>
<tr>
<td>1.2 Review of the Literature</td>
</tr>
<tr>
<td>1.2.1 Scattering theory</td>
</tr>
<tr>
<td>1.2.2 Signal processing</td>
</tr>
<tr>
<td>1.2.3 Adaptive learning</td>
</tr>
<tr>
<td><strong>2. MODELS</strong></td>
</tr>
<tr>
<td>2.1 Introduction</td>
</tr>
<tr>
<td>2.2 General Formulation</td>
</tr>
<tr>
<td>2.3 Born Approximation</td>
</tr>
<tr>
<td>2.3.1 Frequency domain results</td>
</tr>
<tr>
<td>2.3.2 Time domain results</td>
</tr>
<tr>
<td>2.4 Kirchhoff Approximation</td>
</tr>
<tr>
<td>2.4.1 Frequency domain results</td>
</tr>
<tr>
<td>2.4.2 Time domain results</td>
</tr>
<tr>
<td>2.5 Leading Edge Response</td>
</tr>
<tr>
<td><strong>3. SIGNAL PROCESSING AND FEATURE IDENTIFICATION</strong></td>
</tr>
<tr>
<td>3.1 Introduction</td>
</tr>
<tr>
<td>3.2 Finite Bandwidth Effects on Ideal Features</td>
</tr>
<tr>
<td>3.3 Improvement due to Extrapolation</td>
</tr>
<tr>
<td>3.3.1 Iteration</td>
</tr>
<tr>
<td>3.3.2 Polynomial</td>
</tr>
<tr>
<td>3.4 Deconvolution of Original Signal</td>
</tr>
<tr>
<td>3.4.1 Measurement model of Thompson-Gray</td>
</tr>
<tr>
<td>3.4.2 Flaw-derived reference</td>
</tr>
<tr>
<td><strong>4. FEATURE EXTRACTION AND FLAW CLASSIFICATION PROCEDURES</strong></td>
</tr>
<tr>
<td>4.1 Introduction</td>
</tr>
<tr>
<td>4.2 Brief Discussion of Adaptive Learning</td>
</tr>
<tr>
<td>4.2.1 Adaptive learning by parameter adjustment</td>
</tr>
<tr>
<td>4.2.2 Linearly separable system</td>
</tr>
<tr>
<td>4.2.3 Major disadvantages and advantages of adaptive learning</td>
</tr>
<tr>
<td>4.3 Feature Extraction</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

1.1 Statement of the Problem

The field of Nondestructive Evaluation (NDE) is concerned with the nondestructive injection of energy, in various forms, into a material in order to evaluate the internal state of that material. The form of energy that we will be concerned with here is high frequency acoustic energy, i.e., ultrasonic waves. When a propagating ultrasonic wave interacts with a defect in the material, scattered waves are produced which can be monitored and used for evaluation purposes. There are actually three separate stages of this evaluation process. First, there is the problem of flaw detection, the determination that a signal received is actually from an unwanted defect. Second, once a flaw has been detected, there is the problem of flaw classification - the identification of a flaw as cracklike or volumetric. Finally, given a classification there is the problem of flaw characterization - estimating the size, shape, orientation and material properties of the flaw. This thesis will be concerned exclusively with the flaw classification problem. Thus, we will assume that the flaw has already been detected and an identifiable flaw signal is available. Our focus will be on obtaining a new methodology for performing the classification process. As we will show, this methodology will be based on three separate foundations - model studies, signal processing, and adaptive learning. Although these foundations have been used previously in flaw classification and characterization studies, this work applies and
combines elements of these three fields in a number of novel ways:

1. Model Building

We use the method of integral equations to develop the Born and Kirchhoff approximations for estimating the response of volumetric and cracklike flaws, respectively. It is demonstrated that these models, although approximate, can predict the leading edge response of flaws accurately and that there are significant features in this leading edge response that can be used for classification purposes. It should be noted that the Born approximation results we have obtained are based on a new surface integral formulation that may offer significant advantages when higher order Born series terms are needed. For the present classification purposes, however, only the first Born approximation is employed.

2. Signal Processing

Since the features we propose to use for flaw classification are based on theoretical models, it is essential that the differences between the experimental responses obtained from measurements and these theoretical models be minimized. For example, those aspects in the experimental responses that are due to system and material aspects that are unrelated to the flaw itself, such as pulser-receiver circuit characteristics, transducer diffraction effects, attenuation corrections, etc. must be eliminated. One way for accomplishing this elimination is through a deconvolution procedure using the measurement model of Thompson and Gray [1,2]. Here, we will show that a simple and powerful flaw-derived deconvolution procedure can be used instead. This
new procedure has the advantage of not requiring the separate reference and/or attenuation measurements of the measurement model.

Even after nonflaw specific parts of an experimental response are removed through the use of deconvolution procedures, one must deal with the bandlimited nature of all ultrasonic systems. It is necessary to attempt to recover those frequencies that have been lost, particularly at the low frequency end of the spectrum, since such a loss can severely distort the signals from those found in the theoretical models and make the classification process difficult if not impossible. One method that has been used for low frequency restoration has been that of polynomial extrapolation [3]. Although this method can work effectively, it relies on a priori choices of fitting polynomial order and fitting region that, if chosen improperly, can cause the method to fail. We have shown that an alternative extrapolation procedure, based upon the Gerchberg-Papoulis algorithm [4,5], requires much less a priori information than polynomial extrapolation and has equal or better performance characteristics.

3. Adaptive Learning

Even if a "good" set of features are available for flaw classification, the noise and uncertainties always present in any real set of ultrasonic measurements make it difficult to precisely state a decision-making criterion. One method that has been used in the past to develop such criteria is that of adaptive learning [6,7]. Unfortunately, previous adaptive learning approaches (a) have been based on features that are not fundamental, i.e., that have not been closely
tied to analytical or numerical models of the scattering process, and
(b) have used multiple features and nonlinear decision surfaces that
make it impossible to analyse the system when failures occur. We have
attempted to minimize such limitations of the adaptive learning approach
by dealing only with a very few (2-3) model based features. This
allows us to modify the features so that a simple linear decision
surface is sufficient and to graphically plot the surface and feature
values to see where and why failures occur.

In summary, there are four major accomplishments of this thesis.
(1) The development of a subtracted surface integral approach for
obtaining the Born approximation for volumetric flaws. This
formulation, together with one for cracks, based on the Kirchhoff
approximation, are used to extract three fundamental features of the
scattering process that are shown to be useful for classification
purposes. (2) The application of the Gerchberg-Papoulis algorithm for
low frequency extrapolation procedures and the determination of the
conditions under which it provides a useful extrapolation. (3) The
development of a new flaw-derived deconvolution procedure and the
demonstration of its applicability to volumetric flaws. (4) The
successful use of our three fundamental (model-based) features for
classification using an adaptive learning approach.

The description of the details of the application of our models,
signal processing algorithms and adaptive learning methods to flaw
classification are given in the following chapters. First, however, we
will try to place these contributions in perspective by giving a summary
of previous work in these areas.

1.2 Review of the Literature

This work covers a wide variety of fields, methods and applications. Therefore, a complete literature review of all the areas from which we will draw results is impractical. Here, we will merely try to give an indication of the major, directly relevant works in the areas of scattering theory, signal processing, and adaptive learning.

1.2.1 Scattering theory

As mentioned in the introduction, one of the distinguishing aspects of the classification process we will use here is its reliance on classification features that are based on our fundamental knowledge of the scattering properties of flaws. Most of this knowledge comes from the use of elastodynamic scattering theory models. These models are usually based on integral equation formulations using either volumetric or surface integrals [8-10]. Solving such equations exactly is, of course, exceedingly difficult and usually only possible with numerical methods. Unfortunately, numerical solutions are often not particularly useful for solving classification problems, since they do not directly contain information on those features of the scattering process which are most relevant to the classification problem. However, approximate methods, such as the Born and Kirchhoff approximations [11-19], when applied to volume and surface integral equations, have been particularly useful for both classification and sizing. The Born approximation has
been traditionally derived from volumetric integral equations and the Kirchhoff from surface integrals. However, as shown in Chapter Two, a surface integral formulation for the Born approximation is also possible using the ideas developed by Schuster and Smith [20] and Schuster [21] for acoustic scattering problems. Such surface integral approximations are often preferable to volumetric-based approximations because the surface integrals are dimensionally of lower order and therefore usually easier to calculate either analytically or numerically.

The time domain versions of the Born and Kirchhoff approximations give explicit expressions for the response of flaws to an ideal unit impulse function [13,17]. Chen [22], in a recent thesis, has shown that the early arriving waves in this ideal response (leading edge response) contain significant information on flaw properties that can, in principle, be used for flaw sizing purposes. In Chapter Two, we will show that this same leading edge response also contains information directly usable for flaw classification. This is fortunate because the leading edge response is usually the most significant (largest amplitude) portion of the entire flaw signal, so that we are using that part of the flaw response that is most likely to be measurable and reliable.

Although "exact" numerical results have proven to be of little use for directly predicting general classification features, they are very useful for validating classification features and procedures that are based on approximate methods. Methods that are currently available for obtaining a variety of "exact" solutions for this purpose include the
method of separation of variables [23], T-matrix methods [24], MOOT [25], and Boundary Element Methods [26]. This work, in particular, makes use of some scattering solutions obtained via MOOT [25].

1.2.2 Signal processing

Since we are using ideal models here to define the features that are used in the classification process, it is important, as indicated in the introduction, to minimize differences between actual scattering responses and those predicted by our ideal models. In particular, loss of low frequency information due to transducer and system band limitations can produce severe distortions from an ideal response. This has been compensated for in the past through the use of the knowledge of the behavior of the ultrasonic scattering response of flaws at low frequencies [10,27,28] coupled with polynomial extrapolation methods [3]. As shown in Chapter Three, an alternative to this approach uses an extrapolation procedure due to Gerchberg [4] and Papoulis [5]. This method has seen some use previously in NDE studies [29], but with mixed success. As shown in Chapter Three, failure of the Gerchberg-Papoulis method (like all extrapolation methods, including polynomial extrapolation) can be traced to the demands placed on the method and some explicit restrictions must be placed on its application to guarantee it will work. We should note here that it is our view that such restrictions are essential to such extrapolation procedures because without them these and similar methods are ill-posed and unstable. Failure to recognize this fact can lead one to conclude, erroneously.
that such methods are totally without merit.

Another important signal processing task related to the classification process is the removal of all those features in the measured response that are not related specifically to the flaw itself. Such "contaminants" include pulser-receiver circuit characteristics, transducer diffraction effects, material attenuation effects, etc. In the past, the basic approach for eliminating these nonflaw specific features has relied on the use of linear system theory [30,31], the measurement model of Thompson and Gray [1,2], and a Wiener filter approach [32]. As described in Chapter Three, an alternative approach uses a flaw-derived reference method. To our knowledge, the use of this method has not been discussed previously in the literature for scattering problems, although Schmerr and Sieck [33] did use a similar method for a 2-D artificial void scattering problem.

1.2.3 Adaptive learning

To use a set of defined classification features, it is of course essential to have an explicit decision making process. An adaptive learning method provides such a process, one that is adaptable to effects such as noise, measurement and model errors. The basic adaptive learning approach was described in the context of the field of Artificial Intelligence (AI) by Hunt [34]. The same ideas as employed in the closely related field of pattern recognition were described by Duda and Hart [35] and summarized extensively in Sklansky and Wassel [36]. In the NDE field explicitly, the use of adaptive learning methods
for flaw classification and sizing problems was pioneered by Mucciardi and his associates [6,7]. The particular type of adaptive learning system described in Chapter Four is referred to as a linearly separable system [36]. This type of system is, in fact, just the famous perceptron algorithm [37-39] which was developed as a general model of a vision processing system in the 1960s. Although the model has been proven to be inadequate in this respect [39], it remains the oldest and best understood of the parameter adjustment models. The current intense interest in the AI field of neural networks, in fact, is a direct outgrowth of such methods [40].

There are, of course, methods other than adaptive learning that have been applied, in the AI field, to classification problems of the type we are considering here. The most important AI method currently in use in this respect is the rule-based expert system approach. Schmerr and his coworkers [41,42] in fact have been developing such a system for ultrasonic classification and sizing problems following the same approaches as pioneered in diagnostic classification systems such as MYCIN and PROSPECTOR [43]. It should be pointed out, however, that these rule-based expert systems are static, i.e., they do not currently contain learning elements. Thus, the adaptive learning approach here has some unique features that are not readily assimilated into the expert system methodologies.
2. MODELS

2.1 Introduction

As mentioned in the introduction, this work will use features for classifying flaws that are fundamental, i.e., based on analytical models that can reproduce some of the essential features of the scattering process. The two models that will be considered here are the Born approximation for volumetric flaws and the Kirchhoff approximation for cracks. Explicit expressions are derived for the scattering from an ellipsoidal inclusion and an elliptical flat crack and for the scattering features derived from these expressions. The derivation of the volumetric flaw results is based on a new surface integral Born series representation. This representation is used here to develop the first Born approximation which is, of course, identical to the results of the volumetric formulation. However, for higher order terms in the Born series, this surface formulation may offer considerable advantages over the traditional volumetric formulation since only multiple (2-D) surface integrations rather than (3-D) volumetric integrations are necessary. These advantages will not be pursued here since the first Born approximation is sufficient to define the features needed.
2.2 General Formulation

The basic scattering problem considered is the three dimensional geometrical model depicted in Fig. 2.1. The flaw region \( R^* \), bounded by a finite surface \( S^* \), is embedded in an infinite host medium \( R^0 \). The materials in both regions are assumed to be homogeneous and isotropic with elastic stiffness tensors \( C_{ijkl} \), \( C^*_{ijkl} \) and mass densities \( \rho \), \( \rho^* \) for the host and the flaw respectively. For a void, \( C^*_{ijkl} = \rho^* = 0 \).

The governing equations (in surface integral representation) for the harmonic wavefield scattered from \( S^* \) as obtained from the Betti-Rayleigh reciprocal theorem \([8]\) are:

\[
(2.1) \quad \int_{S^*} \left[ u_{ij}^G(r') \tau_{ij;m}^G(r',r) - \tau_{ij}^G(r') u_{ij;m}^G(r',r) \right] n_j^*(r') dA(r') = \begin{cases} 
  u_{m}^{\text{scatt}}(r), & r \in R^0 \\
  -u_{m}^{\text{inc}}(r), & r \in R^* 
\end{cases}
\]

The quantities \( u_{m}^{\text{scatt}} \) and \( u_{m}^{\text{inc}} \) denote the harmonic displacement fields of the scattered waves and incident waves respectively. On the boundary \( S^* \), \( u_{i}^G(r') \) and \( \tau_{ij}^G(r') \) are the displacements and stresses for the total (incident plus scattered) fields defined within \( R^0 \), and \( n_j^*(r') \) is the \( j \)th component of the outward unit normal \( \hat{n}^*(r') \). Similarly, \( u_{ij;m}^G(r',r) \) and \( \tau_{ij;m}^G(r',r) \) are the displacements and stresses associated with an unit point source (Green's function or fundamental solution) applied at the observation point \( r \in R^0 \) in the \( r_m \) direction. As can be seen in Equation (2.1) and henceforth in the rest of this paper, both the vector
Fig. 2.1. Geometrical model of a scatterer
and indicial notations are used. With the Einstein summation convention being in force, a repeated Latin subscript implies a summation with a range of one to three.

Since the boundary displacements and tractions are assumed continuous across $S^*$, i.e.,

\begin{align}
(2.2) & \quad u_i(r') = u_i^*(r') \\
(2.3) & \quad \tau_{ij}(r')n_j^*,(r') = \tau_{ij}^*(r')n_j^*,(r')
\end{align}

where quantities with asterisk are defined within $R^*$, Equations (2.1) can be expressed in terms of interior total displacements and tractions as

\begin{align}
(2.4) & \quad \int_{S^*} \left[ u_i^*(r') \tau_{ij}^G(r',r) - \tau_{ij}^*(r')u_i^G(r',r) \right] n_j^*,(r') dA(r') \\
& = \begin{cases} 
  u_m^{\text{scatt}}(r), & r \in R^0 \\
  -u_m^{\text{inc}}(r), & r \in R^*
\end{cases}
\end{align}

Either Equations (2.1) or (2.4) are called the general exterior surface integral equations.

In a similar fashion, Tan [8] derived the general interior surface integral equations from the Betti-Rayleigh reciprocal theorem

\begin{align}
(2.5) & \quad \int_{S^*} \left[ \tau_{ij}^*(r')u_{ij}^G(r',r) - \tau_{ij}^*(r')u_{ij}^G(r',r) \right] n_j^*,(r') dA(r') \\
& = \begin{cases} 
  0, & r \in R^0 \\
  u_m^*(r), & r \in R^*
\end{cases}
\end{align}
with $u_i$ and $\tau_{ij}$ being the total fields defined in $R^*$, $u_{i;m}^G(x',x)$ and $\tau_{ij;m}^G(x',x)$ being displacements and stresses due to an unit point source applied at a point $r \in R^*$. These Green's functions, $u_{i;m}^G$ and $\tau_{ij;m}^G$, share the same expressions with the Green's functions of the host except all host material properties are replaced by the corresponding properties of $R^*$.

The summation of these two surface integral formulations yields a third surface integral formulation known as the subtracted surface integral equations [20]

$$\int_{S}^* \left[ (\tau_{ij;m} - \tau_{ij;m}^G)u_i^* - (u_{i;m}^G - u_{i;m}^G)\tau_{ij}^* \right] n_j^* d\Sigma(x')$$

$$= \begin{cases} u_{m}^{\text{scatt}}(x), & r \in R^0 \\ u_{m}^{\text{scatt*}}(x), & r \in R^* \end{cases}$$

where $u_{m}^{\text{scatt*}} = u_{m}^* - u_{m}^{\text{inc}}$ in $R^*$. Unlike the first two surface integral formulations, this subtracted surface integral formulation produces a scattered displacement field both in and out of the flaw.

In the far-field approximation, i.e., $x \to \infty$ in $R^0$, with $e^{-i\omega t}$ dependence assumed, the fundamental solution of displacements $u_{i;m}^G$ is found to be [9]
where \( \alpha = \alpha \hat{r} \) and \( \beta = \beta \hat{r} \) are the longitudinal and shear wave number vectors in \( \mathbb{R}^0 \) respectively, \( \hat{r} \) is the unit vector along the \( r \) direction, \( \omega \) is the angular frequency and \( \delta_{im} \) is the Kronecker delta function. By definition, these wave numbers are

\[
(2.8a) \quad \alpha = \frac{\omega}{c_L} \\
(2.8b) \quad \beta = \frac{\omega}{c_T}
\]

and the wave speeds are given by the Lamé elastic constants \( \lambda \) and \( \mu \), and the mass density \( \rho \) of the host:

\[
(2.9a) \quad c_L = \left( \frac{\lambda + 2\mu}{\rho} \right)^{1/2} \\
(2.9b) \quad c_T = \left( \frac{\mu}{\rho} \right)^{1/2}
\]

Consequently, the far-field fundamental stresses \( \tau_{ij;m}^G \) become [9]
where $i = (-1)^{1/2}$. Since the flaw is assumed to be homogeneous and isotropic, the far-field interior fundamental solutions $u_{ij;m}^{G*}$ and $\tau_{ij;m}^{G*}$ share the same expressions given in Equations (2.7)-(2.10) except all host material properties are replaced by the corresponding properties of the flaw.

Adopting the notations of Gubernatis et al. [9], the far-field scattering displacement field can be written in terms of a $\mathbf{f}(k)$ vector

\begin{equation}
(2.11) \quad u_{m}^{\text{scatt}}(\mathbf{r}, \omega) = \hat{r}_{i} \hat{r}_{m} f_{i}(\omega) \frac{e^{i\alpha\mathbf{r}}}{r} + (\delta_{im} - \hat{r}_{i} \hat{r}_{m}) f_{i}(\beta) \frac{e^{i\beta\mathbf{r}}}{r}
\end{equation}

where the dependent factor $e^{-i\omega t}$ is understood and $\mathbf{r} \to \infty$ is a point in $\mathbb{R}^0$. This expression suggests that the far-field scattering displacement can be decoupled into two basic spherical waves: longitudinal and shear. Hence, $\hat{r}_{i} \hat{r}_{m} f_{i}(\omega)$ and $(\delta_{im} - \hat{r}_{i} \hat{r}_{m}) f_{i}(\beta)$ are called the scattered longitudinal and shear amplitude vectors respectively. From Equations (2.4), (2.7) and (2.10), the $\mathbf{f}$ vector for the far-field exterior surface integral formulation becomes [10]
(2.12) \[ f_i(k) = \frac{-k^2}{4\pi\rho^2} \left[ i k C_{ijkl} \delta_j \int_{S^*} \left( u_k^*(r', \omega)e^{-i k \cdot r'} n_i, dA(r') \right) + C_{ijkl}^* \int_{S^*} \left( u_k^*, 1 (r', \omega)e^{-i k \cdot r'} n_j, dA(r') \right) \right] \]

where \( \mathbf{k} = k \mathbf{r} \) with \( k = \alpha \) or \( \beta \) is the scattered wave number vector and the comma in the subscript denotes derivatives with respect to the (Cartesian) coordinates.

Relating the flaw material properties \((\lambda^*, \mu^*, \rho^*)\) to the host material properties \((\lambda, \mu, \rho)\) by

(2.13a) \[ \lambda^* = \lambda + \delta \lambda \]
(2.13b) \[ \mu^* = \mu + \delta \mu \]
(2.13c) \[ \rho^* = \rho + \delta \rho \]

the following approximations hold within the first order approximation of material variations

(2.14a) \[ e^{ik^* \mathbf{r}} \sim e^{ik \mathbf{r}} \]
(2.14b) \[ e^{-ik^* \cdot \mathbf{r}'} \sim e^{-ik \cdot \mathbf{r}'} \]

where \( k^* = k^* \mathbf{r} \) with \( k^* = \alpha^* \) or \( \beta^* \) is the scattered wave number vector in terms of flaw properties. Similar to Equations (2.8) and (2.9)

(2.15a) \[ \alpha^* = \frac{\omega}{c_L} \]
\[ (2.15b) \quad \beta^* = \frac{\omega}{c_T^*} \]

and

\[ (2.16a) \quad c_L^* = \left( \frac{\lambda^* + 2\mu^*}{\rho^*} \right)^{1/2} \]
\[ (2.16b) \quad c_T^* = \left( \frac{\mu^*}{\rho^*} \right)^{1/2} \]

In terms of the approximations given in Equations (2.14), the \( \mathbf{f} \) vector for the far-field interior surface integral formulation is

\[ (2.17) \quad f_1(k^*) = \frac{k^*}{4\pi\rho^*} \sum_{n_j, w_k, l'} \left[ n_j^* w_k^* l' + i k^* n_j^* r_j^* w_k^* \right] e^{-ik \cdot r'} dA(r'). \]

Subsequently, by superposition of both \( \mathbf{f} \) vectors, Equations (2.12) and (2.17), the \( \mathbf{f} \) vector for the far-field subtracted surface integral formulation is obtained.

2.3 Born Approximation

2.3.1 Frequency Domain Results

In dealing with scattering problems by means of integral equations, there are basically two different kinds of formulations: surface and volumetric. The basic equations of these approaches are interrelated. For example, applying the divergence theorem to the exterior surface
integral equations, Equations (2.4), gives the volumetric integral formulation [9]

\[
(2.18) \quad u_m(r) = u_m^{\text{inc}}(r) + \delta \rho \omega^2 \int_{R^*} u_i^* (r') u_i^G (r', r) dV(r') -
\]

\[
\delta C_{ijkl} \int_{R^*} u_{k,l}^* (r') u_i^G (r', i; m(r', r) dV(r')
\]

where \( r \in R^* \cup R^0 \) and \( \delta C_{ijkl} = C_{ijkl}^* - C_{ijkl} \). These two equations are therefore not independent. However, they do not yield the same scattering displacement field inside the flaw except when the scatterer is a void, in which case the total displacement field is zero and thus \( u_m^{\text{scatt}}(r) = - u_m^{\text{inc}}(r) \).

Equation (2.18) is analogous to the Lippmann-Schwinger equation [44] in Quantum Mechanics and can be iterated to develop an infinite perturbation series. This infinite series is known as the Born or Neumann series [45]. Executing this iteration algorithm on Equation (2.18), it becomes

\[
(2.19) \quad u_m^{(n+1)}(r) = u_m^{\text{inc}}(r) + \delta \rho \omega^2 \int_{R^*} u_i^{(n)}(r') u_i^G (r', r) dV(r') -
\]

\[
\delta C_{ijkl} \int_{R^*} u_{k,l}^{(n)} (r') u_i^G (r', i; m(r', r) dV(r')
\]

where \( n = 0, 1, 2, 3, \ldots \) and

\[
(2.20a) \quad u_i^{(0)}(r') = u_i^{\text{inc}}(r')
\]
In three dimensional elastodynamics, this Born series converges as long as the scattering field is weak inside the flaw

\[(2.21) \quad |u_{\text{scatt}}(r)| = |u(r) - u_{\text{inc}}(r)| \ll \sup_{\mathbf{r} \in \mathbb{R}} \{|u_{\text{inc}}(\mathbf{r})|\}\]

and the frequency is relatively small compared with the flaw size which, roughly, can be written as [11]

\[(2.22) \quad \left(\frac{\omega d}{c_T}\right)^2 m_p + \left(\frac{d}{\lambda}\right)^2 m_c < 1\]

where \(d\) is the length scale of \(R^*\) (e.g. the mean distance from the centroid to the boundary of \(R^*\)), \(\lambda\) is the smallest length scale \(\lambda = \min \{d, c_T/\omega\}\) and

\[(2.23a) \quad m_p = \sup_{\mathbf{r} \in R^*} \{||\delta \rho||\}/\rho\]

\[(2.23b) \quad m_c = \sup_{\mathbf{r} \in R^*} \left\{\frac{\delta C_{ijkl}}{\mu}\right\}\]

with \(i,j,k,l = 1,2,3\). The first order of this perturbation series, \(n = 0\), leads to the so-called Born approximation [12]. Physically, this approximation requires the scattering process to be weak so that the total displacement and strain fields inside the flaw can be approximated
by the corresponding incident quantities.

Unlike the volumetric formulation, the exterior surface formulation cannot be used to develop a valid Born series because of the properties of the integral operators [21]. This deficiency can be corrected by appealing to the subtracted surface integral formulation, Equations (2.6). This formulation yields an integral operator for which the Born series should be convergent under the weak scattering conditions. Although a formal proof of this convergence has not yet been made for the elastic wave case, Schuster [21] has demonstrated similar results for the scalar case.

Because the far-field Born approximation of the subtracted surface integral formulation is the summation of both the far-field Born approximations of the exterior and interior surface integral formulations, the latter two approximations will be considered first. To obtain the expression of the far-field Born approximation of the exterior surface integral equation, it is easier to apply the Born approximation to the general expression, Equation (2.4), before the far-field approximation. Following this order, the $f$ vector is found to be

$$(2.24) \quad f_i(k) = \frac{-k^2}{4\pi\rho c^2} \delta_{ijkl} \int_{S^*} u_{k,l}(r') e^{-ik \cdot r'} n_j^* (r') dA(r').$$

In the present work, incident displacement fields are considered to be harmonic plane waves with the dependence $e^{-i\omega t}$ being understood and
(2.25) \( u_1^{\text{inc}}(r', \omega) = a e^{i k^0 \cdot r'} \)

where \( e^i \) and \( e^p \) stand for the propagation and the polarization unit vectors respectively, "a" denotes the wave amplitude and the incident wave number vector \( k^0 \) is defined as

(2.26) \( k^0 = k^0 e^i \) and

(2.27) \( k^0 = \begin{cases} \alpha, & \text{for longitudinal incident waves} \\ \beta, & \text{for shear incident waves} \end{cases} \)

As a consequence of Equations (2.24) and (2.25), the scattered longitudinal and shear amplitude vectors of Equation (2.11) for the far-field Born approximation of the exterior surface integral formulation become

(2.28) \( r_m r_i f_i(\phi) = \frac{r_m \alpha^2 k^0}{4 \pi \rho^2} \left[ \delta \lambda(e^p \cdot e^i)(k^0 r_m \cdot e^i - \alpha) + \right. \\
\left. \delta \mu(e^p \cdot e^i)(k^0 r_m \cdot e^i - \alpha r_m \cdot e^i) + \delta \mu(e^p \cdot e^i)(k^0 - \alpha r_m \cdot e^i) \right] aS(k^0, \alpha) \)

(2.29) \( (\delta_{lm} - r_m r_i) f_i(\beta) = \frac{\beta^2 k^0}{4 \pi \rho^2} \left[ \delta \lambda(e^p \cdot e^i) k^0 (e_m^i - (r_m \cdot e^i) r_m) + \right. \\
\left. \delta \mu(e^p \cdot e^i)(k^0 - \beta r_m \cdot e^i) + \delta \mu(e^p \cdot e^i)(k^0 - \beta r_m \cdot e^i) \right] aS(k^0, \beta) \)
where the variation of the elastic stiffness tensor $\delta C_{ijkl}$ has been replaced by

\begin{equation}
\delta C_{ijkl} = \delta \lambda \delta_{ij} \delta_{kl} + \delta \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk})
\end{equation}

and the shape factor $S(k^O, k)$ is defined as

\begin{equation}
S(k^O, k) = \int_{\mathbb{R}^n} \epsilon^i(k^O - k', \xi) \epsilon^j dV(\xi')
\end{equation}

By applying the Born approximation and substituting

\begin{equation}
C_{ijkl}^* = \lambda^* \delta_{ij} \delta_{kl} + \mu^* (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk})
\end{equation}

into Equation (2.17), the scattered longitudinal and shear amplitude vector for the far-field Born approximation of the interior surface integral formulation are found to be

\begin{equation}
\hat{r}_m \hat{r}_i f_i(\alpha^*) = \frac{r_m \alpha^*}{4\pi} \left[ \hat{r}_m \hat{e} - \frac{k^O}{\rho} \frac{2}{\omega^2} \left\{ \lambda^* (\hat{r}_m \hat{e}) (\hat{e} \cdot \hat{e}) + \mu^* (\hat{r}_m \hat{e}) \right\} \right] aS(k^O, \alpha^*)
\end{equation}
Following Equations (2.13), the above equations can be written in terms of the first order of material variations as

\[
(2.35) \quad \hat{r}_m \hat{r}_i f_i(\hat{\alpha}^*) = \frac{-r_m \alpha^2}{4\pi} \left[ \frac{\delta\rho}{\rho} - \frac{\delta\lambda + 2\delta\mu}{\lambda + 2\mu} (\hat{r}\cdot\hat{e}) \right] \]  

\[
= \frac{-r_m \alpha^2}{4\pi} \left[ \frac{\delta\mu}{\mu} \right] \left( 1 - \frac{\hat{r}\cdot\hat{e}}{\alpha} \right) \left( \hat{r}\cdot\hat{e} \right) aS(k^0, \alpha) 
\]

By definition, the scattering longitudinal and shear amplitude vectors for the subtracted surface integral formulation result from summing up Equations (2.28), (2.29), (2.35) and (2.36)
\[ (2.37) \quad r_m \hat{r}_m f_i(\alpha) = \frac{\alpha^2}{4\pi} B_m^L a S(k^0, \alpha) \]

\[ (2.38) \quad (\delta_{im} - \hat{r}_i \hat{r}_m) f_i(\beta) = \frac{\beta^2}{4\pi} B_m^T a S(k^0, \beta) \]

where the longitudinal and shear material vectors are

\[ (2.39a) \quad B_m^L = \left[ -\frac{\delta \rho}{\rho} \hat{r} \cdot \hat{e}^P - \frac{\delta \kappa}{\rho \omega} \left( \delta \lambda (\hat{r} \cdot \hat{e}^P \cdot \hat{e}^P) + 2 \delta \mu (\hat{r} \cdot \hat{e}^i \cdot \hat{r} \cdot \hat{e}^P) \right) \right] r_m \]

\[ (2.39b) \quad B_m^T = -\frac{\delta \rho}{\rho} (\hat{e}^P_m - (\hat{r} \cdot \hat{e}^P) r_m) - \frac{\delta \mu k^0}{\mu \beta} (\hat{r} \cdot \hat{e}^P)(\hat{e}^i_m - (\hat{r} \cdot \hat{e}^i) r_m) - \frac{\delta \mu k^0}{\mu \beta} (\hat{r} \cdot \hat{e}^i)(\hat{e}^P_m - (\hat{r} \cdot \hat{e}^P) r_m) . \]

As mentioned previously, these results of the far-field Born approximation of the subtracted surface integral formulation are identical with the results of the far-field Born approximation of the volumetric integral formulation.

For an ellipsoidal flaw as is depicted in Fig. 2.2, the far-field scattering displacement field from Equation (2.11) can be described in terms of the longitudinal scattering amplitude \( A \) and shear scattering amplitudes \( B \) and \( C \) along the spherical coordinates \( \hat{r} \), \( \hat{\theta} \), and \( \hat{\phi} \) respectively

\[ (2.40) \quad u^{\text{scatt}}_i(r, \omega) = r_i A e^{i \alpha r} + \Theta_i B e^{i \beta r} + \phi_i C e^{i \beta r} . \]
Fig. 2.2. Scattering configuration for an ellipsoid
Assuming all incident waves illuminate the flaw in the positive z-direction, i.e. \( \hat{\mathbf{e}}^i = \hat{\mathbf{z}} \), for an incident longitudinal plane wave, \( \hat{\mathbf{e}}^P = \hat{\mathbf{e}}^i = \hat{\mathbf{z}} \), the scattering amplitudes are [12]

\[
(2.41a) \quad A = \frac{a\alpha^2}{4\pi} M^L_x S(\alpha^0, \alpha)
\]

\[
(2.41b) \quad B = \frac{a\beta^2}{4\pi} M^T_\Theta S(\alpha^0, \beta)
\]

(2.41c) \( C = 0 \)

where

\[
(2.42a) \quad M^L_x = \frac{\delta \rho}{\rho} \cos \theta - \frac{\delta \lambda + 2 \delta \mu \cos^2 \theta}{\lambda + 2 \mu}
\]

\[
(2.42b) \quad M^T_\Theta = \frac{\alpha \delta \mu}{\beta \mu} \sin 2 \theta - \frac{\delta \rho}{\rho}
\]

and \( \alpha^0 = \alpha \hat{\mathbf{e}}^i \) is the longitudinal incident wave number vector. For an incident shear plane wave polarized in the x-direction, \( \hat{\mathbf{e}}^P = \hat{\mathbf{x}} \), these scattering amplitudes become [12]

\[
(2.43a) \quad A = \frac{a\alpha^2}{4\pi} M^L_x S(\beta^0, \alpha)
\]

\[
(2.43b) \quad B = \frac{a\beta^2}{4\pi} M^T_\Theta S(\beta^0, \beta)
\]
\[(2.43c) \quad C = \frac{a^2}{4\pi} M_\phi^T S(\beta^0, \beta)\]

where

\[(2.44a) \quad M_r^L = \frac{\delta \rho}{\rho} \sin \Theta \cos \phi - \frac{\beta \delta \mu}{\alpha (\lambda + 2\mu)} \sin 2\Theta \cos \phi\]

\[(2.44b) \quad M_\Theta^T = \frac{\delta \rho}{\rho} \cos \Theta \cos \phi - \frac{\delta \mu}{\mu} \cos 2\Theta \cos \phi\]

\[(2.44c) \quad M_\phi^T = \frac{\delta \mu}{\mu} \cos \Theta \sin \phi - \frac{\delta \rho}{\rho} \sin \phi\]

The shape factor \(S(k^0, k)\), given in Equation (2.31), embodies all information about the shape of the scatterer. By contrast, the coefficients \(M_r^L, M_\Theta^T, M_\phi^T\) involve property changes in the flaw. For a convex flaw, the shape factor integral, Equation (2.31), can be evaluated by deforming the coordinates such that \(R^*\) turned into a sphere as demonstrated by Chiou [46] for a two dimensional case. Letting \(q\) be \((k^0 - k)\), Equation (2.31) can be written as

\[(2.45) \quad I(q) = \iiint_{-\infty}^{\infty} \Theta(\chi) e^{i \mathbf{q} \cdot \mathbf{r}} d\mathbf{y}\]

where

\[(2.46) \quad \Theta(\chi) = \begin{cases} 1, & \text{for } \chi \in R^* \\ 0, & \text{otherwise} \end{cases}\]
Then deform the coordinates $\mathbf{y}$ to a new set of coordinates $\mathbf{x}$ through a transformation matrix $A$

$$\mathbf{x} = A \mathbf{y}.$$  

This transformation Jacobian matrix $A$ is chosen to be a real-constant nonsingular matrix such that $\Theta(\mathbf{y})$ becomes axisymmetric in $\mathbf{x}$, i.e. $\Theta(\mathbf{x}) = \Theta(|\mathbf{x}|)$. Equivalently, the inhomogeneity $K^*$ is deformed into a sphere in $\mathbf{x}$. As a consequence,

$$I(q) = \det(A^{-1}) \iiint_{-\infty}^{\infty} \Theta(\mathbf{x}) e^{iK \cdot \mathbf{x}} d\mathbf{x}$$

where $K = A^{-1} \mathbf{q}$ and $\det(A^{-1})$ denotes the determinant of the inverse matrix $A^{-1}$. Choosing a set of spherical coordinates as is depicted in Fig. 2.3, then

$$I(q) = \det(A^{-1}) \int_{r=0}^{\infty} \int_{\phi=0}^{2\pi} \int_{\psi=0}^{\pi} \Theta(r) e^{iKr \cos \psi} r^2 \sin \psi d\psi d\phi dr$$

where $\Theta(\mathbf{x})$ becomes $\Theta(r)$, $K = |K|$ and

$$\Theta(r) = \begin{cases} 1, & \text{for } r \in [0, \rho] \\ 0, & \text{otherwise} \end{cases}$$

with $\rho$ being the radius of the deformed spherical flaw.

After the evaluation of the integral in Equation (2.49),
Fig. 2.3. Spherical coordinates for the shape factor
\[(2.51) \quad I(q) = \frac{4\pi \det(A^{-1})}{K} \int_0^{\rho} \sin(Kr) rdr\

= 4\pi \det(A^{-1}) \frac{\sin(K\rho) - K\rho \cos(K\rho)}{K^3}.
\]

That is,

\[(2.52) \quad S(k^0, k) = 4\pi \det(A^{-1}) \frac{\sin(K\rho) - K\rho \cos(K\rho)}{K^3}
\]

where \(K = |A^{-1}(k^0 - k)|\). For example, if \(R^*\) is an ellipsoid with semiaxes \(a_1^*, a_2^*\) and \(a_3^*\), then

\[(2.53) \quad A = \begin{pmatrix}
\frac{1}{a_1} & 0 & 0 \\
0 & \frac{1}{a_2} & 0 \\
0 & 0 & \frac{1}{a_3}
\end{pmatrix}
\]

and thus

\[(2.54) \quad A^{-1} = \begin{pmatrix}
a_1 & 0 & 0 \\
0 & a_2 & 0 \\
0 & 0 & a_3
\end{pmatrix}.
\]

Which results in a sphere of radius \(\rho = 1\), and

\[(2.55) \quad K = (a_1^2(k_1^0 - k_1)^2 + a_2^2(k_2^0 - k_2)^2 + a_3^2(k_3^0 - k_3)^2)^{1/2}
\]

\[(2.56) \quad S(k^0, k) = 4\pi a_1 a_2 a_3 \frac{\sin K - K\cos K}{K^3}
\]

which agree with the published result [12].
2.3.2 Time Domain Results

Since the scattered wave given in Equation (2.11) is harmonic, integrating the equation along all frequency components yields the following Fourier integrals

\[
\int_{-\infty}^{\infty} u_m^{\text{scatt}}(r, \omega) e^{-i\omega t} d\omega = \frac{r}{r_m^2} \int_{-\infty}^{\infty} f_i(\omega) e^{i\alpha_\omega - i\omega t} d\omega + \frac{(\delta_{im} - \frac{r}{r_m})}{r} \int_{-\infty}^{\infty} f_i(\omega) e^{i\alpha_\omega - i\omega t} d\omega
\]

and thus the corresponding scattering wave in the time domain is obtained. Let the time scattering wave be

\[
u_m^{\text{scatt}}(r, t) = \int_{-\infty}^{\infty} u_m^{\text{scatt}}(r, \omega) e^{-i\omega t} d\omega.
\]

Likewise, the time incident plane wave becomes

\[
u_m^{\text{inc}}(r', t) = \int_{-\infty}^{\infty} u_m^{\text{inc}}(r', \omega) e^{-i\omega t} d\omega
\]

\[
= e^{-p} \int_{-\infty}^{\infty} a e^{-i\omega(t - e^{-i} \cdot r'/c^0)} d\omega
\]

\[
= u_a e^{-p}
\]

where the time incident wave amplitude $u_a$ is the Fourier transform of the frequency incident wave amplitude $a$. 
(2.60) \[ u_a = \int_{-\infty}^{\infty} ae^{-i\omega(t-\frac{r^i}{c^0} \cdot r'/c^0)} d\omega \]

and the incident wave speed \( c^0 = c_L \) or \( c_T \) depending on the incident wave type. From Equation (2.11), (2.37) and (2.38), the general expression of the far-field scattering displacement field under the Born approximation in the time domain is

(2.61) \[ u_{\text{scatt}}(r,t) = \frac{R_L}{r} + \frac{R_T}{r} \]

where

(2.62a) \[ R_L = \int_{-\infty}^{\infty} \frac{r \cdot r}{R_m} f_1(\alpha) e^{i\alpha r - i\omega t} d\omega = \frac{-B_L}{4\pi c_L^2} \frac{d^2}{dt^2} \sum_{\alpha} R^* u_a dV(r') \]

(2.62b) \[ R_T = \int_{-\infty}^{\infty} \left( \delta_{im} - \frac{r \cdot r}{R_m} \right) f_1(\theta) e^{i\beta r - i\omega t} d\omega = \frac{-B_T}{4\pi c_T^2} \frac{d^2}{dt^2} \sum_{\alpha} R^* u_a dV(r') \]

with \( B^L_m \) and \( B^T_m \) given in Equations (2.39). For the special case of an impulsive incident plane wave

(2.63) \[ u_a(t-\frac{r^i}{c^0} \cdot r'/c^0) = u_o \delta(t-\frac{r^i}{c^0} \cdot r'/c^0) \]

with \( \delta(t-\frac{r^i}{c^0} \cdot r'/c^0) \) being the Dirac delta function, the above result has a particularly simple geometrical interpretation. Making this substitution and defining a normalized "time" \( s \) through


(2.64) \[ s_\gamma = \frac{c^0}{q_\gamma} (t - \frac{r'}{c_\gamma}) \]

where \( \gamma = L \) or \( T \) and \( q_\gamma = |q_\gamma|, q_\gamma = q_\gamma \hat{q}_\gamma \) and

(2.65) \[ q_\gamma = \frac{\hat{r} \cdot r'}{c_\gamma} \]

the far-field time response functions can be written as

(2.66) \[ R_\gamma \left( \frac{c^0}{q_\gamma} \right)^{\frac{1}{3}} = \frac{B}{4\pi c_\gamma^2 q_\gamma^3} \int_{R'} \delta(s_\gamma - q_\gamma \cdot \xi') dV(\xi') \]

with \( \gamma = L \) or \( T \). The integral in Equation (2.66) can be directly performed because the \( \delta \) function simply samples the volume element \( dV(\xi') \) at \( \xi = s_\gamma \) with \( \xi = \xi' \cdot q_\gamma \) as is shown in Fig. 2.4, and the volume element itself can be written as a function of \( \xi \) only as \( dV(\xi') = A(\xi) d\xi \), where \( A(\xi) \) is the cross-sectional area of the flaw in the \( q_\gamma \) direction. Thus

(2.67) \[ R_\gamma = \frac{B}{4\pi c_\gamma^2 q_\gamma^3} \frac{d^2 A(s_\gamma)}{ds_\gamma^2} \]

For a specific example, consider the area function \( A(s_\gamma) \) for an ellipsoidal flaw with semiaxes \( a_1, a_2 \) and \( a_3 \) oriented along the coordinates \( x, y \) and \( z \) respectively as is depicted in Fig. 2.5. Imagine
Fig. 2.4. Cross-sectional area $A(q_{\gamma \cdot r})$ of a flaw
Fig. 2.5. $A(s_{y})$ of an ellipsoid
an infinite plane wave traveling along its normal direction \(-\hat{q}_\gamma\). It would start intercepting the ellipsoid at the boundary point which is a distance \(h_\gamma\) away from the center \(o\). As it progresses, the intercepting area \(A(s_\gamma)\) increases to its maximum as the distance \(s_\gamma\) decreases to 0. When it leaves the center \(o\) and approaches the opposite boundary point, \(A(s_\gamma)\) decreases back to 0. Therefore, this area function \(A(s_\gamma)\) can be expressed in terms of a delta integral bounded by two Heaviside functions \(H()\) [47]

\[
A(s_\gamma) = \iint_{-\infty}^{\infty} \delta(s_\gamma - \hat{q}_\gamma \cdot \gamma) [H(\hat{q}_\gamma \cdot \gamma + h_\gamma) - H(\hat{q}_\gamma \cdot \gamma - h_\gamma)] d\gamma dV(\gamma)
\]

This expression becomes obvious when one applies the sifting property of the delta function

\[
\iint_{-\infty}^{\infty} \delta(s_\gamma - \hat{q}_\gamma \cdot \gamma) [H(\hat{q}_\gamma \cdot \gamma + h_\gamma) - H(\hat{q}_\gamma \cdot \gamma - h_\gamma)] d\gamma dV(\gamma)
\]

\[
= \int_{-\infty}^{\infty} \delta(s_\gamma - y_q) [H(y_q + h_\gamma) - H(y_q - h_\gamma)] A(y_q) dy_q
\]

\[
= A(s_\gamma) [H(s_\gamma + h_\gamma) - H(s_\gamma - h_\gamma)]
\]

where \(y_q = \hat{q}_\gamma \cdot \gamma\). By deforming the coordinates \(\gamma\) to a new set of coordinates \(x\) through a transformation Jacobian matrix \(B\)
\[
(2.70) \quad \mathbf{x} = \mathbf{B} \gamma
\]

the flaw \( R^* \) can be transformed into a sphere \( R^*_c \) centered at the origin of \( \mathbf{x} \). Equation (2.68) becomes

\[
(2.71) \quad A(s, \gamma) = \det(B^{-1}) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(s - \mathbf{B}^{-1} \mathbf{q} \cdot \mathbf{x}) \left[ H(\mathbf{B}^{-1} \mathbf{q} \cdot \mathbf{x} + h \gamma) - H(\mathbf{B}^{-1} \mathbf{q} \cdot \mathbf{x} - h \gamma) \right] dV(x)
\]

where \( \mathbf{K} = \mathbf{B}^{-1} \mathbf{q} \gamma \) and \( \mathbf{K} \) and \( \mathbf{K} \) being the magnitude and unit vector of \( \mathbf{K} \). By comparing with Equation (2.68), the above integral is merely the cross-sectional area at \( s\gamma/K \) from the center (Fig. 2.6) and this area is bounded by the two boundary points at \( r = h \gamma/K \) along the \( \mathbf{K} \) direction with \( r \) being the radius of the sphere. From the geometry in Fig. 2.6, we find
Fig. 2.6. Cutoff area of a sphere
\[
(2.72) \int_{-\infty}^{\infty} \delta(s - \frac{s}{K}) \frac{d^2}{d^2 \frac{x}{K^2}} \left[ H(\frac{s}{K} + \frac{h}{K}) - H(\frac{s}{K} - \frac{h}{K}) \right] \text{d}V(x)
\]

\[
= \pi \left( \frac{s^2}{K^2} - \frac{h^2}{K^2} \right) \left[ H(\frac{s}{K} + \frac{h}{K}) - H(\frac{s}{K} - \frac{h}{K}) \right].
\]

Therefore, from this result, Equation (2.71) becomes

\[
(2.73) \quad A(s) = \frac{\pi \det(B^{-1})}{K} \left( \frac{s^2}{K} - \frac{h^2}{K^2} \right) \left[ H\left(\frac{s}{K} + \frac{h}{K}\right) - H\left(\frac{s}{K} - \frac{h}{K}\right) \right].
\]

For an ellipsoid,

\[
(2.74) \quad B = \begin{bmatrix}
a_1 & 0 & 0 \\
0 & a_2 & 0 \\
0 & 0 & a_3
\end{bmatrix}
\]

and hence

\[
(2.75) \quad B^{-1} = \begin{bmatrix}
a_1 & 0 & 0 \\
0 & a_2 & 0 \\
0 & 0 & a_3
\end{bmatrix}
\]

and \( r = 1 = h/K \). As a result,

\[
(2.76) \quad A(s) = \frac{\pi a_1 a_2 a_3}{h \gamma} \left( 1 - \frac{s^2}{h \gamma} \right) \left[ H(s + h \gamma) - H(s - h \gamma) \right]
\]

where
(2.77) \[ h_\gamma = K = |\mathbf{B}^{-1}q_\gamma| = (a_1^2 q_{\gamma x}^2 + a_2^2 q_{\gamma y}^2 + a_3^2 q_{\gamma z}^2)^{1/2} \]

and \( q_{\gamma x} = q_{\gamma y} \) etc. Subsequently, the first and second derivatives of \( A(s_\gamma) \) are (Fig. 2.7)

(2.78) \[ \frac{dA(s_\gamma)}{ds_\gamma} = \frac{-2ma_1 a_2 a_3}{h_\gamma^3} \left[ H(s_\gamma + h_\gamma) - H(s_\gamma - h_\gamma) \right] \]

(2.79) \[ \frac{d^2 A(s_\gamma)}{ds_\gamma^2} = \frac{2ma_1 a_2 a_3}{h_\gamma^2} \left[ \delta(s_\gamma + h_\gamma) + \delta(s_\gamma - h_\gamma) \right] - \frac{2ma_1 a_2 a_3}{h_\gamma^3} \left[ H(s_\gamma + h_\gamma) - H(s_\gamma - h_\gamma) \right] \]

where \( \delta() \) is the Dirac delta function.

To demonstrate some specific examples, the scattering configuration of Fig. 2.3 is again considered. For an impulsive longitudinal incident plane wave, \( \mathbf{e}^i = \mathbf{e}^o = \hat{z} \), the longitudinal and shear time response functions are

(2.80a) \[ R_m^L \left( \frac{L}{L} \right) = \frac{-u_0 c_L H_L^L}{4\pi q_L^2} \frac{d^2 A(s_L)}{ds_L^2} r_m \]

(2.80b) \[ R_m^T \left( \frac{T}{T} \right) = \frac{-u_0 c_T H_T^T}{4\pi q_T^2 c_T^2} \frac{d^2 A(s_T)}{ds_T^2} \theta_m \]
Fig. 2.7. Area function $A(s)$ and its derivatives for an ellipsoid
where the material factors \( M^L_T \) and \( M^T_\Theta \) are given in Equations (2.42). For an impulsive shear incident plane wave polarized in the \( x \)-direction, \( \hat{e}^P = \hat{x} \), the time response functions are

\[
R^L_m = \frac{-u_0 c_T M^L_T}{4\pi c_T q_L} \frac{2A(s_L)}{ds_L^2} \quad R^T_m = \frac{-u_0 c_T (M^T_\Theta_m + M^T_\phi_m)}{4\pi q_T} \frac{2A(s_T)}{ds_T^2}
\]

where \( M^L_T \), \( M^T_\Theta \) and \( M^T_\phi \) are given in Equations (2.44).

2.4 Kirchhoff Approximation

2.4.1 Frequency Domain Results

Borrowing the idea that Kirchhoff first proposed for the field of geometrical optics, an approximation has been developed to cope with scattering from crack-like flaws in elastodynamics [14]. Basically, this Kirchhoff approximation is a method of specifying the physical properties of a crack in terms of the jumps in displacement and stress across it.

Considering only stress-free cracks, the scattering displacement field from Equations (2.1) become
Designating the crack surface which is facing the incident wave to be $S^B$ and the other side $S^D$, as is shown in Fig. 2.8, Equation (2.82) can be reduced to the following form for an ideal (zero thickness) crack [9,10]

\[ u_{\text{m}}^{\text{scatt}}(\xi, \omega) = \int_{S^B} u_i(\xi', \omega) \tau^{G}_{ij;m}(\xi', \xi) n^*_j(\xi') dA(\xi') . \]

Equation (2.83)

\[ u_{\text{m}}^{\text{scatt}}(\xi, \omega) = \int_{S^B} \Delta u_i(\xi', \omega) \tau^{G}_{ij;m}(\xi', \xi) n^*_j(\xi') dA(\xi') \]

where

\[ n^*_j(\xi') \bigg|_{\xi' \in S^B} = - n^*_j(\xi') \bigg|_{\xi' \in S^D} \]

and $\Delta u_i(\xi', \omega)$ is the displacement discontinuity between the front surface $S^B$ and the back surface $S^D$. According to the Kirchhoff approximation [14]

i) $S^B$ works as a perfect reflector and

ii) $S^D$ is assumed strictly dark,

the scattering is therefore dependent upon the front surface displacement field only

\[ u_{\text{m}}^{\text{scatt}}(\xi, \omega) = \int_{S^B} u_i(\xi', \omega) \tau^{G}_{ij;m}(\xi', \xi) n^*_j(\xi') dA(\xi') . \]

At the far-field observation point, similar to the Born approximation, the Kirchhoff approximation returns a scattering
Fig. 2.8. Crack scattering geometry
displacement field in terms of both the longitudinal and shear spherical waves

\[ u_m^{\text{scatt}}(r, \omega) = \frac{r^{-1} f_1(\delta_{im} - r_i r_m)}{r} + \frac{r^{-1} f_1(\beta) e^{i \beta r}}{r} \]

where the \( f \) vector is obtained by the substitution of Equations (2.10) into (2.85)

\[ f_1(k) = \frac{-ik^3 C_{ijkl} r_j}{4 \pi \omega^2} \int_{S^B} u_k e^{-ik \cdot r'} n_i^* (r') dA(r') \]

and \( k = k \hat{r} \) with \( k = \alpha \) or \( \beta \). If the crack surface is flat, i.e. \( n_j^*(r') = n_j \) is constant everywhere on \( S^B \), then both the longitudinal and shear scattering amplitude vectors can be simplified to

\[ r^m_{ri} f_1(\alpha) = \frac{-i \alpha n_j^{B; \alpha}}{4 \pi} \int_{S^B} u_k(\xi', \omega) e^{-ik \cdot \xi'} dA(\xi') \]

\[ (\delta_{im} - r_i r_m) f_1(\beta) = \frac{-i \beta n_j^{B; \beta}}{4 \pi} \int_{S^B} u_k(\xi', \omega) e^{-ik \cdot \xi'} dA(\xi') \]

where

\[ b_{ij;m}^{G;\alpha} = \left[ \left( 1 - \frac{2 c_T^2}{c_L^2} \right) \delta_{ij} + \frac{2 c_T^2}{c_L^2} r_i r_j \right] r_m \]

\[ b_{ij;m}^{G;\beta} = r_j \delta_{im} + r_i \delta_{jm} - 2 r_i r_j r_m \]
These equations match the results published in the literature [15-17].

Let the incident displacement field be a harmonic plane wave

\[(2.91) \quad u^{\text{inc}}(r', \omega)^{\lambda} = A e^{i \omega t} e^{i k_{\lambda} r'} \]

with dependence $e^{i \omega t}$ being understood, and where $e_{\lambda}^{i}$ and $d_{\lambda}^{i}$ are the incident propagation and polarization unit vectors respectively. The wave amplitude is $A$ and the wave number $k_{\lambda} = \alpha$ or $\beta$ depends on whether the wave type is longitudinal ($\lambda=L$) or shear-vertical ($\lambda=SV$). Since Greek indices are used to denote wave types, they do not obey the summation convention.

Employing the above Kirchhoff approximation, the total displacement therefore is [17]

\[(2.92) \quad u(r', \omega)^{\lambda} = u^{\text{inc}}(r', \omega)^{\lambda} + u^{\text{scatt}}(r', \omega)^{\lambda} = A a(\Theta_{\lambda}) e^{i \omega t} e^{i k_{\lambda} r'} \]

where the vector $a(\Theta_{\lambda})$, involving reflection coefficients, propagation and polarization unit vectors, can be found in [16,17]. Following Sedov and Schmerr [17], the vertical incident angle $\Theta_{\lambda}$ is illustrated in Fig. 2.9. By making the substitution of Equations (2.91) into (2.88) and (2.89), the far-field scattering displacement field becomes

\[(2.93) \quad u^{\text{scatt}}(r, \omega)^{\lambda} = u^{\text{sc}}(r, \omega)^{\alpha \lambda} e^{i \alpha \varphi} + u^{\text{sc}}(r, \omega)^{\beta \lambda} e^{i \beta \varphi} \]

with both the longitudinal and shear scattering amplitude vectors being
Fig. 2.9. Incident and reflected waves of a crack
(2.94) \( \{u_m^{sc}(\Sigma, \omega)\} = \frac{-\Delta a_1(\Theta, \lambda) \hat{h}_1}{4\pi} B_{ij;\lambda} G_{ij;\lambda} \)

where \( \xi = \alpha \) or \( \beta \) and the material vectors \( B_{ij;\lambda} \) are given in Equations (2.90). The function \( I^\lambda(k_{\xi}) \) has an integral

(2.95) \( I^\lambda(k_{\xi}) = ik_{\xi} \int_{S_{\mathbf{B}}} \exp(-ik_{\xi} q_{\lambda\xi} \cdot \mathbf{r}') dA(\mathbf{r}') \)

which is the two dimensional analog of the shape factor defined in Equation (2.31) and the vector \( q_{\lambda\xi} \) in the exponential is defined as [17]

(2.96) \( q_{\lambda\xi} = \mathbf{r} - \frac{c x_{\xi}}{c} \)

where there are four possible combinations of incident and reflected wave types: \( \lambda, \xi = L, SV \) and \( \mathbf{r} \) is the unit vector in the observation direction as is shown in Fig. 2.10. Details of the evaluation of the above integral in the \( I^\lambda(k_{\xi}) \) function can be found in [46] which employed the methods similar to those used in Equations (2.45)-(2.52) to compute the shape factor integral of Equation (2.31). The result is found to be

(2.97) \( \int_{S_{\mathbf{B}}} \exp(-ik_{\xi} q_{\lambda\xi} \cdot \mathbf{r}') dA(\mathbf{r}') = \frac{2\det(A^{-1})}{k_{\xi} \rho} J_1(k_{\xi} \rho) \)

where \( J_1() \) is the ordinary Bessel function of order one and \( \rho = |A^{-1} q_{\lambda\xi}| \). For an elliptical crack with semiaxes \( b_1 \) and \( b_2 \).
Fig. 2.10. Observation coordinates of crack scattering
\[ \det(A^{-1}) = b_1 b_2 \quad \text{and} \quad \psi = (b_1^2 \xi_1^2 + b_2^2 \xi_2^2)^{1/2} \quad \text{with} \quad \xi_i \quad \text{being the} \ i^{th} \ \text{component of} \ \xi \ \text{in the} \ b_i \ \text{semiaxis direction. Therefore} \]

\[ (2.98) \quad I^\lambda(k \xi) = \frac{2 \pi i b_1 b_2}{\rho} J_1(k \xi \rho) . \]

In case of normal incidence, \( \xi \xi' = 0 \), hence

\[ (2.99) \quad I^\lambda(k \xi) = \imath \kappa \xi b_1 b_2 \]

which is linearly dependent on the frequency \( \omega \).

2.4.2 Time Domain Results

To obtain the far-field scattering displacements of a crack in the time domain, Equation (2.93) must be summed over all frequency components. After making this summation, the results are

\[ (2.100) \quad u_{m}^{\text{scatt}}(r, t) = \int_{-\infty}^{\infty} \{u_{m}^{\text{scatt}}(r, \omega)\}^\lambda e^{-i \omega t} d\omega \]

\[ = \frac{R^\alpha_m}{r} + \frac{R^\beta_m}{r} \]

where

\[ (2.101) \quad R^k_m = \int_{-\infty}^{\infty} \{u_{m}^{\text{scatt}}(r, \omega)\}^k \lambda e^{i kr} e^{-i \omega t} d\omega \]
with \( k = \alpha \) or \( \beta \). Similarly, the time domain incident plane wave, Equation (2.91), becomes

\[
(2.102) \quad u^{\text{inc}}(r',t) = \int_{-\infty}^{\infty} (u^{\text{inc}}(r',\omega)) e^{-i\omega t} d\omega
\]

\[
= u_a \delta(t)
\]

with the amplitude \( u_a \) being the Fourier transform of the frequency amplitude \( A(\omega) \)

\[
(2.103) \quad u_a = \int_{-\infty}^{\infty} A(\omega) \exp(-i\omega(t - \frac{\Delta r}{c})/\omega) d\omega.
\]

From Equations (2.94), (2.101) and (2.103)

\[
(2.104) \quad R^\xi_m = \frac{a(\xi,\alpha) c^2 G_i \xi}{4 \pi c \xi} \int_{-\infty}^{\infty} u_a \left(t + \frac{q \lambda \xi \cdot r' - r}{c \xi}\right) dA(r')
\]

where \( \xi = \alpha \) or \( \beta \), \( c_\alpha \) and \( c_\beta \) are equivalent to \( c_L \) and \( c_T \) respectively. Particularly, if the incident displacement is an impulsive plane wave

\[
(2.105) \quad u_a(t + \frac{q \lambda \xi \cdot r' - r}{c \xi}) = u_0 \delta(t + \frac{q \lambda \xi \cdot r' - r}{c \xi})
\]

then the above integral will have a simple geometric interpretation.

Making this substitution and defining a normalized "time" \( s_{\lambda \xi} \) through

\[
(2.106) \quad s_{\lambda \xi} = \frac{c \xi}{q \lambda \xi} (t - \frac{r}{c \xi})
\]
with \( q_{\lambda \xi} = |q_{\lambda \xi}| \) and \( \dot{q}_{\lambda \xi} = q_{\lambda \xi} \dot{q}_{\lambda \xi} \), the above far-field time response functions become

\[
R_m^\xi = \frac{a_i(\Theta)_{ijm} G_{ij} B_{ijm} c_{ij} \xi d}{4 \pi q_{\lambda \xi} S L} \int_{S L} \delta(s \lambda \xi - \dot{q}_{\lambda \xi} \xi') dA(\xi') .
\]

Note that for normal incidence \( q_{\lambda \xi} \) is perpendicular to the crack surface or \( \dot{q}_{\lambda \xi} \xi' = 0 \) and hence these time response functions behave like the derivative of the Dirac delta function. For incidences other than normal to the crack surface, the integral in Equation (2.107) can be evaluated to be

\[
\int_{S L} \delta(s \lambda \xi - \dot{q}_{\lambda \xi} \xi') dA(\xi') = \frac{D_{\lambda \xi}(s \lambda \xi/\cosh)}{|\cosh|}
\]

where \( \eta \) is the angle between \( q_{\lambda \xi} \) and its projection onto the crack surface as is shown in Fig. 2.11 and \( D_{\lambda \xi} \) is the cross-sectional length of the crack at \( \dot{q}_{\lambda \xi} \xi' \) and perpendicular to the projection of \( q_{\lambda \xi} \). As a result,

\[
R_m^\xi = \frac{u a_i(\Theta)_{ijm} G_{ij} B_{ijm} c_{ij} \xi d}{4 \pi |\cosh| q_{\lambda \xi}^2} D_{\lambda \xi}(s \lambda \xi/\cosh) .
\]

Specifically, for an elliptical crack with semiaxes \( b_1 \) and \( b_2 \) and which is oriented with an angle \( \gamma \) away from the \( \psi \)-axis in Fig. 2.11 as is shown in Fig. 2.12, the distance function \( D_{\lambda \xi} \) and its derivative are
Fig. 2.11. Cross-sectional length of a crack
Fig. 2.12. Orientation of an elliptical crack
found to be

\[
\begin{align*}
D_{\lambda\xi}(s_{\lambda\xi}/\cosh) &= \frac{2b_1b_2[(b_1^2\cos^2\gamma + b_2^2\sin^2\gamma) - (s_{\lambda\xi}/\cosh)^2]^{1/2}}{b_1^2\cos^2\gamma + b_2^2\sin^2\gamma} \\
\frac{dD_{\lambda\xi}(s_{\lambda\xi}/\cosh)}{ds_{\lambda\xi}} &= \frac{-2b_1b_2s_{\lambda\xi}/\cosh^2\eta}{(b_1^2\cos^2\gamma + b_2^2\sin^2\gamma)[(b_1^2\cos^2\gamma + b_2^2\sin^2\gamma) - (s_{\lambda\xi}/\cosh)^2]^{1/2}}
\end{align*}
\]

where \(|s_{\lambda\xi}/\cosh| < (b_1^2\cos^2\gamma + b_2^2\sin^2\gamma)^{1/2}\), otherwise both \(D_{\lambda\xi}\) and its derivative are zero. As is shown in Fig. 2.13, the derivative of \(D_{\lambda\xi}\) is axisymmetric with square root singularities at both the leading and the trailing edges of the response. These singularities are the "flash point" contributions, obtained from the geometrical ray theory [18,19], and they are the distinct features of the scattering response from a crack.

2.5 Leading Edge Response

As can be seen in Equations (2.67) and (2.79), the impulse response signal of a volumetric flaw modeled by the Born approximation consists of a leading delta function (the left-most spike in Fig. 2.14a). On the other hand, for a flat crack modeled by the Kirchhoff approximation, the early time response has a square root singularity (Fig. 2.14c) as shown in Equations (2.109) and (2.111). In Fig. 2.14, these models are compared with the bandlimited "exact" numerical results obtained by the
Fig. 2.13. $D_{\lambda \xi}$ and its derivative
Fig. 2.14. Comparisons among Born, Kirchhoff and the "exact" results
Method of Optimal Truncation (MOOT) [25]. If only the early time response are being considered, both the Born and Kirchhoff models agree very well with the MOOT results. However, when looking at the complete signal, both Born and Kirchhoff show obvious discrepancies in the late response. As is shown in Fig. 2.14, the exact result for a volumetric flaw does not show a flat plateau as is predicted by Born. Moreover, instead of a second delta function reflected by the back edge as is seen in Born, MOOT exhibits a creeping wave. For crack signals, the Kirchhoff approximation and the exact result both predict a pair of flash points. Nevertheless the Kirchhoff waveform misses the later arriving Rayleigh wave which is seen as the "tail" of the exact result.

Even though the Born approximation is a long-wavelength, weak-scattering theory, it has also been successfully used for sizing strong volumetric scatterers like voids [48]. Chen [22], using elastodynamic ray theory, recently gave some exact results for voids that may help to explain this unexpected success of the Born approximation.

Recall that the Born response for a volumetric flaw is (see Equation (2.67))

\[ (2.112) \quad R_m^E = \text{const} \frac{d^2A(s)}{ds^2} \]

where
with $\xi = \alpha$ or $\beta$ being the reflective wave type. By shifting the zero of time to the front pulse, i.e., replacing $s$ by $(s - h_\gamma)$, Equation (2.79) can be written as

\[
\begin{align*}
2na_1a_2a_3 & = -B_m c_0^3 u_0 \frac{2na_1a_2a_3}{4\pi c_0^3 \xi^3} \\
& \frac{d^2A(s)}{ds^2} = \frac{2na_1a_2a_3}{h_\gamma^2} \left[ \delta(s) + \delta(s - 2h_\gamma) \right] - \\
& \frac{2na_1a_2a_3}{h_\gamma^3} \left[ H(s) - H(s - 2h_\gamma) \right].
\end{align*}
\]

After this shifting, the signal starts at $s = 0$ as is depicted in Fig. 2.15. Defining the leading edge response function $L_m^\xi(s)$ to be the integration of this shifted signal across the early time response, then

\[
L_m^\xi(s) = \int_0^s R_m^\xi ds = \text{const.} \frac{dA(s)}{ds} \bigg|_0^s = c_0 + c_1 s
\]

where $c_0$ and $c_1$ are constants. That is, based on the Born approximation, the leading edge response function of a volumetric flaw behaves like a step and a linear term as is shown in Fig. 2.15 in the early time response. Notice that the polarity of a Born signal depends on the sign (either positive or negative) of the constant in Equation (2.113). For volumetric inclusions with acoustic impedance relatively lower than the host material, including voids, the Born signal predicts
(a) Born (low impedance inclusion) (c) Kirchhoff (crack)

(b) Born (high impedance inclusion)

Fig. 2.15. Leading edge responses for Born and Kirchhoff signals
both delta functions in the response must be negative. On the other hand, for inclusions with higher impedance, both delta functions will be positive.

For a crack, recall from Equation (2.109) that the impulse response based on the Kirchhoff approximation is

\[
R_m^\xi(s) = \text{const.} \frac{dD_{\lambda\xi}(s)}{ds} \cosh(\theta \lambda)
\]

(2.116)

where

\[
\text{const} = \frac{u_0 a_i(\theta \lambda) n_i c_{\xi} \xi_{ij}}{4\pi \cos \theta |q_{\lambda\xi}^2|}
\]

(2.117)

where \( \xi = \alpha \) or \( \beta \) being the reflective wave type. With a shifting of the zero of time to the beginning of the signal as is shown in Fig. 2.15, Equation (2.111) becomes

\[
\frac{dD_{\lambda\xi}(s)}{ds} = \frac{b_1 b_2}{\Phi^2 \cosh(\theta \lambda)} \left( \frac{s}{\Phi \cosh(\theta \lambda)} \right)^{1/2} \left[ 1 - \left( \frac{s}{2 \Phi \cosh(\theta \lambda)} \right) \right]^{1/2}
\]

(2.118)

where

\[
\Phi = \left( b_1^2 \cos^2 \gamma + b_2^2 \sin^2 \gamma \right)^{1/2}
\]

(2.119)

Similarly, integrating this shifted signal across the early time response yields the following leading edge response function
(2.120) \[ L_m^E(s) = \left. \int_0^s R_m^E ds = \text{const.} D_{\lambda E}(s/\cosh) \right|_{s=0}^s = c_{1/2}s^{1/2} + c_{3/2}s^{3/2} + o(s^{5/2}) \]

where \( c_{1/2} \) and \( c_{3/2} \) are constant coefficients and \( o(s^{5/2}) \) denotes the rest of terms have an order of at least \( s^{5/2} \).

Since the leading edge response function \( L_m^E(s) \) differs very distinctively for flaw signals based on the Born or Kirchhoff approximation, one can, in principle, key on those features (differences) to do flaw classification. In view of Fig. 2.15, there are clearly two features in the leading edge response function,

i) \( F_1 \): the value at the beginning of the leading edge response function, i.e., \( s = 0 \),

ii) \( F_2 \): the slope right after the beginning of the leading edge response function.

For a volumetric flaw with low (high) acoustic impedance, \( F_1 \) is negative (positive) but not zero and \( F_2 \) is positive (negative), while for a crack, \( F_1 \) is zero and \( F_2 \) is always negative.

For an ideal (infinite bandwidth) system, \( F_1 \) and \( F_2 \) would always suffice to classify the flaw as crack-like or volumetric. However, as is shown in the next chapter, finite bandwidth effects can, under certain circumstances render \( F_2 \) unreliable. Thus, it is necessary, in those cases to replace \( F_2 \) by a different feature \( F_3 \). We have chosen to define \( F_3 \) as the absolute ratio between the amplitude of the front edge response and the amplitude of its immediately following response, of
opposite polarity (sign) (as will be defined more precisely in the subsequent chapters), in the time response function \( R_m^E(s) \). For a volumetric flaw, we expect

\[
\text{(2.121)} \quad F_3 = \frac{\text{amplitude of the step response}}{\text{amplitude of the front edge response}} = \left| \frac{A_H}{A_L} \right|
\]

since the leading edge response contains both a Dirac delta function and step contributions (Fig. 2.16a). For a crack, we either expect a value of zero for \( F_3 \) (Fig. 2.16b) or, if the crack were small or viewed at near normal incidence, a value of 1 (Fig. 2.16c) for the ideal model signals. Thus, as long as the \( F_3 \) value for the volumetric case is not too close to either one or zero, this will be a useful distinguishing feature. This feature, \( F_3 \), might seem very ridiculous at first glance since it depends on the strength of either a Dirac delta function or a square root singularity and both of them are infinite. However, this feature will prove to be useful when \( F_3 \) is defined in terms of the responses of a bandlimited system as can be seen later in Chapters Three and Four.

Before obtaining an explicit expression for \( F_3 \) for a volumetric flaw from Equations (2.114), one must adjust the coefficient of the front pulse (delta function) response so that it agrees with the coefficient of the step response in dimension (this is necessary since the strength of the delta function itself is not dimensionless). Substituting Equations (2.64) into (2.114), the front pulse response term becomes
(a) Volumetric flaw (Born)

(b) Large crack (Kirchhoff)

(c) Small crack (Kirchhoff)

Fig. 2.16. Ideal time response for model signals
which now has a coefficient with a dimension that matches the Heaviside step function in Equation (2.114). Therefore for a Born signal

\[
F_3 = \frac{\text{const.} \frac{2n_a a_2 a_3}{h_Y^2}}{\text{const.} \left( \frac{2n_a a_2 a_3}{h_Y^3} \right)} = \left| \frac{c_0}{q_Y h_Y} \right|.
\]

This result, without the absolute sign, agrees with the first term found in Chen's thesis [22] for the ratio between the "step" and the front delta response. Chen, actually employed elastodynamic ray theory to obtain two terms in the numerator of the expression for \( F_3 \) which rendered it exact. He showed however, that this "second term" would vanish if the host material has a Poisson's ratio equal to 1/3. In practice, most materials under investigation have a Poisson's ratio very close to 1/3 and because of that, the above result, Equation (2.123), is also close to "exact" for Born signals with unlimited bandwidth.

In case of a pulse-echo setup with longitudinal incidence only, note that \( F_3 \) becomes simply

\[
F_3 = \frac{c_L}{2h_Y}.
\]
3. SIGNAL PROCESSING AND FEATURE IDENTIFICATION

3.1 Introduction

In measuring ultrasonic signals, acoustic transducers are employed as senders and receivers. All acoustic transducers, however, are sensitive only within a finite frequency range. Subsequently, all ultrasonic measurements are bandlimited. This bandlimitation changes the form of a signal including its leading edge features.

The loss of high frequencies causes the existence of finite rise times for the pulses in the leading edge response of the flaw. The first section of this chapter, therefore, examines the effects of such rise times on our definition and use of features in the leading edge response.

On the other hand, loss of low frequencies causes our signals to sit on a wavy base line and induces considerable nonideal "structure". This corruption also can severely affect our ability to distinguish between a volumetric flaw and a crack. If the loss of the low frequency end is "small" enough, one can recover this low frequency portion of the signal by employing extrapolation techniques. In a later section, both the extrapolation by iteration and extrapolation by polynomial fit will be examined.

An experimental signal response is always contaminated with noise (both acoustic and electronic) and system effects that are not flaw related. As mentioned in the introduction, the process which tries to remove all nonflaw dependent features except noisiness in the signal
response is called "deconvolution". Here, two different deconvolutions will be discussed and compared. One way to deconvolve a signal is through the measurement model developed by Thompson and Gray [1,2] which uses basically a reference signal taken from a flawless sample. A new type of deconvolution considered by the present work is a self-derived reference which employs a reference signal extracted from the flaw signal itself in conjunction with a Wiener filter [32]. This latter approach is particularly attractive because it does not require the use of a separate reference signal or the calculation of diffraction and attenuation corrections.

3.2 Finite Bandwidth Effects on Ideal Features

In Chapter 2, the leading edge response was found for both volumetric flaws and cracks. For a volumetric flaw, a Dirac delta function is expected in the leading edge response. On the other hand, a crack would return a square root singularity in the front edge response. Although these features are obviously different, they are also based on an assumption that there is no bandlimitation in the frequency domain. For a realistic measurement, finite bandwidth is always inevitable.

To see the effects of finite bandwidth, the bandlimited version of the delta function will be examined in the following. By definition, the spectral components of a delta function in the time domain is merely a constant (unity) function, i.e.,
(3.1) \[ \delta(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} d\omega. \]

If this constant function has only nonzero values inside the bandwidth 
0 \leq \omega_0 \leq |\omega| \leq \omega_1, then, the above integral becomes [49]

\[ (3.2) \left( \int_{-\omega_1}^{-\omega_0} + \int_{\omega_0}^{\omega_1} \right) \frac{e^{-i\omega t}}{2\pi} d\omega = \frac{\omega_1}{\pi} \sin(\omega_1 t) = \frac{\omega_0}{\pi} \sin(\omega_0 t) \]

\[ = \frac{1}{\pi} \cos \left( \frac{\omega_1 + \omega_0}{2}t \right) \sin \left( \frac{\omega_1 - \omega_0}{2}t \right) \]

which is a cosine function modulated by a sinc function. Notice that
the above integration must include negative frequency components in
order to guarantee a real time function. If only the high frequency end
is lost, i.e., \( \omega_0 = 0 \) (Fig. 3.1a) then, the bandlimited delta function
\( b(t, \omega_1) \) becomes

\[ (3.3) b(t, \omega_1) = \frac{\omega_1}{\pi} \frac{\sin(\omega_1 t)}{\omega_1 t} \]

which is a sinc function as is shown in Fig. 3.1b. Note that the number
of significant figures shown in these graphs (and many subsequent ones)
is an internal feature of the Symbolics 3670 LISP-machine used for these
calculations. In many cases, data displayed (particularly experimental
(a) a step function with bandwidth 0-20 MHz in the frequency domain

(b) a sinc function with bandwidth 0-20 MHz in the time domain

(c) integration of the sinc function

Fig. 3.1. A bandlimited Dirac delta function
data) will actually contain far fewer "true" significant figures. This sinc function has a maximum at \( t = 0 \) and the maximum value is equal to \( \omega_1 / \pi \). Its zeros nearest the origin are at \( \omega_1 t = \pi \) and \( -\pi \) or \( t = \pi / \omega_1 \) and \( -\pi / \omega_1 \). The time difference between these two zeros, \( 2\pi / \omega_1 \), becomes the "rise time" in the integrated time domain as is depicted in Fig. 3.1c. With \( \omega_1 \to \infty \), this rise time is zero and we have the case in an ideal Born signal. For a finite high frequency \( \omega_1 \), the rise time \( \Delta t_{\text{Born}} \) of a Born signal becomes

\[
(3.4) \quad \Delta t_{\text{Born}} = \frac{2\pi}{\omega_1}.
\]

A comparison with the examples shown in Figs. 3.2 and 3.4 for \( f_1 = \omega_1 / 2\pi = 20 \text{ MHz} \) (megahertz) shows good agreement for the rise time of the integrated time domain response and Equation \( (3.4) \). On the other hand, the rise time in a Kirchhoff signal (which is by definition the time measurement between the beginning and the minimum in the integrated time signal) does not change very much from the infinite bandwidth results when the high frequencies are lost (see Figs. 3.3 and 3.5).

In view of the above results, it is necessary to redefine the first and second features, \( F_1 \) and \( F_2 \), which were introduced in the last section of Chapter 2 to incorporate the fact that the high frequency end of a real measurement is always finite. Since the Born rise time is a function of the maximum frequency while the Kirchhoff rise time is not significantly affected, the following new definitions are necessary.

1. \( F_1 \): the value measured at the Born rise time (for all
Fig. 3.2. Born signals for a 100 µm spherical void
(a) time domain

(b) integrated time domain

Fig. 3.3. Kirchhoff signals for a 100 \( \mu \text{m} \) (effective) circular crack
(a) time domain

(b) integrated time domain

Fig. 3.4. Born signals for a 400 µm spherical void
(a) time domain

![Graph of Kirchhoff signals in time domain]

(b) integrated time domain

![Graph of Kirchhoff signals in integrated time domain]

Fig. 3.5. Kirchhoff signals for a 400 μm (effective) circular crack
responses).

2. **F2**: the slope measured immediately after the Born rise time (for all responses).

Subsequently, for a volumetric flaw with low (high) acoustic impedance, \( F_1 \) is negative (positive) and \( F_2 \) is positive (negative), whereas for a crack with a rise time bigger than the Born rise time, both \( F_1 \) and \( F_2 \) are always negative. Whenever the effective crack size is so small that the Kirchhoff rise time (\( \Delta t_{\text{Kirchhoff}} \)) is less than the Born rise time (\( \Delta t_{\text{Born}} \)), \( F_2 \) will be unable to distinguish the flaw type. With an upper frequency bound \( \omega_1 \), the smallest effective crack radius \( a_e \) where classification is theoretically possible with these features can be calculated by equating the Born and Kirchhoff rise times,

\[
\Delta t_{\text{Born}} = \Delta t_{\text{Kirchhoff}}.
\]

To find an expression for this limiting \( a_e \), consider a typical Kirchhoff signal for a small crack with a finite maximum frequency as shown in Fig. 3.3. To estimate the time width between the two zeros of the front pulse in Fig. 3.3a, the two sections, from the front zero to the pulse tip and from the pulse tip to the second zero will be examined separately. Comparing with a bandlimited delta function of a Born signal as is shown in either Fig. 3.2a or 3.4a, the front half of the first pulse in the Kirchhoff signal (Fig. 3.3a) exhibits great similarity. Therefore, the time width of this front half is taken to be \( \Delta t_{\text{Born}}/2 \). The second half of the Kirchhoff front pulse is merely the reflection of the first half of the crack, therefore, the traveling time
it takes is \( \frac{2a_e}{c} \) where \( c \) is the phase velocity in the host material.

Summing up these results, the Kirchhoff rise time is

\[
\Delta t_{\text{Kirchhoff}} = \frac{\Delta t_{\text{Born}}}{2} + \frac{2a_e}{c}.
\]

Hence, from Equations (3.4), (3.5) and (3.6)

\[
\frac{\pi}{\omega_l} = \frac{2a_e}{c}
\]

\[
a_e = \frac{\pi c}{2\omega_l}.
\]

For example, if the host is titanium alloy, which has a longitudinal phase velocity to be \( 6.07 \times 10^3 \) m/sec, together with a maximum frequency being 20 MHz, this effective radius is found to be

\[
a_e = 76.0 \ \mu m
\]

where \( 10^6 \ \mu m \) (micron) = 1 m. It should be noted that the effective radius \( a_e \) depends not only on the geometry but also on the look angle. This dependency can be seen in Fig. 2.5 and Equation (2.77) for a volumetric flaw. For a crack, the "distance" between the two flash points \( 2a_e \) as being shown in Fig. 3.6 is

\[
2a_e = 2ac \cos \theta
\]

where \( \theta \) is the look angle measured from the crack surface. In case of a
Fig. 3.6. Effective radius for a flat crack
pulse-echo setting and the incident wave is normal to the crack, i.e., \( \theta = 90^\circ \), then \( a_e \rightarrow 0 \). This would move both flash points together such that one "sits" on the top of the other. In view of Equation (2.107), this response merely demonstrates the behavior of the derivative of the Dirac delta function.

For small cracks or cracks which are measured at an angle at or near normal incidence, it therefore is possible for \( \Delta t_{\text{Kirchhoff}} \) to be less than or equal to \( \Delta t_{\text{Born}} \). When this occurs, the slope feature \( F_2 \) will not be a reliable indicator of flaw type and it is necessary to couple \( F_1 \) with a different feature such as feature \( F_3 \) described in the previous chapter. Before this feature can be used realistically, however, it also must be modified to account for finite bandwidth effects. First, consider a volumetric flaw. In this case, the loss of high frequency components will not only replace the delta function by a sinc function but also the step function by its bandlimited version. To accommodate the changes in a signal due to finite bandwidth, the expression of the ratio feature \( F_3 \) introduced in the last chapter must be modified by appropriate factors.

Since the bandlimited delta function is a sinc function with a maximum value \( \omega_1/\pi \) at \( t = 0 \) as can be seen in Equation (3.3), Equation (2.123) becomes

\[
F_3 = \frac{c^0}{|q_\gamma h_\gamma \omega_1|} \pi
\]

in order to incorporate the change in the delta function.
To find out the effect due to the change in the step function, the following analysis will be considered. Assuming \( f(t) \) is a unit step function in the time domain with a span from \(-t_o/2\) to \(t_o/2\) as is illustrated in Fig. 3.7a.

\[
(3.13) \quad f(t) = \begin{cases} 
1, & t \in [-t_o/2, t_o/2] \\
0, & \text{otherwise}
\end{cases}
\]

the Fourier transform of this function is just a sinc function (Fig. 3.7b),

\[
(3.14) \quad F(\omega) = \int_{-\infty}^{\infty} f(t)e^{i\omega t} \, dt \\
= t_o \cdot \frac{\sin(\omega t_o/2)}{\omega t_o/2}
\]
as is expected. Bandlimiting it with a frequency range \([0, \omega_1]\) and then performing the inverse Fourier transform, the bandlimited version of \( f(t) \) can be obtained

\[
(3.15) \quad f_b(t) = \frac{1}{2\pi} \int_{-\omega_1}^{\omega_1} F(\omega)e^{-i\omega t} \, d\omega \\
= \frac{1}{\pi} \left\{ \text{Si}[\omega_1(t + t_o/2)] + \text{Si}[\omega_1(t_o/2 - t)] \right\}
\]

where \( \text{Si}[\cdot] \) is the Sine integral defined by [50]

\[
(3.16) \quad \text{Si}(x) = \int_{0}^{x} \frac{\sin(t)}{t} \, dt.
\]
(a) time domain (infinite bandwidth)

\[ f(t) \]

(b) frequency domain (0-20 MHz)

(c) Sine integral function (integral of (b))

Fig. 3.7. A bandlimited step function
By the substitution $t' = (t + t_o/2)$, $f_b(t)$ becomes

$$f_b(t) = \frac{1}{\pi} \left( \text{Si}[\omega_1 t'] + \text{Si}[\omega_1(t_o - t')] \right).$$

For a large $t_o$ and a fixed $t'$, $\text{Si}[\omega_1 t']$ converges to $\pi/2$ [50], therefore

$$f_b(t) = \frac{1}{\pi} \left( \frac{\pi}{2} + \text{Si}[\omega_1(t_o - t')] \right).$$

As for $\text{Si}[\omega_1(t_o - t')]$, it has a value of approximately $1.18\pi/2$ at the first maximum as can be seen in Fig. 3.7c. Hence,

$$f_b(t) \bigg|_{\text{1st max}} = \frac{1}{\pi} \left( \frac{\pi}{2} + \frac{1.18\pi}{2} \right) = 1.09.$$

Integrating this result into the ratio feature $F3$, Equation (3.12) becomes

$$F3 = \frac{1.09\pi c^0}{|q| h_\gamma \omega_1}$$

which reflects the ratio between the amplitude of the immediate opposite peak after the front pulse and the amplitude of the front pulse.

For small cracks, based on Kirchhoff model, this ratio is expected to be 1 because of the fact that the second flash point is the immediate opposite peak after the front pulse (Fig. 3.3a). As for large cracks, $F3$ is expected to be either less than or equal to zero since there are
local "wiggles" between the flash points due to the loss of high frequency components as is shown in Fig. 3.5a.

To see how accurate this ratio, Equation (3.20), predicts F3 for volumetric flaw signals, the following parameters are considered

\[ c^0 = 6.07 \times 10^3 \text{ m/sec (longitudinal wave speed in titanium alloy)}, \]

\[ |q^1| = 2 \text{ (a pulse-echo setting)}, \]

\[ f_1 = \omega_1/2\pi = 20 \text{ MHz} \]

and F3 then becomes

\[ F3 = \frac{82.57 \mu m}{h} \]

This equation is then tabulated in Table 3.1 to compare with other measurements. The results computed from the equation is called "theory" in the table.

Significant discrepancies are noticed between Equation (3.21) and the other values in the table. The source of these discrepancies can be seen from Fig. 3.4. The step function part of the amplitude ratio in F3 was obtained from a measurement of the peak immediately following the bandlimited leading edge delta function in Fig. 3.4a. However, this step function value is also the slope just following the integrated delta function (step) in Fig. 3.4b. In an ideal infinite bandwidth system this slope would coincide with the value obtained from a straight line drawn from the leading edge step response peak to the origin. In Fig. 3.4b, we see that high frequency "wiggles" distort this straight line and cause F3 to have a large overestimated value. Since it is
Table 3.1. Comparisons of F3 for volumetric flaws in titanium alloy

<table>
<thead>
<tr>
<th>$h_Y$ (µm)</th>
<th>F3</th>
<th>theory</th>
<th>Born</th>
<th>MOOT</th>
<th>experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>75</td>
<td>1.10</td>
<td>1.74</td>
<td></td>
<td></td>
<td>0.64</td>
</tr>
<tr>
<td>100</td>
<td>0.83</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>0.41</td>
<td>0.35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>239</td>
<td>0.34</td>
<td>0.66</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>242</td>
<td>0.28</td>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>0.28</td>
<td>0.64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>310</td>
<td>0.27</td>
<td>0.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>397</td>
<td>0.21</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>0.21</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>401</td>
<td>0.21</td>
<td>0.55</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>450</td>
<td>0.18</td>
<td>0.52</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>800</td>
<td>0.10</td>
<td>0.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>0.08</td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1200</td>
<td>0.07</td>
<td>0.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1400</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
difficult in general to compensate analytically for such differences in our expression for $F_3$, the question is can $F_3$ still be used reliably as a feature? Recall we are using the difference in $F_3$ from the values of zero or one to be an indicator of a volumetric flaw. As Table 3.1 shows, even with such discrepancies $F_3$ does appear to display such difference except for very small or large flaws. Thus, $F_3$ will be retained for the present work.

Besides losing high frequency components, a real signal measurement also misses some low frequency components. In the case of a Dirac delta function as is displayed in Equation (3.2), the loss of low frequency end results in a cosine modulation of the function in the time domain. Similar effects are found in other signals. To illustrate this phenomenon, a Born signal without the low frequency end is shown in Fig. 3.8. As can be seen in Fig. 3.8, the base line of the signal oscillates because of the lack of low frequency components. A similar modulation is thus expected also in the integrated time signal and therefore all feature measurements would be corrupted. In some cases, such modulation can be removed by applying an appropriate extrapolation scheme at the low frequency end. Two different extrapolation algorithms will be examined in the next section.
(a) a Born signal in time domain with bandwidth 0-20 MHz

(b) a Born signal in time domain with bandwidth 2-20 MHz

Fig. 3.8. Effect of missing low frequency components
3.3 Improvement Due to Extrapolation

Due to the finite bandwidth of acoustic transducers and beam spreading effects, all real signal measurements are expected to miss some low frequency components and probably most of the high frequency components. Finding ways to reconstruct the low frequency end of a signal is the main concern of this section. As is shown here, the restoration of signals below the low frequency limit can be accomplished in some cases. There exist very fundamental mathematical reasons why such restoration beyond the low limit should be possible. These reasons rest on two basic mathematical principles which are listed here as theorems. The proof of these principles can be found in the book written by Guillemin [51].

Theorem 3.1 The Fourier transform of a spatially bounded function is an entire function, i.e., analytic throughout the entire frequency domain.

Theorem 3.2 If any analytic function in the frequency plane is known exactly in an arbitrarily small (but finite) region of that plane, then the entire function can be found (uniquely) by means of analytic continuation.

It should be pointed out, however, that although these theorems suggest it is possible to recover the low frequency portion of the signal, they do not, unfortunately, provide a means for accomplishing this purpose. Since analytic continuation is inherently an ill-posed problem, the development of such means is no small task.

To put our problem in perspective, consider an impulsive plane wave
scattering off an isolated, interior flaw with finite geometry. This scattered signal, theoretically, consists of an infinite chain of various waves: longitudinal, shear and surface radiation which can, for example, include creeping waves for a volumetric flaw [52] or Rayleigh waves for a crack [53,54]. In practice, however, all measurements are limited to a finite span of time. Even with such a limitation, a scattered signal still retains considerable complexity as can be seen, for example, in experimental results [55–58]. By time gating and controlling the orientation of the receiving transducer, one can always choose to reduce some of this complexity and receive only finite portions of the total scattered response. Even though this bounded measurement is not complete, it can still contain a significant amount of physical information about a flaw. If we can assume this finite measurement to be the scattered signal, then by Theorem 3.1, its Fourier transform is an entire function. Because of the finite bandwidth limitation, only a finite section of this spectral function is reliable. However, based on Theorem 3.2, resolution below the lower frequency limit can be restored by means of analytic continuation.

Two different ways of doing the analytic continuation using extrapolation techniques will now be examined.

3.3.1 Iteration

One way to do extrapolation is by means of an iterative, successive approximation approach as proposed by Gerchberg [4] and Papoulis [5]. To demonstrate how their algorithm works, a spectral function $G(f)$ of a
finite flaw is assumed to have reliable values within the frequency range \((f_0, f_1)\) as is depicted in Fig. 3.9a. Then \(G(f)\) is transformed to the spatial domain and becomes \(g(t)\). Since \(g(t)\) must be bounded, all values which are beyond its span \((t_0, t_1)\) are set to zero as is shown in Fig. 3.9b. After the zeroing, the signal is transformed back to the frequency domain. Because the values within \((f_0, f_1)\) before the first iteration are believed to be reliable, these values are always restored to their original values. Therefore, transforming the signal back and forth between the time and frequency domains iteratively in this manner, corrections are being made outside \((f_0, f_1)\) in the frequency domain and inside \((t_0, t_1)\) in the time domain during each iteration. The convergence of this algorithm can be proved \([5,59]\). Based on the examples tested in the present work, it often converges in less than 10 iterations.

Since the algorithm could be applied to a time signal as well as an integrated time signal, it was tested on both signals, using Born and Kirchhoff models. From Figs. 3.10 to 3.13, it can be seen that the Gerchberg-Papoulis algorithm works better in the integrated time domain for volumetric (Born) signals and works better in the time domain for cracks (Kirchhoff). Since the flaw type, whether it is volumetric or cracklike, is generally an unknown in a test, neither one of the domains is more favorable to work with. To be consistent, the algorithm must be always applied to only one domain. For the present work, this algorithm will henceforth be applied only to the integrated time domain.

In order to demonstrate when the algorithm will work or fail, three
(a) spectral function

\[ G(f) \]

\[ f_0 \quad \text{allowed to vary} \quad f_1 \quad \text{allowed to vary} \]

(b) spatial function

\[ g(t) \]

\[ t_0 \quad \text{allowed to vary} \quad t_1 \quad \text{reset to zero} \]

Fig. 3.9. Gerchberg-Papoulis iterative algorithm
(a) extrapolated in the time domain

![Graph of extrapolated signal in the time domain]

(b) extrapolated in the integrated time domain

![Graph of extrapolated signal in the integrated time domain]

Fig. 3.10. Two extrapolated (by iteration) integrated time signals of Born for a 800 μm-radius void
Fig. 3.11. The integrated time signal with bandwidth 0-20 MHz of Born for a 800 μm-radius void
(a) extrapolated in the time domain

(b) extrapolated in the integrated time domain

Fig. 3.12. Two extrapolated (by iteration) integrated time signals of Kirchhoff for a 600 μm-effective-radius crack
Fig. 3.13. The integrated time signal with bandwidth 0-20 MHz of Kirchhoff for a 600 μm-effective-radius crack.
Born signal and two Kirchhoff signals of various flaw sizes were considered and the results are displayed in Figs. 3.14 to 3.23. With 10 iterations being assigned to each extrapolation, it takes 18 CPU seconds on a Symbolics 3670 single-user LISP-workstation to perform an extrapolation. For all these examples, both the transmitted and received waves are longitudinal. Only the real or imaginary part of the spectral signals are shown for the Born or Kirchhoff models, respectively.

As can be seen in Figs. 3.14, 3.15, 3.22 and 3.23, whether it is a Born or Kirchhoff signal, the algorithm can restore the low frequency components as long as the cutoff at the low frequency end is less than a quarter of the first cycle in the spectral signal. Whereas if the cutoff is about one quarter of the first cycle, it still works for Born but not Kirchhoff as are shown in Figs. 3.16, 3.17, 3.20 and 3.21. When the cutoff is close to half of the first cycle, it fails completely as are demonstrated in Figs. 3.18 and 3.19.

In the above extrapolated signals, Figs. 3.15a, 3.17a, 3.19a, 3.21a and 3.23a, two "spikes" are seen surrounding each signal. The location of these spikes are the boundary points which define the spatial size of the signal. During the extrapolation process, all points beyond the spatial signal are set to zero, causing discontinuities to occur at these two boundary points and, hence, spikes at these locations.
(a) frequency domain (real) with bandwidth 2-20 MHz

(b) integrated time domain with bandwidth 2-20 MHz

Fig. 3.14. A bandlimited Born signal for a 400 μm-radius void
(a) extrapolated (by iteration) integrated time domain with bandwidth 2-20 MHz

(b) integrated time domain with bandwidth 0-20 MHz

Fig. 3.15. Comparison of two integrated time signals of Born for a 400 μm-radius void
(a) frequency domain (real) with bandwidth 2-20 MHz

(b) integrated time domain with bandwidth 2-20 MHz

Fig. 3.16. A bandlimited Born signal for a 600 μm-radius void
(a) extrapolated (by iteration) integrated time domain with bandwidth 2-20 MHz

(b) integrated time domain with bandwidth 0-20 MHz

Fig. 3.17. Comparison of two integrated time signals of Born for a 600 μm-radius void
(a) frequency domain (real) with bandwidth 2-20 MHz

(b) integrated time domain with bandwidth 2-20 MHz

Fig. 3.18. A bandlimited Born signal for a 1000 μm-radius void
(a) extrapolated (by iteration) integrated time domain with bandwidth 2-20 MHz

(b) integrated time domain with bandwidth 0-20 MHz

Fig. 3.19. Comparison of two integrated time signals of Born for a 1000 μm-radius void
(a) frequency domain (imaginary) with bandwidth 2-20 MHz

(b) integrated time domain with bandwidth 2-20 MHz

Fig. 3.20. A bandwidth Kirchhoff signal for a 300 μm-effective-radius crack
(a) extrapolated (by iteration) integrated time domain
with bandwidth 2-20 MHz

(b) integrated time domain with bandwidth 0-20 MHz

Fig. 3.21. Comparison of two integrated time signals of Kirchhoff
for a 300 μm-effective-radius crack
(a) frequency domain (imaginary) with bandwidth 2-20 MHz

(b) integrated time domain with bandwidth 2-20 MHz

Fig. 3.22. A bandlimited Kirchhoff signal for a 400 μm-effective-radius crack
(a) extrapolated (by iteration) integrated time domain with bandwidth 2-20 MHz

(b) integrated time domain with bandwidth 0-20 MHz

Fig. 3.23. Comparison of two integrated time signals of Kirchhoff for a 400 \( \mu \text{m} \)-effective-radius crack
3.3.2 Polynomial

Besides the iterative scheme, a second extrapolation technique is by polynomial fit. In order to be successful with the polynomial fit, a test function should behave close to such a polynomial. Theoretically, the far-field scattering amplitude of the displacement field $u^S(r, \omega)$ is found to behave as a power series of frequency. Expanding the scattering displacement field as

$$(3.22) \quad u^S(r, \omega) = \sum_{n=0}^{\infty} u_n^S(r) (i\omega)^n$$

where all the coefficients $u_n^S(r)$ are real functions, it has been found that the zeroth and first degree scattering amplitudes, $u_0^S(r)$ and $u_1^S(r)$, vanish for a volumetric scatterer which is localized in all directions [10,27] or an elliptical crack [28]. Hence

$$(3.23) \quad u^S(r, \omega) = \sum_{n=2}^{\infty} u_n^S(r) (i\omega)^n .$$

The real part of the displacement scattering amplitude consists of all even powers of frequency, starting with $\omega^2$, and the imaginary part consists only of odd powers of frequency, starting with $\omega^3$. That is, the real part is an even function of frequency and the imaginary part is an odd function. A polynomial fit can thus be designed so that the real part of the scattering amplitude is extrapolated by a polynomial of the form.
(3.24) \[ R(\omega) = \sum_{j=1}^{n} a_{2j} \omega^{2j} \]

and the imaginary part by a polynomial of the form

(3.25) \[ I(\omega) = \sum_{j=1}^{n} a_{2j+1} \omega^{2j+1}. \]

The coefficients \(a_2, a_3, \ldots, a_{2n+1}\) are determined by a polynomial regression on the 2\(n\) spectral components of the scattering amplitude starting with the minimum frequency \(\omega_{\text{min}}\). It has been found out that, such polynomial fitting with high order polynomials, \(n > 6\), is very unstable. In practice, \(n = 3\) will typically give a smooth fit to the data. For the present work, all polynomial fittings are based on the order \(n = 3\). Therefore, the coefficients \(a_2, a_3, \ldots, a_7\) can be estimated from the following six simultaneous equations

(3.26) \[ \omega a = b \]

where the frequency matrix \(\omega\), (unknown) coefficient vector \(a\) and the inhomogeneous vector \(b\) are

(3.27) \[ \omega = \begin{bmatrix} 2 & 0 & 0 & \omega_0^4 & \omega_0^6 & 0 & 0 & 0 \\ \omega_1^2 & 0 & 0 & \omega_1^4 & \omega_1^6 & 0 & 0 & 0 \\ \omega_2^2 & \omega_2^4 & \omega_2^6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega_0^3 & \omega_0^5 & \omega_0^7 & 0 & 0 \\ 0 & 0 & 0 & \omega_1^3 & \omega_1^5 & \omega_1^7 & 0 & 0 \\ 0 & 0 & 0 & \omega_2^3 & \omega_2^5 & \omega_2^7 & 0 & 0 \end{bmatrix} \]
In these equations, $\omega_0 = \omega_{\text{min}}$, $\omega_1 = \omega_{\text{min}} + \delta \omega$ and $\omega_2 = \omega_{\text{min}} + 2\delta \omega$ are the three consecutive sampling points starting at $\omega_{\text{min}}$ along the frequency axis. The measured values $b_j$'s can be found by comparing the corresponding parts in the following equations

\begin{align*}
(3.30a) \quad b_2 + ib_3 &= u^s(\omega_0) \\
(3.30b) \quad b_4 + ib_5 &= u^s(\omega_1) \\
(3.30c) \quad b_6 + ib_7 &= u^s(\omega_2).
\end{align*}

One of the drawbacks of this technique, as can be seen in Equation (3.27), is that the frequency matrix $\omega$ is relatively ill-conditioned because the three sampling frequencies $\omega_0$, $\omega_1$ and $\omega_2$ are very close. In the actual calculations, the real and imaginary parts of the spectral signal are extrapolated separately. For either parts, the frequency matrix $\omega$ is found to have a condition number of the order $10^6$ for both the Born and Kirchhoff models. Even though the condition number is large, the estimated coefficient vector $a$ is found to be reasonably
accurate.

To demonstrate the results, two examples from Kirchhoff signals are given in Figs. 3.24 to 3.26. The extrapolated results, when compared with the integrated time signals which are not bandlimited at the low frequency end, show considerable improvement around the flaw center. However, it cannot bring the energy down to the horizontal axis around both spatial boundary points. Therefore it does not provide as good an estimation of flaw features as the iterative approach except in finding the flaw center which is defined as the point where the minimum is located. Details of how to identify the flaw center is given in the Appendix. From Figs. 3.25 and 3.26, this extrapolation technique starts to fail when the low frequency end cutoff is close to half of the first cycle in the spectral signal.

In view of the results of these extrapolation techniques, it seems that they both rely on the gross structure of the spectral signals. It should be pointed out, however that the structure of a spectral signal can be distorted severely merely by a shifting of its spatial signal in the time domain. Therefore, determining the zero of time [52] and then shifting the time signal accordingly before doing the extrapolation is very important for the success of both methods.
(a) extrapolated (by polynomial fit) integrated time domain with bandwidth 2-20 MHz

(b) integrated time domain with bandwidth 0-20 MHz

Fig. 3.24. Comparison of two integrated time signals of Kirchhoff for a 300 μm-effective-radius crack
(a) frequency domain (imaginary) with bandwidth 2-20 MHz

![Frequency Domain Graph](image)

(b) integrated time domain with bandwidth 2-20 MHz

![Time Domain Graph](image)

Fig. 3.25. A Kirchhoff signal for a 800 μm-effective-radius crack
(a) extrapolated (by polynomial fit) integrated time domain with bandwidth 2-20 MHz

(b) integrated time domain with bandwidth 0-20 MHz

Fig. 3.26. Comparison of two integrated time signals of Kirchhoff for a 800 µm-effective-radius crack
3.4 Deconvolution of Original Signal

In ultrasonic tests, an electrical voltage is usually the source of energy for the incident ultrasonic wave in the specimen under investigation. At the other end of the system, the monitored output energy is again in the form of an electrical voltage. This output voltage not only contains information about the flaw but also the characteristics of each system component involved. The process of isolating the flaw response from the rest of the testing system is often referred to as the deconvolution or restoration problem in ultrasonic signal processing [29].

In Fig. 3.7, a generalized ultrasonic testing system is displayed. The first component of the system is a pulser. A pulser is an electrical network which drives a transducer to produce an acoustic pulse [30]. The piezoelectric ultrasonic transducer is the most common means of producing or receiving ultrasonic pulses [60,61]. Since it is virtually impossible to transfer energy from an ultrasonic probe into a specimen via an air gap, some form of coupling material must be provided. This can be done by smearing a clean grease onto the specimen surface as with contact probes, or alternatively the whole inspection can be carried out in a water bath. To eliminate the need for a coupling fluid, other means of generating and receiving waves may be considered such as electromagnetic acoustic transducers [60]. The characteristic of a transducer is very complicated. Some of its basic properties can be understood by modeling the transducer as a simple piston source which sits on an ideal baffled surface [62]. Besides the
Fig. 3.27. A generalized ultrasonic testing system
effect of its own electrical and mechanical characteristics, a transducer affects the total system response through the nature of the diffracted sound beam it produces. When travelling inside a material, this sound beam loses energy to material attenuation, affecting also the total system response. A receiver basically contains an electrical coupling network, an amplifier and an analog gate [30]. The amplifier is needed because the electrical signals produced by the receiving transducer are low in amplitude, and the signal-processing electronics require higher voltages [30]. A gate, if present, provides the means to separate a desired section of a signal from the others. After the receiver, a transient recorder [63] may be present if it is necessary to convert the analog output voltage signal to a digital form.

In order to analyse the system, we adopt the idea of Frederick and Seydel [31] and consider each component of the system as a linear time-invariant (LTI) system [64]. Although such an assumption may not be always appropriate, it does provide a good basis for assessing the performance of system components [30].

The behavior of a LTI system is completely described by its impulse time response or its equivalent frequency response. These two responses are linked by the Fourier transform.

Based on such a modeling, a block diagram [30] of the ultrasonic testing system is depicted in Fig. 3.28. On both sides, the symbols for the impulse and frequency response given for each component are also shown. For both the material attenuation frequency responses, \( \alpha_1(\omega) \) and \( \alpha_2(\omega) \) are the frequency dependent attenuation coefficients [65,66].
Impulse Time Response

\[ v_1(t) \quad \text{input voltage} \quad V_1(\omega) \]
\[ b_1(t) \quad \text{pulser} \quad B_1(\omega) \]
\[ x_1(t) \quad \text{sending transducer} \quad X_1(\omega) \]
\[ c_1(t) \quad \text{transducer diffraction} \quad C_1(\omega) \]
\[ m_1(t) \quad \text{material attenuation} \quad \exp(-\alpha_1(\omega)d_1) \]
\[ f(t) \quad \text{flaw} \quad F(\omega) \]
\[ m_2(t) \quad \text{material attenuation} \quad \exp(-\alpha_2(\omega)d_2) \]
\[ c_2(t) \quad \text{diffraction to receiver} \quad C_2(\omega) \]
\[ x_2(t) \quad \text{receiving transducer} \quad X_2(\omega) \]
\[ r(t) \quad \text{receiver} \quad R(\omega) \]
\[ g(t) \quad \text{gating} \quad G(\omega) \]
\[ s(t) \quad \text{digitization} \quad S(\omega) \]
\[ v_0(t) \quad \text{output voltage} \quad V_0(\omega) \]

Impulse Frequency Response

Fig. 3.28. Elements of an ultrasonic system modeled as a linear, time-invariant system
While the coefficient $\alpha_1(\omega)$ and distance $d_1$ are the measurements for the acoustic path between the transmitting transducer and the flaw, $\alpha_2(\omega)$ and $d_2$ are the measurements for the path between the flaw and the receiving transducer.

Hence, the output frequency response is

$$V_0(\omega) = S(\omega)G(\omega)R(\omega)X_2(\omega)C_2(\omega)e^{-\alpha_2(\omega)d_2}F(\omega)e^{-\alpha_1(\omega)d_1}C_1(\omega).$$

In principle, all the responses plus the input and output voltages can be measured and therefore the flaw response can be obtained by direct deconvolution

$$F(\omega) = \frac{V_0(\omega)}{S(\omega)G(\omega)R(\omega)B_1(\omega)V_1(\omega)}.$$ 

However, this is impractical. The following sections will discuss two different deconvolution techniques trying to recover the flaw response $F(\omega)$ or $f(t)$ in a practical situation.

3.4.1 Measurement Model of Thompson-Gray

One of the methods to deconvolve the flaw signal $F(\omega)$ was proposed by Thompson and Gray [1] in 1983. Its derivation is based on the assumption that the relative flaw size is sufficiently small in comparison to the incident beam width so that at the flaw location, the incident wave can be considered as a quasiplane wave [1,2,67].
For simplicity, consider for example, a simple pulse-echo contact experiment as is shown in Fig. 3.29a. In a pulse-echo setup, the same transducer is used as the transmitter and the receiver and thus $X_1(\omega) = X_2(\omega) = X(\omega)$. From the geometry, both the incident and scattered waves travel the same path and hence $\alpha_1(\omega) = \alpha_2(\omega) = \alpha(\omega)$, $d_1 = d_2 = d_o$. For the present measurement model, errors due to gating and digitization are ignored, $G(\omega) = S(\omega) = 1$. To estimate the transducer diffraction factor $C_1(\omega)$ or $C_2(\omega)$, the transducer is considered as a piston source sitting on a baffled surface as is shown in Fig. 3.30 where the specimen is on the side $z \geq 0$. By considering the specimen as a fluid medium, both diffraction factors are found to be equivalent for a pulse-echo configuration $C_1(\omega) = C_2(\omega) = C(\omega)$ and for small flaws near the center of the beam [68]

\begin{equation}
C(\omega) = C(\omega, d_o) = \frac{-ik}{2\pi} \int_{A_t} \frac{e^{iKR}}{R} dA
\end{equation}

where $A_t$ is the transducer surface and $R = |x - x_o|$. Therefore, $C(\omega)$ can be calculated once the flaw is located, i.e., $d_o$ is known. This integral, in Equation (3.33), is a special case of the Rayleigh diffraction integral [1] with the integrand $(e^{ikR})/(4\pi R)$ being the Green's function for the present case. In an immersion setup in which both the transducer and the specimen are in a fluid medium, the same technique, Equation (3.33), with the integrand replaced by the
(a) a flawed sample

(b) a flawless sample

Fig. 3.29. Pulse-echo contact experiments
Fig. 3.30. Transducer modeling

\[ x = (x, y, z) \]

\[ x_o = (x_o, y_o, z_o) \]
appropriate Green's function is found to be impractical. Instead, the factor is derived based on the reciprocity relation [1].

From the above discussion,

\[
V_{o}(\omega) = R(\omega)X^{2}(\omega)C^{2}(\omega)e^{-2\alpha(\omega)d}F(\omega)B(\omega)V_{i}(\omega).
\]

Here, a reference setup as is shown in Fig. 3.29b is considered. The specimen used has the same material as the flawed specimen. By measuring the back surface echo, the output voltage is

\[
V_{o}^{\text{ref}}(\omega) = R(\omega)X^{2}(\omega)D(\omega)e^{-2\alpha(\omega)d}B(\omega)V_{i}(\omega)
\]

where \(D(\omega)\) is the total diffraction effect due to propagation to the free surface and back. Similarly, based on the above transducer model, this factor is found to be [68]

\[
D(\omega) = e^{2ikd}D(d)
\]

where

\[
D(d) = \frac{-ik e^{-2ikd}}{2\pi} \int_{A_{t}} \int_{A_{t}} \left(\frac{e^{ikR}}{R} dA\right) dA.
\]

Again, \(D(\omega)\) can be calculated as soon as the flaw is located.

From Equations (3.34) and (3.35), the flaw response can be written as
On the right hand side, the attenuation coefficient $\alpha(\omega)$ can be measured in a separate experiment \[65, 66\] on the same kind of material. The flaw location $d_0$ can be calculated by timing delay from the output signal. The back surface distance $d$ is given by the reference configuration.

When both output signals are received, $v_o(t)$ and $v_0^{\text{ref}}(t)$, they can be transformed to $V_o(\omega)$ and $V_0^{\text{ref}}(\omega)$ by a FFT algorithm. As a result, all the quantities on the right hand side of Equation (3.38) are known.

However, at low signal-to-noise-ratio deconvolution by straight division is unstable. Thus, Equation (3.38) must be modified to handle this situation. One successful method that has been employed in this regard is the Wiener filter \[32\].

To understand the origin of the Wiener filter concept, let us rewrite the flaw signal as the output signal $O(\omega)$ divided by the reference signal $R(\omega)$

\[
(3.39) \quad F(\omega) = \frac{O(\omega)}{R(\omega)}.
\]

If both measurements contain noise elements given by $n(t)$ and $m(t)$, respectively, whose Fourier transforms are $N(\omega)$ and $M(\omega)$, respectively, then
(3.40) \[ F(\omega) = \frac{O(\omega) + N(\omega)}{R(\omega) + M(\omega)}. \]

For small noise, Equation (3.39) is recovered. However, if \( O(\omega) < N(\omega) \) at certain frequencies

(3.41) \[ F(\omega) = \frac{N(\omega)}{R(\omega)} \]

which is only a measurement of the noise rather than the actual signal. To avoid this, both the denominator and the numerator of Equation (3.40) are multiplied by \([R(\omega) + M(\omega)]^*\)

(3.42) \[ F(\omega) = \frac{\left[ O(\omega) + N(\omega) \right] \left[ R(\omega) + M(\omega) \right]^*}{\left[ R(\omega) + M(\omega) \right] \left[ R(\omega) + M(\omega) \right]^*} \]

\[ = \frac{O(\omega)R^*(\omega) + O(\omega)M^*(\omega) + N(\omega)R^*(\omega) + N(\omega)M^*(\omega)}{|R(\omega)|^2 + |M(\omega)|^2 + M(\omega)R^*(\omega) + R(\omega)M^*(\omega)} \]

where ( )* denotes complex conjugation. Quantities such as \( O(\omega)M^*(\omega) \) etc., can be proved to be the cross-correlation of their time signals \( o(t) \) and \( m(t) \) [69]. Assuming the noise terms are uncorrelated with \( O(\omega) \) and \( R(\omega) \) and with each other, then

\( O(\omega)M^*(\omega) = N(\omega)R^*(\omega) = N(\omega)M^*(\omega) = M(\omega)R^*(\omega) = R(\omega)M^*(\omega) = 0 \) and

(3.43) \[ F(\omega) = \frac{O(\omega)R^*(\omega)}{|R(\omega)|^2 + |M(\omega)|^2}. \]

With this expression, as long as \(|M(\omega)|^2\) is small,
regardless whether \(0(\omega)\) is small or not. An equivalent form of Equation (3.43) is

\[
(3.45) \quad F(\omega) = \frac{0(\omega)}{R(\omega)} \left( \frac{|R(\omega)|^2}{|R(\omega)|^2 + |M(\omega)|^2} \right). 
\]

Therefore, this concept is to apply the Wiener filter \(W(\omega)\) to the original deconvolution, Equation (3.39),

\[
(3.46) \quad F(\omega) = \frac{0(\omega)}{R(\omega)} W(\omega)
\]

where

\[
(3.47) \quad W(\omega) = \frac{|R(\omega)|^2}{|R(\omega)|^2 + Q^2}.
\]

The term \(Q\) is taken as a constant

\[
(3.48) \quad Q = c |R(\omega)|_{\text{max}}
\]

as a simple measure of the noise term. The index \(c\) is chosen to be 0.1 for the examples in this section.

To see how well the measurement model works, two experimental examples for a two-to-one oblate spheroidal void as shown in Fig. 3.31 are exhibited in Figs. 3.32 and 3.34. The first case (Fig. 3.32) is for
Fig. 3.31. Scattering configuration of a 2-to-1 oblate spheroid
(a) output voltage $V_0(\omega)$

\[ 3.0404 \times 10^{-6} \text{S} \]

(b) flaw response $F(\omega)$ by measurement model

\[ 0.0049759745 \times 10^{-4} \text{S} \]

Fig. 3.32. Original and deconvolved scattering magnitudes at $\theta = 176^\circ$ for a 2-to-1 oblate spheroidal void
(a) Wiener-filtered flaw response $F(\omega)W(\omega)$ by measurement model

(b) Wiener-filtered numerical (MOOT) data

Fig. 3.33. Comparison of frequency magnitudes ($\Theta = 176^\circ$) of filtered measurement model results and filtered numerical (MOOT) data
(a) output voltage $V_o(\omega)$

(b) flaw response $F(\omega)$ by measurement model

Fig. 3.34. Experimental scattering magnitudes ($\theta = 137^\circ$) for a 2-to-1 oblate spheroidal void
(a) Wiener-filtered flaw response $F(\omega)W(\omega)$ by measurement model

(b) Wiener-filtered numerical (MOOT) data

Fig. 3.35. Comparison of frequency magnitudes ($\theta = 136^\circ$) of filtered measurement model results and filtered numerical (MOOT) data
\[ \Theta = 176^\circ \] while the second (Fig. 3.34) is at \( \Theta = 136.76^\circ \). In both figures, the top graphs are the spectral magnitudes for the (undeconvolved) output voltages while the bottom ones are the results after being deconvolved by the measurement model (in all cases shown, material attenuation was small so that attenuation corrections were ignored). These results may be compared with the complementary theoretical (numerical) results obtained by MOOT [25] as shown in Figs. 3.33b and 3.35b. Notice that, the look angles \( \Theta \) between the experimental signals and the MOOT signals are slightly different. This small difference is not deemed to be significant.

3.4.2 Flaw-derived Reference

With an impulsive input signal, the output signal from a flaw in an ultrasonic test usually contains a leading pulse plus other later arriving responses. All these signal components are inevitably corrupted by the system characteristics. If the scatterer is a volumetric flaw, then the leading pulse is the front surface echo which is the same as from a perfect reflector at the flaw location but normalized by the unknown flaw Gaussian curvature and reflection coefficient. Since this front pulse contains also all the characteristics of the testing system except the flaw, it can be extracted and used as a reference signal (with an unknown normalization factor) in the deconvolution process. This loss of the information on the normalization factor is acceptable if only relative amplitude information is needed from the deconvolution process. For our
classification features, this restriction is acceptable. We also should
point out that for sizing methods such as the Born method [48], this is
also acceptable. Fig. 3.36 shows a block diagram of the algorithm which
extracts the reference signal from the output voltage.

Realizing the front pulse contains significant high frequency
components, both real and imaginary parts of the original signal $V_0(\omega)$
are multiplied by a Hanning window in order to get rid of the low
frequency components and thus make it easier to identify the front
pulse. By definition, the Hanning window is basically the shifted
cosine function with only half of a cycle

$$H(\omega) = 0.5\left[1 - \cos\left(\frac{2\pi t}{T}\right)\right]$$

for $t \in [0,T]$.

After multiplication by the Hanning window, the signal is then
transformed to the time domain. In the time domain, the front pulse is
identified. Every point is set to zero except the front pulse. This
front pulse is then shifted to the origin (to remove any large linear
phase terms) and is transformed back to the frequency domain as the
reference signal.

To illustrate this algorithm, an example of the original scattering
measurement at $\Theta = 176^\circ$ for a two-to-one oblate spheroidal void is given
in Figs. 3.37 and 3.38. In Fig. 3.37, the effect of the Hanning window
is shown when comparing the two time signals. Clearly, after the
Hanning window, the front pulse becomes more distinctive. In Fig. 3.38.
Fig. 3.36. Algorithm for the flaw-derived reference signal
(a) output voltage $v_o(t)$

![Graph of $v_o(t)$ before the Hanning window]

(b) $v_o(t)$ after the Hanning window

![Graph of $v_o(t)$ after the Hanning window]

Fig. 3.37. The output voltage $v_o(t)$ before and after the Hanning window.
(a) $V^\text{ref}_o(t)$

(b) $V^\text{ref}_o(\omega)$ (magnitude)

Fig. 3.38. Flaw-derived reference
the results of the flaw-derived reference signal are shown. In view of the above process, the reference signal $V_{o}^{\text{ref}}(\omega)$ is always the same as the output voltage $V_{o}^{o}(\omega)$ except it does not have the flaw response and it has an extra response due to the Hanning window $H(\omega)$ and an unknown normalization factor $F_{o}$. Therefore

$$V_{o}^{\text{ref}}(\omega) = \frac{F_{o} V_{o}^{o}(\omega)}{F(\omega)} H(\omega) \quad (3.50)$$

or equivalently

$$F(\omega) = \frac{F_{o} V_{o}^{o}(\omega) H(\omega)}{V_{o}^{\text{ref}}(\omega)} \quad (3.51)$$

For simplicity, this flaw signal is written as

$$F(\omega) = \frac{O(\omega)}{R(\omega)} \quad (3.52)$$

where $O(\omega) = F_{o} V_{o}^{o}(\omega) H(\omega)$ is the Hanning windowed effective input voltage and $R(\omega) = V_{o}^{\text{ref}}(\omega)$ is the reference signal. Again, based on the same reason as was discussed in the measurement model, Equation (3.52) is replaced by

$$F(\omega) = \frac{O(\omega)}{W(\omega)} \quad (3.53)$$

where $W(\omega)$ is the Wiener filter given in Equation (3.47). The Wiener index $c$ is chosen to be 0.1 for the following examples (the same value
was used for the examples in the measurement model). To demonstrate the deconvolution process, experimental examples using the flaw-derived reference are compared with numerical calculations using Wiener filtered MOOT results [25] in Figs. 3.39 and 3.41. Notice that both signals have slightly different look angles $\theta$. In Figs. 3.40 and 3.41, the original MOOT signals are also given. Inspecting the MOOT signals after multiplication by the Wiener filter, it seems the Wiener index, $c = 0.1$, could be chosen significantly smaller if it was desirable to have the Wiener filter to have less of a distortion effect on the original signals.
(a) flaw response by the flaw-derived reference

(b) Wiener-filtered numerical (MOOT) data

Fig. 3.39. Comparison of frequency magnitudes ($\theta = 176^\circ$) deconvolved by the flaw-derived reference with the Wiener-filtered numerical (MOOT) data
Fig. 3.40. Original MOOT spectral magnitude for a 2-to-1 oblate spheroidal void at $\theta = 175^\circ$. 
(a) flaw response by the flaw-derived reference

(b) Wiener-filtered numerical (MOOT) data

Fig. 3.41. Comparison of frequency magnitudes ($\theta = 136^\circ$) deconvolved by the flaw-derived reference with the Wiener-filtered numerical (MOOT) data
Fig. 3.42. Original Moot spectral magnitude for a 2-to-1 oblate spheroidal void at $\theta = 135^\circ$
4. FEATURE EXTRACTION AND FLAW CLASSIFICATION PROCEDURES

4.1 Introduction

Because of the inevitable presence of noise and distortion from the measuring devices, the classification of flaws cannot always be solved by completely deterministic procedures. Therefore one approach that can be employed is the use of rule-based expert system techniques based on heuristic knowledge. In fact, as reported in the Artificial Intelligence literature, many of the successful expert systems are classification schemes [70-73]. Currently, a rule-based ultrasonic flaw classification expert system of this type is being developed at the Iowa State Center for Nondestructive Evaluation [41,42].

Another approach for dealing with classification problems that has come from the Artificial Intelligence field is based on adaptive learning methods. Such learning schemes usually rely on some kind of discrimination or decision function [74-77] and is the approach chosen for the present work.

In the following sections, the basic concepts of adaptive learning and the learning algorithm which is being employed in this work will be introduced. Discussion of some of the criticisms of adaptive learning and how the present work deals with them will be also considered.

The success of any classification depends upon the choice of features. In general, extracting a set of good features so that the job of classification can be done reliably is the most difficult part of a classification scheme. Therefore, we will also discuss the practical
ways to extract flaw features from an acoustic signal.

4.2 Brief Discussion of Adaptive Learning

Learning, ranging from simple rote memorization tasks to tackling complex scientific principles, has always been a part of our daily routines. Even though humans are very experienced in learning, we have yet to understand the fundamental mechanisms of learning. Until we have that understanding, it will be impossible to build a machine that can truly learn like a human being [78]. For this reason we adopt the idea of Forsyth and Rada [74] to define a learning machine or algorithm as a computer system which can improve its performance at a given task over time without re-programming. By this definition, a learning system must be able to learn adaptively. Traditional rule-based expert systems [75] do not have this capability and hence are not considered as learning systems. Currently, adaptive learning systems are often discussed in the field of pattern recognition where various numerical techniques are used for classifying patterns. Patterns are sets of features that are extracted from input data. Each pattern can be represented as an ordered set of numbers, where each number is the value of a feature. Geometrically, a pattern is a vector represented by a point in the feature space. Viewed in this light, the adaptive learning system we will use here can also be considered as a pattern recognition system.
4.2.1 Adaptive Learning by Parameter Adjustment

There are a variety of adaptive learning techniques [74,77] among which, parameter adjustment (PA) is one of the simplest forms of learning. Basically, a PA system is an explicit model containing free parameters. These parameters may be adjusted to improve performance. Learning consists of changing the free parameters based on observed performance [77]. One of the earliest and most effective PA systems was the checker-playing program proposed by Samuel [79,80]. In his program, he used a polynomial with sixteen parameters as its static evaluation function.

A PA system in general has two stages of processing:
1. feature extraction from input (numerical) data resulting in patterns in the feature space,
2. pattern classification of patterns, assigning them to one of two or more classes in the classification space.

Selecting the features to be extracted from data and processing them into a form that makes classification possible is crucial to the success of any classifier. This part will be discussed in the last section of the current chapter.

For a system that has \( n \) features, a pattern \( \mathbf{x} \) is a vector with \( n \) values in the \( n \)-dimensional feature space (hyperspace),

\[
(4.1) \quad \mathbf{x}^T = (x_1 \ x_2 \ x_3 \ ... \ x_n)
\]

where \( \mathbf{x}^T \) denotes the transposed vector of \( \mathbf{x} \). Each of these values may
be the value of the extracted feature itself, or the resultant value from scaling, normalization, or other processing of a feature. If the features are chosen such that all patterns belonging to the same class cluster together and there is no overlapping of clusters, hyperplanes can be set up to separate the hyperspace into regions that represent the different classes. The mathematical equations for these hyperplanes are usually called decision functions or discrimination functions. If clusters are overlapping, the decision function approach will not work with deterministic functions and statistical functions are required to determine the likelihood that a pattern is in a particular class [35,36]. Sometimes, a different selection or processing of the features can improve the separation among different classes. For the deterministic case, the coefficients of the decision functions are adjusted through a supervised training procedure. This type of learning is basically experience-based and therefore, the more cases of association observed, the more likely the association is to be generally true [74].

4.2.2 Linearly Separable System

From the compactness hypothesis [81], patterns or feature vectors of a given class are assumed in some sense nearer to all patterns in that class than to all or most of the patterns in other classes. The patterns in a given class occupy a region in the feature space which is called a class region [36]. When the class regions do not overlap, the classes are said to be separable. If, for every class region, a
hyperplane (whose expression is a linear function) can be placed to separate that region from all other class regions, the classes are said to be linearly separable. A rigorous proof of linear separability under certain assumptions can be found in [36].

Ideally, a system which is not linearly separable can always be transformed into a linear separable one by doing nonlinear distortion in the original hyperspace (feature space) [82]. However, this transformation usually results in a new hyperspace with drastically high dimensionality and very expensive computing effort [34]. In practice, a nonlinearly separable system is dealt with by using statistical functions or other techniques based on the individual situation [36].

Consider now a system having k classes that are separable. Then there exists a set of functions \( \{g_i\} \) such that [34]

\[
g_i(x) \geq g_j(x) \quad \text{for all } j
\]

if and only if the feature vector or pattern \( x \) is a member of class \( i \). These classes are said to be linearly separable if the functions are linear in terms of the features, i.e.,

\[
g_i(x) = w_i^T \cdot x + v_{i0}
\]

where \( x \) is the feature vector, \( w_i \) is the weight vector containing the adjustable parameters and \( v_{i0} \) is a constant. The dimension for both \( x \) and \( w_i \) is equal to the number of features. Equation (4.3) also can be written as
where $\mathbf{v}_i$ and $\mathbf{y}$ are augmented weight and feature vectors respectively such that

\begin{align*}
(4.5a) \quad \mathbf{v}_i^T &= (v_{i1}^1 v_{i1}^2 \ldots v_{i1}^{n+1}) \\
&= (v_{10}^i v_{i1}^1 v_{i1}^2 \ldots v_{i1}^n) \\
(4.5b) \quad \mathbf{y}_i^T &= (y_1^1 y_1^2 \ldots y_1^{n+1}) \\
&= (1 x_1^1 x_1^2 \ldots x_1^n). 
\end{align*}

In general, $k$ decision functions are needed for a system with $k$ classes. In particular, considering a system with only two classes, only one hyperplane is needed to separate the classes as is shown in Fig. 4.1 with two features. Assuming there exist two decision functions $g_1(\mathbf{y})$ and $g_2(\mathbf{y})$, we can combine them to form a single decision function

\begin{equation}
(4.6) \quad g(\mathbf{y}) = g_1(\mathbf{y}) - g_2(\mathbf{y}).
\end{equation}

From Equation (4.4)

\begin{equation}
(4.7) \quad g(\mathbf{y}) = (\mathbf{v}_1^T - \mathbf{v}_2^T) \cdot \mathbf{y} = \mathbf{v}^T \cdot \mathbf{y}
\end{equation}

where $\mathbf{v} = \mathbf{v}_1 - \mathbf{v}_2$. Therefore, from Equation (4.2), the relation

\begin{equation}
(4.8) \quad g(\mathbf{y}) > 0
\end{equation}

is satisfied if and only if $\mathbf{y}$ is a member of class 1. For this case, a learning algorithm based on an error correction procedure can be
Fig. 4.1. A decision function, $g(x) = 0$, separates the two class regions.
developed [34,36]. The elements of this error correction procedure are as follows.

1. Let \( v(n) \) be the approximation used to classify \( y(n) \) at trial \( n \).
2. In the first trial \( (n = 1) \), the weight vector can be set to an arbitrary vector or \( v^T(1) = (0 \ 0 \ ... \ 0) \).
3. Given \( y(n) \), the following case rules are to be followed.
   - **case 1:** \( y(n) \) ∈ class 1. Here we have two possibilities.
     - a. If \( v^T(n)y(n) > 0 \), \( y(n+1) = v(n) \).
     - b. If \( v^T(n)y(n) \leq 0 \), an error has occurred.
       The adjustment is \( y(n+1) = y(n) + c y(n) \) where \( c \) is an arbitrary positive constant.
   - **case 2:** \( y(n) \) ∈ class 2. Again there are two possibilities.
     - a. If \( v^T(n)y(n) < 0 \), \( y(n+1) = v(n) \).
     - b. If \( v^T(n)y(n) \geq 0 \), \( y(n+1) = y(n) - c y(n) \).

To understand why this algorithm would work, we examine the decision function in Equation (4.7) and express it by the definition of vector dot product

\[
(4.9) \quad g(y) = ||v|| ||y|| \cos \theta
\]

where \( || \cdot || \) denotes the norm of a vector and \( \theta \) is the angle between \( v \) and \( y \). In Equation (4.9), we have applied the fact that the norm of a vector is equivalent to the norm of its transpose. Since the norm of a vector is always nonnegative, the decision function is therefore a measure of the extent to which \( v \) and \( y \) point in the same or opposite directions [76]. For this algorithm, a \( v \) must be found which points in
the same direction as the feature vectors for class 1 examples and in
the opposite direction for class 2 feature vectors. In the situation of
misclassification as in case 1b for example, adding the positive
quantity cy(n) to v(n) causes it to point more in the direction of y(n).
Similarly, in case 2b, subtracting cy(n) from v(n) moves it to point
more in the opposite direction to y(n). The result is that v^T.y always
improves after adjustment. By adding feature vectors y(n) to v, it
causes v move toward the mean of y(n)

$$v_c = \frac{1}{N} \sum_{n=1}^{N} y(n).$$

(4.10)

It can be shown that this training procedure converges to a solution
vector in a finite number of trials whenever a solution exists [36].
When no solution exists, the procedure produces a weight vector which
remains bounded for any size of the training set [83]. It should be
noted that this algorithm is not restricted to two-category systems.
The training procedure also can be generalized for a multiclass problem
[36].

4.2.3 Major Disadvantages and Advantages of Adaptive Learning

Adaptive learning, like all techniques, has a number of significant
advantages and disadvantages when compared with other methods [74,77].
In this work, we have attempted to minimize the disadvantages as
follows.

1. One of the criticisms of adaptive learning methods is that the
knowledge that the system gains during its training phase is opaque [74]. In the training process, a set of coefficients (numbers) are optimized, but the meaning of these optimal coefficients is usually not apparent.

The system we will propose here does not overcome this problem. In fact, we think this "opaqueness" will always be a characteristic of adaptive learning schemes. However, in cases where only the results of the classification process is of paramount importance, this might be considered as an actual convenience.

2. Another objection raised against adaptive learning schemes is that the features used are numeric and thus are computable but not intelligible [74].

In the present work, however, since all features are extracted from known models, they do have direct physical significance in terms of the classes they represent. Thus, we feel a model-based adaptive learning system can eliminate this negative aspect of "traditional" adaptive learning schemes.

3. A third disadvantage of adaptive learning systems is the inability to understand why errors occur in the classification process because of the difficulty of representing decision surfaces and feature vectors in an intelligible manner for multi-dimensional feature spaces [74].

Here, however, we will always use a maximum of two features at any time. Thus, the decision surface and feature vectors are all easily represented graphically.
4. Finally, an objection can be raised to the assumption of linear separability, since few problems are directly of this nature. Our answer to this objection is that all three features we use are extracted from known models. With appropriate normalization and processing of the features, as will be discussed in the next section, this model-based system can, in fact, be proven to be (at least theoretically) linearly separable.

By employing an adaptive learning system, we also gain certain advantages over other approaches. Two of the major advantages are given in the following.

1. The system can adapt its decision making process as conditions change. One of the major competitors of adaptive learning systems are rule-based expert system approaches and to date these systems have been static systems where the knowledge base is essentially fixed.

2. Adaptive learning systems can tolerate noise and model errors and incorporate them easily into the decision making process [74]. Other methods, such as rule-based expert systems also must deal with errors and noise to function properly. However, such accommodations are usually made through the use of heuristic confidence factors which must be chosen a priori. If the classification environment is substantially different from the expert system designers expectations, large errors could occur in the classification process because of these differences. In some cases, this could require constant readjustments of the confidence factors, resulting in a
system based more on "fudge" factors than on facts.

4.3 Feature Extraction

As described in Chapter 3, all three features that we are relying on to do classification are taken from either a time or an integrated time signal. The locations of these signals on the time axis depend on the time scale used while the signals are being collected. Therefore, it is essential that the flaw signal be correctly detected on the time scale before feature extraction is possible. We will not, however, consider explicitly the detection process. Instead, we will assume that the flaw signal has previously been properly identified.

For a flaw signal in the time domain, the first portion of the response is called here the leading edge response. When the incident plane wave is a Dirac delta function, the ideal leading edge response contains a pulse which is either a Dirac delta function for a volumetric flaw or a square root singularity for a crack. With a finite bandwidth, ringing is induced everywhere before and after the leading edge pulse for both cases. This phenomenon was demonstrated in Figs. 3.2a to 3.5a. Notice that the period of the ringing for a bandlimited delta function is different from the one for a square root singularity. Basically, one may rely on this phenomenon and assign the first extrema around which there is no larger peak within a time interval $\alpha \Delta t$ to be the leading edge pulse. Here $\Delta t$ is the period measured in the ringing of a bandlimited delta function, $\alpha$ is a "fudge factor" determined by experience which tries to account for the difference for a bandlimited
square root singularity and the effect of noise. However, such a search does not include considerations such as distortion of the amplitude value due to bandlimitation, noise, etc. Therefore, a more detailed search procedure is necessary to properly identify the leading edge pulse. Details of this complete search method can be found in the Appendix. Once the leading edge pulse location is known, the beginning of the signal can be estimated.

To obtain the integrated time signal, the time signal is integrated once along the time axis. In the present work, the first zero crossing point of the integrated time signal is used as the beginning of the signal. Hence, from the beginning of the integrated time signal, one can measure both features F1 and F2 at a distance of the Born rise time which is defined in Equation (3.4) and is shown in Fig. 4.2.

In order to control the size of both features F1 and F2, some kind of normalization is necessary. Before measuring F1, the integrated time signal is normalized by its extrema. Therefore value of F1 is expected to be around 1 or -1 for an inclusion with high or low acoustic impedance respectively. In case of a crack, we expect F1 ≥ -1 in general, depending on the effective crack size.

Before normalizing the feature F2, an estimate of the flaw size in time is defined as

\[ (4.11) \Delta t_{center} = \text{time between the first extrema from the beginning of an integrated time signal and its following zero crossing point.} \]

For synthetic data using either a volumetric (Born) or crack (Kirchhoff)
(a) Born signal

(b) Kirchhoff signal

Fig. 4.2. Born rise time ($\Delta t_{\text{Born}}$) and size estimate ($\Delta t_{\text{center}}$) measured in the integrated time signal.
model, $\Delta t_{\text{center}}$ is essentially the exact measurement of the effective flaw size as is illustrated in Fig. 4.2. Even for more exact synthetic data, this $\Delta t_{\text{center}}$ is still very much the measurement of the effective flaw size as is shown in Fig. 4.3 [25]. To normalize $F_2$, it is divided by the slope defined by the straight line connecting the first extrema at the beginning of an integrated time signal and its following zero crossing point. Since the time between this extrema and zero crossing point is just $\Delta t_{\text{center}}$, we have, for example (see Fig. 4.3)

\[
F_2^{\text{normalized}} = \frac{F_2}{A} \frac{1}{\Delta t_{\text{center}}}
\]

The third feature $F_3$ is the amplitude ratio for the time signal. In practice, this feature is taken as the ratio between the following absolute amplitudes: $A_2$ the largest local extrema (which has the opposite polarity comparing to the leading edge pulse) within an interval of $\beta \Delta t_{\text{Born}}$ measured from the leading edge pulse tip and $A_1$ the value at the leading edge pulse tip. When there is no local extrema within a small time interval from the leading edge pulse (taken to be $\beta \Delta t_{\text{Born}}$ from the leading edge pulse tip), $F_3$ is set to zero. The choice $\beta = 2$ was found to be adequate for all the experimental results we have considered. Figs. 4.4 and 4.5 show the measurements of $F_3$ by using model signals. Although theoretically $F_3$ should be either zero or around one for an ideal (infinite bandwidth) crack signal, for bandlimited crack signals, $F_3$ can actually have a negative value as can
(a) 2-to-1 oblate spheroidal void

(b) circular crack

Fig. 4.3. Size estimate ($\Delta t_{\text{center}}$) measured in the integrated time signals of NDT data.
Fig. 4.4. Measuring $F_3$ in a volumetric (Born) signal
(a) small crack

(b) large crack

Fig. 4.5. Measuring $F_3$ in a crack (Kirchhoff) signal
be seen in Fig. 4.5b. In the present work, we will also set $F_3 = 0$ whenever a negative value is obtained. Therefore, the value of $F_3$ is expected to either 0 or close to 1 for cracks. However, for volumetric flaws which are not too small or large (see Table 3.1), recall $F_3$ should have a value between 0 and 1. This means that we could have volumetric (v) and crack (c) feature vectors (points) lying in the feature space of $F_1$-$F_3$ as shown in Fig. 4.6. It is clear from that figure that volumetric flaws and cracks are not linearly separable for this choice of features. This problem can be solved, however, by calculating and using a feature $F_3'$ instead of $F_3$, where

$$F_3' = F_3(1 - F_3).$$

In this case, along the $F_3'$ axis, cracks are always at zero or very close to zero and volumetric flaws are always at points bigger than zero as is shown in Fig. 4.7. For the sake of convenience, the prime in $F_3'$ will be dropped and $F_3$ is always understood as $F_3'$ in the rest of this thesis.

Although we have three features, only two of them, either $F_1$ and $F_2$ or $F_1$ and $F_3$ will be considered in each classification session. As is discussed in Chapter 2 when $F_3$ was first introduced, $F_3$ will be keyed on only when $F_2$ is not reliable.

Theoretically, misclassification in the feature space of $F_1$ and $F_2$ will not occur until the effective flaw size is so small such that its rise time $\Delta t_{\text{flaw}}$ is
Fig. 4.6. Feature space of F1 and F3 (original)
Fig. 4.7. Feature space of F1 and F3'

- low impedance
- high impedance
- cracks
(4.14) \( \Delta t_{\text{flaw}} = \Delta t_{\text{Born}} \)

where \( \Delta t_{\text{flaw}} \) is the time difference between the beginning point and its following extrema in the integrated time signal and \( \Delta t_{\text{Born}} \) was given in Equation (3.4). However, since a finite time segment after \( \Delta t_{\text{Born}} \) is needed to estimate \( F_2 \), if the flaw is a crack, the endpoint of this time segment can fall beyond the first minimum of the integrated time signal even when

(4.15) \( \Delta t_{\text{flaw}} \geq \Delta t_{\text{Born}} \)

resulting in a misclassification (Fig. 4.8a). With the above consideration and after inspecting different results obtained from experiments or MOOT [25], we decided that as long as the flaw signal is sufficiently large, i.e., \( \Delta t_{\text{center}} \geq 2 \Delta t_{\text{Born}} \), \( F_2 \) will be reliable (Fig. 4.8b) and the feature space of \( F_1 \) and \( F_2 \) will be considered. This is true because for both volumetric flaws and cracks (Fig. 4.8b) this restriction will guarantee that we have a sufficient time interval available after \( \Delta t_{\text{Born}} \) to use as a finite time segment to estimate the slope feature \( F_2 \). The class regions for this hyperspace are shown in Fig. 4.9a. For a flaw signal that has a \( \Delta t_{\text{center}} \) such that \( \Delta t_{\text{Born}} \leq \Delta t_{\text{center}} < 2 \Delta t_{\text{Born}} \), the feature space of \( F_1 \) and \( F_3 \) will be used and the flaw classes in this space are shown in Fig. 4.9b. Otherwise, i.e., if it is found \( \Delta t_{\text{center}} < \Delta t_{\text{Born}} \), the flaw is considered too small to be classifiable.

In using the above selected features, we note that the
classification scheme that we use in the present work is the so-called Perceptron algorithm [37-39] given in section 4.2.2 except the positive constant $c$ is always chosen to be 1.
(a) misclassification by F2 for a crack

\[ \Delta t_{\text{flaw}} = \Delta t_{\text{center}} \]

(b) correct classification by F2

\[ \Delta t_{\text{Born}} = \Delta t_{\text{flaw}} \leq \frac{\Delta t_{\text{center}}}{2} \]

Fig. 4.8. Measuring the slope feature F2
(a) F1-F2 classification space

-1 0 1

F1

F2

inclusions with low impedance

inclusions with high impedance

cracks

(b) F1-F3 classification space

-1 0 1

F1

F3

inclusions with low impedance

inclusions with high impedance

Fig. 4.9. Classification spaces of F1-F2 and F1-F3
5. DEMONSTRATION EXAMPLES

The success of any adaptive learning system relies on the (supervised) training on a set of good and "real" data. In our case, we chose the synthesized data from MOOT [25] to do the training for both feature spaces F1-F3 and F1-F2. These MOOT data consist only of two flaws, a two-to-one oblate spheroidal void and a circular crack. With the frequency range chosen to be 0-20 MHz, the long semiaxis of the spheroid has a length of 483 μm and the radius of the crack is also 483 μm. As is shown in Fig. 5.1, the scattering configuration can be considered as either pulse-echo or pitch-catch with the incident wave traveling along the negative z-direction. Only longitudinal-longitudinal scattering is considered in all the examples demonstrated in this chapter. For the frequency range 0-20 MHz, the Born rise time is found to be

\[ \Delta t_{\text{Born}} = 4 \times 0.00978 \mu\text{sec} . \]

As is discussed in the previous chapter, flaws that have a \( \Delta t_{\text{center}} \) such that \( \Delta t_{\text{Born}} \leq \Delta t_{\text{center}} < 2\Delta t_{\text{Born}} \) will be classified in F1-F3 feature space while those having a \( \Delta t_{\text{center}} \) such that \( \Delta t_{\text{center}} \geq 2\Delta t_{\text{Born}} \) will be considered in F1-F2 feature space. In order to be consistent, these restrictions were followed when choosing the MOOT data to do the training. Two sets of data, each for one class, were chosen to train the classifier in F1-F3 space as is seen in Table 5.1. The reasons for picking only two data sets are
(a) volumetric flaw (2-to-1 oblate spheroid)

(b) crack

Fig. 5.1. Angular position of the receiving transducer
Table 5.1. Data used to train the classifier in F1-F3 space

<table>
<thead>
<tr>
<th>source</th>
<th>flaw class</th>
<th>look angle $\Theta$ (deg.)</th>
<th>F1</th>
<th>F3</th>
<th>$\Delta t_{\text{center}}$ (usec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOOT</td>
<td>spheroid</td>
<td>20</td>
<td>-0.98</td>
<td>0.23</td>
<td>7x0.00978</td>
</tr>
<tr>
<td>MOOT</td>
<td>crack</td>
<td>30</td>
<td>-0.94</td>
<td>0.0</td>
<td>6x0.00978</td>
</tr>
</tbody>
</table>
1. there was only one crack data set that had $\Delta t_{\text{center}} < 2\Delta t_{\text{Born}}$
   available and
2. the chosen spheroid data has a feature vector $(F_1,F_3)$ which is the
   closest to the crack's and thus is expected to be most effective in
   training the classifier.

In this case, it took twenty three training trials before the decision
function converged to the line

(5.2) $0.5 F_1 + 2.53 F_3 = 0$ .

In the training, we decided not to guess a threshold constant, but
instead, we forced the decision function to go through the origin.
Since both classes are in the same quadrant of the $F_1$-$F_3$ space, choosing
a zero threshold might have slowed down the rate of convergence of the
decision function.

To test this decision plane (line), several data sets from either
experiment, M00T, or Kirchhoff approximation were used. Results were
tabulated in Table 5.2 and were also shown in Fig. 5.2. All but two
were correctly classified. Notice that in Fig. 5.2, both misclassified
feature vectors are located on the negative $F_3$ axis. These two
misclassification cases are experimental data which, for reasons not yet
understood, lost the high frequency oscillation in the leading part of
the "step function" and therefore have a normalized amplitude ratio
feature $F_3$ equal to zero.

For big flaws, i.e., those have a $\Delta t_{\text{center}} \geq 2\Delta t_{\text{Born}}$, it took only
three training trials to have the decision function converge to
Table 5.2. Data tested by the classifier in F1-F3 space

<table>
<thead>
<tr>
<th>source</th>
<th>guessed flaw class</th>
<th>look angle Θ (deg.)</th>
<th>F1</th>
<th>F3</th>
<th>Δt center (usec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>experiment</td>
<td>crack1</td>
<td>22.2</td>
<td>-0.96</td>
<td>0.0</td>
<td>4x0.00978</td>
</tr>
<tr>
<td>experiment</td>
<td>crack1</td>
<td>27.4</td>
<td>-1.0</td>
<td>0.0</td>
<td>5x0.00978</td>
</tr>
<tr>
<td>experiment</td>
<td>volumetric</td>
<td>3.9</td>
<td>-0.72</td>
<td>0.22</td>
<td>6x0.00978</td>
</tr>
<tr>
<td>MOOT</td>
<td>volumetric</td>
<td>0</td>
<td>-0.82</td>
<td>0.22</td>
<td>5x0.00978</td>
</tr>
<tr>
<td>MOOT</td>
<td>volumetric</td>
<td>5</td>
<td>-0.80</td>
<td>0.21</td>
<td>5x0.00978</td>
</tr>
<tr>
<td>MOOT</td>
<td>volumetric</td>
<td>15</td>
<td>-0.87</td>
<td>0.24</td>
<td>6x0.00978</td>
</tr>
<tr>
<td>MOOT</td>
<td>volumetric</td>
<td>25</td>
<td>-0.85</td>
<td>0.25</td>
<td>7x0.00978</td>
</tr>
<tr>
<td>experiment</td>
<td>crack</td>
<td>25</td>
<td>-0.97</td>
<td>0.0</td>
<td>4x0.00978</td>
</tr>
<tr>
<td>experiment</td>
<td>crack</td>
<td>30</td>
<td>-1.0</td>
<td>0.0</td>
<td>7x0.00978</td>
</tr>
<tr>
<td>Kirchhoff</td>
<td>crack</td>
<td>15</td>
<td>-1.0</td>
<td>0.0</td>
<td>4x0.00978</td>
</tr>
<tr>
<td>Kirchhoff</td>
<td>crack</td>
<td>20</td>
<td>-0.99</td>
<td>0.0</td>
<td>5x0.00978</td>
</tr>
<tr>
<td>Kirchhoff</td>
<td>crack</td>
<td>30</td>
<td>-0.93</td>
<td>0.0</td>
<td>7x0.00978</td>
</tr>
</tbody>
</table>

1 misclassification
Fig. 5.2. Classification space for F1-F3

+ volumetric flaws
x cracks
(5.3) 0.05 F1 - 1.35 F2 = 0.

Such a fast convergence of the decision function was expected since the two classes, volumetric flaw and crack, are located in separate quadrants of the F1-F2 space and the decision function was again chosen to go through the origin. Three separate data sets, tabulated in Table 5.3, were used for the training. For the testing, five spheroidal void data sets from experiment and five crack data sets from MOOT (since no crack data of this size from experiment was available) were considered. All of them were correctly classified as can be seen in both Table 5.4 and Fig. 5.3.

In general, these results show that the classification features and method we have used are working reasonably well. Figs. 5.2 and 5.3 show the advantages of dealing with only a few features so that a visualization of the classification results is possible. It would, of course, have been more satisfying if additional experimental data were available to test a wider range of flaw shapes and "noise" levels.
Table 5.3. Data used to train the classifier in F1-F2 space

<table>
<thead>
<tr>
<th>source</th>
<th>flaw class</th>
<th>look angle $\Theta$ (deg.)</th>
<th>F1</th>
<th>F2</th>
<th>$\Delta t_{\text{center}}$ (usec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOOT</td>
<td>spheroid</td>
<td>30</td>
<td>-0.92</td>
<td>0.99</td>
<td>9x0.00978</td>
</tr>
<tr>
<td>MOOT</td>
<td>spheroid</td>
<td>35</td>
<td>-1.0</td>
<td>1.24</td>
<td>10x0.00978</td>
</tr>
<tr>
<td>MOOT</td>
<td>crack</td>
<td>40</td>
<td>-0.87</td>
<td>-0.36</td>
<td>10x0.00978</td>
</tr>
</tbody>
</table>
Table 5.4. Data tested by the classifier in F1-F2 space

<table>
<thead>
<tr>
<th>source</th>
<th>guessed flaw class</th>
<th>look angle θ (deg.)</th>
<th>F1</th>
<th>F2</th>
<th>Δt&lt;sub&gt;center&lt;/sub&gt; (usec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>experiment</td>
<td>volumetric</td>
<td>43.2</td>
<td>-1.0</td>
<td>1.40</td>
<td>9x0.00978</td>
</tr>
<tr>
<td>experiment</td>
<td>volumetric</td>
<td>70.5</td>
<td>-1.0</td>
<td>1.70</td>
<td>18x0.00978</td>
</tr>
<tr>
<td>experiment</td>
<td>volumetric</td>
<td>48.1</td>
<td>-0.86</td>
<td>0.52</td>
<td>9x0.00978</td>
</tr>
<tr>
<td>experiment</td>
<td>volumetric</td>
<td>27.4</td>
<td>-1.0</td>
<td>0.81</td>
<td>8x0.00978</td>
</tr>
<tr>
<td>M00T</td>
<td>crack</td>
<td>48.1</td>
<td>-1.0</td>
<td>0.62</td>
<td>11x0.00978</td>
</tr>
<tr>
<td>M00T</td>
<td>crack</td>
<td>50</td>
<td>-0.82</td>
<td>0.60</td>
<td>13x0.00978</td>
</tr>
<tr>
<td>M00T</td>
<td>crack</td>
<td>60</td>
<td>-0.75</td>
<td>-0.66</td>
<td>13x0.00978</td>
</tr>
<tr>
<td>M00T</td>
<td>crack</td>
<td>70</td>
<td>-0.71</td>
<td>-0.76</td>
<td>13x0.00978</td>
</tr>
<tr>
<td>M00T</td>
<td>crack</td>
<td>80</td>
<td>-0.74</td>
<td>-0.57</td>
<td>13x0.00978</td>
</tr>
<tr>
<td>M00T</td>
<td>crack</td>
<td>90</td>
<td>-0.70</td>
<td>-0.91</td>
<td>15x0.00978</td>
</tr>
</tbody>
</table>
Fig. 5.3. Classification space for F1-F2

+ volumetric flaws
x cracks
6. DISCUSSION AND CONCLUSIONS

As we have seen in the previous chapters, it is possible to develop a methodology for flaw classification that is simple yet effective. Although the number of training and test samples available for validating these methods were limited (as they always seem to be in NDE problems - flaws are scarce!) there is enough "elasticity" in the methods to allow them to succeed even in more difficult cases. To guarantee such success, however, it may be necessary to examine these techniques more closely. Some areas for future investigation include:

1. Models

The model-based features we are using have been based on the assumption that the flaw surfaces are smooth. The effect that surface roughness has on obtaining these features, therefore, definitely needs consideration. The choice of leading edge features is probably a plus in this respect because it is likely that these features are considerably more robust than other features (such as creeping waves, Rayleigh waves, resonances, etc.) in the presence of surface roughness. However, this assumption needs verification. Also, it should be pointed out that our models assume an isotropic medium surrounding the flaw. The effects of material anisotropy on these features as found, for example, in composites or welds also needs attention. Finally, we note that the surface integral Born series, that was presented here only to the first order approximation, should be extended to higher order terms. Even if this extension can only be done numerically, it may be possible to use the higher order terms to develop a rapid means for solving the
direct scattering problem. This type of solution capability would be considerable value in the treatment of inverse scattering problems, for example.

2. Signal Processing

The success of the flaw-derived deconvolution and low frequency extrapolation methods have been demonstrated. Particularly in the area of deconvolution, however, there are a number of extensions and improvements that could be considered. First, it should be pointed out that the deconvolution procedures described in chapter three have only been validated on experimental signals from volumetric flaws. Although simulation studies have suggested that these same procedures can apply to cracks, verification of this has not been possible because of a lack of "real" cracklike flaws. Second, the Hanning filter step in the deconvolution procedure needs further examination. The amount of assistance that this filter provides in isolating the leading edge flaw reference pulse is rather minimal and the choice of other filters may be better.

3. Adaptive Learning

In the present work the classification process only considered two classes of flaws - volumetric flaws and cracks. The features we used, however, are sufficient to easily change the problem output to three classes - high impedance volumetric flaws, low impedance volumetric flaws, and cracks. Since these three classes are (theoretically) linearly separable in the current feature spaces, the only modification necessary to our adaptive learning procedure for accomplishing this
extension is the use of more than one decision surface (plane). To
verify such a three-category classification problem it will be necessary
(again) to have a good set of "real" high and low impedance flaws to
consider. Such a training/test set, unfortunately, is not currently
available. The sets we did use for our two-category classification
problem were, however, much more extensive than those used in previous
adaptive learning classification experiments [6]. This was possible
because our flaw features are all fundamental, i.e., based on models,
and "exact" numerical model results were available for testing purposes.

In summary, we wish to emphasize that the current classification
methodology is based on a different philosophy than found in most
adaptive learning approaches. Instead of employing large numbers of
loosely defined "features" and allowing the system to sort out which, if
any, of those features are important, we have chosen, as indicated
previously, to rely instead on very few but fundamental features. When
employed "correctly", i.e., on those NDE problems where the basic
assumptions of the method are not grossly violated, we feel that this
approach has considerable merit. However, we also note that the recent
development of large self-modifying neural networks, containing new and
powerful learning algorithms [84], makes the reexamination of the more
traditional adaptive learning methods also of considerable interest.
7. REFERENCES


I would like to express my sincere appreciation and gratitude to my major professor Dr. Lester W. Schmerr. With his great understanding and insight in various fields such as Ultrasonic Nondestructive Evaluation, Artificial Intelligence, and Signal Processing, he has supported and guided me patiently through this research. For his invaluable effort, I am deeply grateful.

I would also like to thank Mr. Samuel J. Wormley and Dr. Timothy A. Gray who provided me with both experimental and numerical scattering data.

I would also like to direct my appreciation to Professors Donald O. Thompson, R. Bruce Thompson, John P. Basart, Harry L. McMaken, and Steve F. Russell for offering advice and serving on my committee. The financial support of both the Department of Engineering Science and Mechanics and the Iowa State Center for Nondestructive Evaluation throughout this research is acknowledged with gratitude.

Finally, to my wife Patricia, I am specially thankful. Her understanding, support, encouragement, and love have helped me continuously to work hard with great enthusiasm.
9. PROCEDURES FOR IDENTIFYING THE LEADING EDGE PULSE AND CENTER OF A FLAW SIGNAL

To identify the leading edge pulse and center of a flaw signal, we have developed procedures which assume the response is in the time domain and the following primary quantities are known.

1. A measure of the noise level $N$ and signal-to-noise ratio $SNR$.
2. Locations and values of all maxima and minima above a threshold (we use 20% of the magnitude of the biggest pulse as the threshold in the present work).
3. Estimate of $\Delta t$ of oscillations due to limited bandwidth and $\delta_2/\delta_1$ for amplitude of oscillations (see Fig. 9.1).

Based on these primary quantities, we can identify a candidate-for-the-leading-edge-pulse (clep) in the signal which is defined as the location and value of the first extremum where a larger value does not exist within a time interval $\alpha \Delta t$. The "$\alpha$" here is a "fudge" factor and has a value of 1.42 in the present work. After the clep is found, we can follow the procedures given in Figs. 9.2 and 9.3 to search for the leading-edge-pulse (lep) and then the center of the flaw can be identified based on the lep (see Fig. 9.4). In Figs. 9.2 to 9.4, all the prepositions (P1:, etc.) and conclusions (C1:, etc.) are listed in the following.

P1: the value of the clep is positive.

P2: there is no previous negative minimum within $\alpha \Delta t$ of the clep whose magnitude $\geq \delta_2/\delta_1$ times the magnitude of the clep.

P3: there is no signal $> \delta_2/\delta_1$ times the magnitude of the clep
Fig. 9.1. Oscillations of a sinc function which has a frequency range 0–20 MHz
Fig. 9.2. Procedures for finding lep when clep is positive
Fig. 9.3. Procedures for finding lep when clep is negative
Fig. 9.4. Procedures for estimating the flaw center
outside of a time range $\omega_\Delta t$ from the clep.

P4: there is moderate to strong evidence of a linearly increasing amplitude in the frequency domain.

P5: there is a previous negative minimum within $\omega_\Delta t$ of the clep whose magnitude $> (\delta_2/\delta_1 + m_1)$ times the magnitude of the clep where $m_1$ is a "noise" term.

P6: there is a previous positive maximum within $\omega_\Delta t$ of the clep whose magnitude $> (\delta_2/\delta_1 + m_2)$ times the magnitude of the clep where $m_2$ is a noise term.

P7: the value of the lep is negative.

P8: a minimum occurs in the once integrated signal within a time interval $\leq \omega_\Delta t$ from the first-zero-crossing before the lep.

C1: the previous extremum is the lep.

C2: the clep is the lep.

C3: the center of the flaw is at the first extremum of the twice integrated signal.

C4: the center of the flaw is at the first minimum of the once integrated signal.
10. COMPUTER PROGRAMS

10.1 Programs for Visual Calculator

10.2 Programs for Expert System