1990

Model-based ultrasonic flaw classification and sizing

Chien-Ping Chiou

Iowa State University

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

Part of the Mechanical Engineering Commons

Recommended Citation


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

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

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

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

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

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

Chiou, Chien-Ping, Ph.D.

Iowa State University, 1990
Model-based ultrasonic flaw classification and sizing

by

Chien-Ping Chiou

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

1990
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>GENERAL INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>PART I. A QUASI-PULSE-ECHO TECHNIQUE FOR ULTRASONIC FLAW CLASSIFICATION</td>
<td>6</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>8</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>9</td>
</tr>
<tr>
<td>CLASSIFICATION TECHNIQUES</td>
<td>12</td>
</tr>
<tr>
<td>THEORETICAL PREDICTIONS</td>
<td>18</td>
</tr>
<tr>
<td>EXPERIMENTAL VERIFICATIONS</td>
<td>25</td>
</tr>
<tr>
<td>SUMMARY AND DISCUSSION</td>
<td>31</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>33</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>34</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>41</td>
</tr>
<tr>
<td>PART II. NEW ASPECTS OF MODEL-BASED ULTRASONIC FLAW SIZING</td>
<td>63</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>65</td>
</tr>
<tr>
<td>INVERSE SCATTERING MODEL</td>
<td>67</td>
</tr>
<tr>
<td>CORRECTION OF BANDWIDTH EFFECT IN CRACK DATA</td>
<td>70</td>
</tr>
<tr>
<td>EQUIVALENT FLAW SIZING</td>
<td>71</td>
</tr>
<tr>
<td>SIZING AND CLASSIFICATION</td>
<td>76</td>
</tr>
<tr>
<td>SPHERICAL HARMONICS EXPANSION ALGORITHM</td>
<td>79</td>
</tr>
<tr>
<td>PART III. AN ADAPTIVE NEURAL NETWORK MODEL FOR FLAW SIZING IN NONDESTRUCTIVE EVALUATION</td>
<td></td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td></td>
</tr>
<tr>
<td>ULTRASONIC FLAW SIZING PROBLEM IN NDE</td>
<td></td>
</tr>
<tr>
<td>NEURAL NETWORK MODEL</td>
<td></td>
</tr>
<tr>
<td>ADAPTIVE TRAINING ALGORITHM</td>
<td></td>
</tr>
<tr>
<td>NETWORK TRAINING AND SIZING RESULTS</td>
<td></td>
</tr>
<tr>
<td>CONCLUSION</td>
<td></td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td></td>
</tr>
<tr>
<td>REFERENCES</td>
<td></td>
</tr>
</tbody>
</table>
GENERAL INTRODUCTION

Defects exist in all structural materials. Some defects are benign while others are harmful and may cause serious safety problems. Thus, based on both safety and economic considerations, it is important to be able to detect and characterize flaws without damaging the host materials. The purpose of this research effort is to develop such nondestructive evaluation (NDE) techniques for flaw characterization. Particularly, we are concerned with the development of more advanced and reliable methods to do ultrasonic flaw classification and sizing where we obtain the type, size, shape, and orientation information of a single isolated flaw from ultrasonic immersion or contact measurements.

Ultrasonic Flaw Classification

Obtaining information on flaw type falls into the flaw classification category. Making distinctions between harmless and dangerous defects has great economic value, since this knowledge can extend the service time of structures or parts even if a flaw has been detected. The correct flaw type information also allows one to choose the appropriate algorithm for flaw sizing and hence improve the sizing accuracy.

Many current ultrasonic classification methods are heuristic or empirical, and often lacks of solid theoretical support [1]. Recently, however, more quantitative techniques have appeared in the literature. Among those new approaches are ray tracing and signal amplitude analyses [2], the use of pattern recognition [3-4], and the application of expert systems [5].

In the first topic of this work, we present a new flaw classification technique that is both quantitative and simple. It employs the time separation
and amplitude difference of mode-converted diffracted signals in a quasi-pulse-echo configuration to distinguish smooth vs. sharp-edged flaws. Both experimental and theoretical results are obtained that validate this method. A discussion is also given of the ways in which this technique might be practically implemented in the field.

**Ultrasonic Flaw Sizing**

In ultrasonics, obtaining the flaw size, shape and orientation information is known as the solution of an inverse problem of elastic wave scattering. Present field inspections have often simply relied on the comparison of the amplitude of flaw signals with those of standard references, like flat-bottom holes, of known sizes [6-7]. These methods, however, are not quantitative enough to provide the information needed for modern fracture mechanics studies and do not meet the demand for more strict safety and reliability guarantees. Thus, one has to appeal to more detailed inverse solutions. Unfortunately, due to complicated flaw geometries and mode coupled boundary conditions, exact theoretical inverse scattering solutions are extremely difficult to obtain. Hence, much attention has focused on the use of approximate techniques like geometrical ray methods [8], the Kirchhoff approximation (for planar cracks) [9] and the Born approximation (for volumetric flaws) [10]. Based on these techniques, a few simple models have received detailed study and several inverse algorithms have been developed [11-13].

Practically, one of the major obstacles in these model-based methods, especially for sizing small volumetric flaws, is the “zero-of-time” problem; that is, the determination of a reference point for aligning scattering data. Although some solutions to this problem do exist [14], it may not be possible always to
locate such references for volumetric flaws, due to the interference of different wave modes such as creeping waves. Fortunately, for isolated cracks, this problem does not arise [13].

Mathematically, inverse problems are also often ill-posed and non-unique [15-16]. One practical way around these difficulties is through smoothing assumptions and the collection of large amounts of scattering data. Such approaches are then essentially flaw imaging techniques [17-18], which usually also need time-consuming data processing. If an imaging approach is not practical, then other alternatives need to be considered. One approach is to require that not all the details of the flaw be resolved, but that instead an estimate be made of a simple flaw shape which best fits the available scattering data and is able to represent the major aspects of the flaw. This type of constrained inverse scheme is called “equivalent flaw sizing”. Equivalent flaw sizing makes sense from a practical standpoint since it is often not possible to use more detailed flaw size information, even if it is available, in existing codes and standards. Previous work has considered fitting isolated volumetric flaws with ellipsoids [12] and planar flaws (cracks) with ellipses [13], respectively, using separate non-linear least squares approaches. Later, both schemes were unified [19] into a single nonlinear least squares algorithm.

In the second subject of this dissertation, we present a new equivalent flaw sizing algorithm where the sizing method reduces to solving both a linear least squares problem and a standard eigenvalue problem. This method represents an important advance over previous algorithms [12-13], since it is noniterative and avoids the solution of a difficult nonlinear least squares problem. It also performs as well as previous methods on both theoretical and
experimental data. We also describe another practical model-based methodology that uses exactly the same scattering data but is not strictly limited to ellipsoidal flaw shapes. This method involves directly fitting the flaw surface geometry in terms of a series of spherical harmonics. Finally, we include a discussion of the effects of classification information on the equivalent flaw sizing problem, and the correction of scattering data errors for cracks due to the finite bandwidth of the interrogating probe.

Another approach for solving flaw sizing problems that we have considered makes use of an artificial neural network [20]. Such network models, consisting of layers of processing cells, have been shown capable of inverting rather arbitrary mappings from incomplete or noisy data [21-22]. The inversion is carried out through the "learning" acquired by the presentation of input-output examples.

In the last topic of this research work, we demonstrate the applicability of a neural network model for our equivalent sizing problem. Here a multilayered feed-forward network is trained by the backpropagation algorithm and the generalized delta rule [23] to capture the mapping between input parameters (directly obtainable from ultrasonic experiments) and equivalent flaw parameters for an isolated crack. It is shown that the network can be trained on noise-free theoretical data, and then used directly to obtain equivalent flaw size and orientation values from experimental data. One of the disadvantages of a neural network is that the training times can be prohibitive when using standard algorithms. Thus, we also demonstrate an enhanced adaptive training scheme that can make the training process more efficient.
Explanation of Dissertation Format

This dissertation is written in an alternative format in compliance with the regulation of the thesis office of Iowa State University. The dissertation includes a general introduction followed by three parts and a general summary.

These parts contain three different topics of research work that either have already been published or will be submitted for publication. Part one presents a new experimental technique for ultrasonic flaw classification and will be submitted to Ultrasonics. The second part, which describes new methods and results for equivalent flaw sizing, is an extension of a paper which appeared in Review of Progress in Quantitative Nondestructive Evaluation, 9A, pp. 117-124. This extended paper will be submitted to the Journal of the Acoustical Society of America. The third part discusses the use of an adaptive neural network model for solving ultrasonic crack sizing problems. This last paper will be submitted to Communications in NDE. Finally, a general summary is given at the end of this dissertation.
PART I. A QUASI-PULSE-ECHO TECHNIQUE
FOR ULTRASONIC FLAW CLASSIFICATION
A QUASI-PULSE-ECHO TECHNIQUE
FOR ULTRASONIC FLAW CLASSIFICATION

Chien-Ping Chiou and Lester W. Schmerr

Center for NDE and
Department of Engineering Science and Mechanics
Iowa State University
Ames, Iowa 50011
ABSTRACT

One of the important flaw characterization tasks in the field of ultrasonic nondestructive evaluation (NDE), is to provide flaw type information by analyzing the flaw responses acquired during an inspection. Here, we present a new quasi pulse-echo ultrasonic classification technique that utilizes the time separation and amplitude difference of mode-converted diffracted signals to distinguish between smooth vs. shape-edged flaw geometries.

Experiments with cylindrical cavities, surface-breaking fatigue cracks and slag inclusions have been used to test the practicality of this approach. All results of these tests show good consistency in the separation of smooth vs. sharp-edged flaws provided that the signal-to-noise ratio is sufficient. Furthermore, the scattering feature used for classification in this method is also verified by detailed elastodynamic scattering calculations.
INTRODUCTION

Recent safety problems in the aviation and nuclear industries have greatly emphasized the need for more reliable ultrasonic nondestructive evaluation (NDE) techniques for locating and characterizing flaws in materials. In the flaw characterization process, it is important to obtain quantitative information on both the flaw type (classification) and the flaw size so that fracture mechanics calculations can be performed to predict the remaining service life of a part. Here we will concentrate on the problem of flaw classification only by analyzing the flaw responses acquired from ultrasonic inspections.

Ultrasonic classification methods usually rely on either time domain or transformed (usually frequency) domain data. A variety of signal features are extracted from these data and then used as the basis for the classification process. Signal features commonly recognized by the ultrasonic community include time domain waveform characteristics, time-of-flight measurements, time/frequency domain amplitude ratios and frequency spectrum analysis [1-3]. Many of these features are of a qualitative and empirical nature. In the distinction between crack-like and volumetric flaws, for example, it is known that defects of volumetric type tend to possess longer pulse durations than crack-like defects, and exhibit different waveform “dynamics” during a transducer scan. In addition to the above deterministic approach, the classification task can be also carried out from a statistical standpoint as described in the work of [4]. Also note that in [5] an attempt was made directly to correlate the defect shape, its stress concentration factor and the time signal amplitude ratio.

In the last decade, significant progress has been made in combining signal processing techniques and artificial intelligence concepts for classification
analysis. In [6-7], an integrated feature-based approach using signal processing and pattern recognition schemes was taken and later extended to weldment inspection. A similar effort can be found in [8]. Mucciardi and his co-workers developed an adaptive learning network [9-10] that combines a statistical approach and the application of a neural network-like model to NDE classification problems. More recent research along this line was the work of [11]. Another artificial intelligence approach involves the use of a rule-based expert system. The expert system FLEX, developed by Schmerr et al. [12], is an example of such a system which has been able to classify general volumetric vs. crack-like defects by using only limited ultrasonic data.

The generality and reliability of a classification method are highly affected by the availability of various signal features. Standard techniques, however, have rarely employed features beyond specularly reflected signals. This is unfortunate since diffracted responses, particularly from sharp-edged defects, contain significant information useful in the classification process. Not until recently have advanced techniques [13-14] emerged to justify the use of these features. Even so, the potential of using diffracted signal features has not yet been fully exploited.

In this work, we present a new feature that utilizes the mode-converted diffracted/scattered data in a near pulse-echo configuration to classify the local defect geometry into one of two general classes: (1) crack-like flaw or (2) volumetric flaw. These two classes of defects are modeled here by two canonical geometries, namely, a circular cavity and a through-thickness surface-breaking crack. Theoretical predictions for classifying these two model scatterers are examined, and are further verified by experiment.
The rest of the paper is organized as five sections. In the first section, the new classification concept is proposed and compared with three existing techniques. The second section presents theoretical model calculations based on the elastodynamic Kirchhoff approximation and Wiener-Hopf diffraction solutions. The third section describes the experimental specimens, test procedures and compares the experimental results with the theoretical predictions. In the fourth section, conclusions and a discussion for future work are addressed. Finally, more detailed derivations of the analytical Kirchhoff model used in the second section are given in an appendix.
CLASSIFICATION TECHNIQUES

The classification techniques considered here are based on the principles of elastic wave propagation in solids. Particularly, we consider the presence of mode-converted scattered or diffracted wave phenomena. In the discussions below, the general signal patterns from flaws with either a smooth surface or a sharp edge are examined. Next, a survey of three classification techniques utilizing diffracted/scattered responses follows. The new classification concept is then introduced and a possible implementation of this concept is discussed.

Flaw Signal Features

When an ultrasonic pulse impinges on a defect, its interaction with the defect results in a series of scattered signal trains. These signals represent a set of "signatures" to be used for identifying the defects. Typical ultrasonic signals from a planar crack-like defect are illustrated in Fig. 1a as obtained in an immersion or contact testing. The primary signals consist of reflected responses, surface traveling waves, edge diffracted waves and head wave components (not shown). If the crack surface is relatively flat and smooth, the specularly reflected wave is usually the strongest available signal. In contrast, the cylindrically spreading edge-diffracted responses ("flashpoints") are relatively weaker than the reflected response. Note that mode conversions can occur at both the crack face in the reflected wave case and the crack edge in the diffracted wave case. As shown in Fig. 1, and in all subsequent sections, we will denote a longitudinal wave by L and a transverse shear wave by T.

For a volumetric void-like defect, the returned signal pattern can be depicted by Fig. 1b. In pulse-echo scanning the signal pattern contains only a
strong reflected wave of the same mode as the incident wave while in the pitch-catch case additional mode-converted reflections are present. Also present is the weaker "creeping" wave that travels around the void's periphery and radiates back to the host medium. The fundamental feature difference between a crack-like flaw and a volumetric flaw, as we now observe, is that in pulse-echo testing a significant mode-converted diffracted wave component can exist for a crack-like defect which does not exist for a volumetric defect. In the following, we will see how this difference plays an important role in the ultrasonic flaw classification problem.

**Related Techniques**

The so-called satellite-pulse technique (SPT) developed by Gruber [13] has appeared in the literature for some time. Based on a pattern recognition scheme, this single-probe classification method makes use of the signal features (as discussed above) along with time-of-flight and amplitude measurements. For a volumetric flaw, the specularly reflected echo and the weaker "satellite" creeping pulse are identified by scanning and used for both sizing and classification (recall Fig. 1a); for an internal planar crack, the diffracted waves from both crack tips and the specularly reflected wave from the crack face are used instead (Fig. 1b). If the defect is of a surface-breaking type such as a fatigue crack, then the reflected response from the crack/surface corner replaces the feature of the second tip-diffracted signal. Note that some features obtained from the frequency modulated spectra of internal cracks are also used in this method.

In contrast, a two-probe classification method based on a variation of Delta Technique was proposed in [14]. The Delta Technique [15] utilizes a receiving
transducer positioned normal to the specimen surface and one or more transmitting transducers sending ultrasonic angular beams into the specimen. Thus the receiver, transmitters and the tested spot form a triangular (delta) geometry. The "variation" of the standard delta technique in [14] refers to interchanging the receiver and transmitter positions and incorporating the use of diffracted/head wave signals. As Fig. 2a illustrates, for a smooth volumetric flaw, both signals $T_1$ and $T_{\text{diff}}$ are reflected waves and $T_1$ reflected from the flaw-top is stronger than $T_{\text{diff}}$ reflected from the backwall. In contrast, in Fig. 2b $T_1$ is the crack top-edge diffracted signal, $T_{\text{latt}}$ is the crack-bottom "head-lateral" wave generated by the redirected backwall reflection and $T_1$ is weaker than $T_{\text{latt}}$. Hence, by using signal ratios, the classes of flaw types can be separated.

Another two-probe redirected-reflection scheme named the LLT-technique was recently proposed [16]. Fig. 3 shows the basic idea of classification using LLT inspection. Note that this technique does not use the feature of a diffracted wave but applies reference signals of a calibration reflector such as a flat-bottom hole or a side wall. The classification factor, CF, for crack-like defects generally had values of less than 10, while small volumetric flaws gave a CF above 10 (Fig. 3). However, for volumetric flaws of increasing size, this method tends to be misleading. It should be pointed out that the present LLT-technique is essentially similar to the work of [17] which was published earlier.

The New Concept

Consider an incident pulse of $T$ type launched in a pulse-echo setup as shown in Figure 4. If the unknown flaw is of smooth volumetric type such as a void, inclusion or piece of slag, the local flaw geometry near the position where the incident wavefront first strikes the flaw can be approximated as a planar
specular reflector parallel to the wavefront. Elastodynamics theory then predicts that \( T \), the amplitude of the specularly reflected \( T \) wave, is large and no mode conversion occurs, i.e., the amplitude of the mode-converted reflected \( L \) wave (as denoted by \( L \)) is zero. However, for a sharp edge reflector such as a crack, \( T \) instead is the amplitude of the edge-diffracted \( T \) wave, and \( L \) is the amplitude of the mode-converted edge-diffracted \( L \) wave. Then by the geometric theory of diffraction, (as will be discussed in next section) for a crack the magnitude of \( L \) is expected to be of the same order as \( T \) provided that the incident wave direction lies in an optimal range. Thus, the existence, in pulse-echo measurements, of a significant amplitude ratio \( L/T \), of mode-converted \( L \) vs. same-mode \( T \) (or, similarly, \( T/L \) ratio for incident \( L \) wave) can be used to classify unknown flaws. To the authors' best knowledge, this is the first attempt to use diffracted and mode-converted responses in such a manner for classification purposes.

In this new method, the distinction between a smooth surface vs. a sharp edge is on the basis of point-wise information, i.e., detailed scanning of a transducer, as in the satellite pulse method, is not required. Also, the method does not rely on the existence of an auxiliary scattering surface (e.g., backwall). Finally, although the mode-converted \( L \) wave is small in amplitude, it is the first-arriving signal at the transducer so, with a sufficient signal-to-noise ratio, it is easily identified from the other waves.

**Implementation Considerations**

To implement the proposed classification concept, we need to examine the optimal ways to generate and send an ultrasonic beam into the specimen and to receive the responses through some device at the interface (Fig. 4). For effective flaw interrogation, an obliquely incident beam is often chosen. If the probe
device has an impedance mismatch with the specimen, secondary mode conversions occur at the interface for both returned L and T signals resulting in an additional pair of T and L signals. Either of these two pairs of signals could then in principle be received by a properly oriented longitudinal/shear probe combination. If we only consider using longitudinal receiving probes then two basic approaches are possible. First, a single transducer (or phased array) could be operated in the pulse-echo mode to receive both L' and L'' signals (see Fig. 4) at oblique angles. Alternatively, two probes could be used in a setup with one serving as transmitter/receiver for the same-mode pulse-echo L' signal and the other being the receiver for the mode-converted pitch-catch L'' signal.

From previous work on ultrasonic crack detection, a shear wave with an incident angle between 40° and 50° has been shown (see, for example, [18]) to give good detectability. Based on these considerations, in this work we decided to use oblique shear waves in this range to test the L/T ratio classification concept. The first approach, which was unsuccessful, (Fig. 4), was implemented in an immersion mode, and the second approach in contact mode (Fig. 5a).

After further examination, however, the second approach was found also difficult to implement due to geometrical and material constraints. In order to mount two minimally spaced probes on the same wedge, a certain impedance difference between the wedge and the specimen is required so that L' and L'' can be adequately separated in angle. Choosing a metal wedge that is made of a material different from the specimen would require unrealistically large wedge thickness; whereas plastic materials provide enough angular separation with less thickness, but possess much higher attenuation. Thus, some compromise is necessary. To reduce the attenuation in the plastic wedge, we relaxed the
advantage of a pulse-echo geometry by replacing the single wedge with two wedges and positioned one probe beside the other one. As shown in Fig. 5b, one probe mounted on wedge I serves as transmitter and receiver (T/R₁) of T waves in the specimen while the other probe mounted on wedge II is used as a second receiver (R₂) of mode-converted L waves. This dual wedge-probe system was then operated in a near pulse-echo configuration (Fig. 6). The two scanning paths (L and T) were designed to "focus" on prescribed flaw locations in our steel specimens. For different samples, the setup will have to be adjusted. However, the main purpose here was to verify the new concept, and not to account for general situations. The experimental details of the use of this dual wedge system is given in a later section. It will be shown that since the opening angle θ (Fig. 6) can be made quite small the original pulse-echo concept is still applicable in this quasi-pulse-echo configuration.
THEORETICAL PREDICTIONS

As described in last section, the contact experiment will be conducted in a near pulse-echo geometry in which L and T signals are separated by an opening angle $\theta$ (Fig. 6). In this section, we investigate the scattering problems of volumetric and crack-like flaws due to an incident plane shear wave in such a configuration. Particularly, we concentrate on the theoretical response of two model scatterers, namely, a cylindrical cavity and a semi-infinite crack.

Cylindrical Cavity and the Elastodynamic Kirchhoff Approximation

The problem of elastic shear wave scattering by a cylindrical cavity has received considerable previous attention [19-22]. The exact solution for an incident plane harmonic shear wave is available by the separation of variables method [19]. This solution includes direct specularly reflected responses and other components such as creeping waves. The time domain version can then be obtained via analytical transform methods or numerical means such as the fast Fourier transform (FFT). Direct time domain solutions due to an ideal delta function incident pulse can also be considered [20]. For our case, since the early arriving specular reflected response is of primary interest and cannot be easily separated out from other components of the above mentioned solutions, we instead employ the elastodynamic Kirchhoff approximation (which hereafter is abbreviated as EKA) directly to extract specularly reflected responses from the solution.

The Kirchhoff approximation originated in the field of geometrical optics [23]. By observing the close resemblance between high frequency elastic wave motion and geometrical optics, we can adapt the same idea to elastodynamics with minor modifications. Kirchhoff approximation basically states that at high
frequencies the scatterer surface is split into an illuminated side and a dark side. The illuminated side is considered locally to act as a perfect plane reflector (Fig. 7) such that the scattered wave field there can be approximated by the reflected wave field from such a plane surface. On the dark side shadow, the total field is assumed zero. This assumption, known as the zero-th order approximation, is in fact the lowest term of a more general uniform crack opening displacement (COD) theory [24]. The major disadvantage of the zero-th order EKA is that for a planar scatterer it lacks accuracy at observation angles far from the specular direction [25]. Since we will apply the EKA here to volumetric scatterers in a near pulse-echo configuration, the EKA approximation is quite adequate. In fact, when only the specularly reflected contribution of the EKA is used, as considered below, the EKA is exact.

The EKA has been extensively used in calculating the diffraction of planar cracks of simple shapes [25-26] and in solving inverse problems [27]. However, less attention has been paid to applying it for volumetric scatterers. In this work we present an integral formulation that is valid for a generally smooth, convex 2-D cavity of arbitrary shape and is readily extended to a three-dimensional geometry. The boundary line integral is then evaluated via the stationary phase method in the high frequency limit. Here the ratio of L to T signals in Fig. 7 is considered by using the results in the Appendix. More details of the derivation can be found there.

From cases (3) of eq. (A-23) in the Appendix, the T→T far-field scattered displacement in the backscattered direction under EKA can be expressed as

\[
[u_{sc}^{TT}]_{\theta=0} \sim -A \left( \frac{r}{2y} \right)^2 \exp[ik_T(y - 2r)] Y^T
\]
which corresponds to signal T. Here \( \hat{Y}^T \) is the shearing polarization unit vector perpendicular to \( \hat{Y} \) (see Fig. A-1) and \( r \) is the cylinder radius. As expected, eq. (1) indicates a cylindrical spreading wave scattered from the cavity with a decay rate of \( y^{-1} \). From cases (4) of (A-23), the T→L displacement corresponding to signal "ray" L (making an angle \( \theta \) with respect to signal "ray" T) is

\[
(2) \quad u^{sc/L} \sim A \left( \frac{r \sin \phi}{\sin \theta} \right) \cos(\theta - \phi) \exp \left[ ik_L \left( y - r \frac{\sin \theta}{\sin \phi} \right) \right] R^T_L (\phi) \hat{Y};
\]

\[ \phi = \tan^{-1} \frac{\sin \theta}{\kappa + \cos \theta}. \]

Note that eq. (2) contains an additional trigonometrical factor and the T→L displacement reflection coefficient. It is known that the signal amplitudes in frequency and time domains are equivalent for a linear time-invariant system. Hence for measured responses at equal distances the L/T time signal ratio is readily computed from (1) and (2) as:

\[
(3) \quad \frac{L}{T} = \left| \frac{u^{sc/L} (\theta \geq 0^\circ)}{u^{sc/T} (\theta = 0^\circ)} \right| \sim \left( \frac{2 \sin \phi}{\sin \theta} \right)^2 \cos(\theta - \phi) R^T_L (\phi); \quad \phi = \tan^{-1} \frac{\sin \theta}{\kappa + \cos \theta}.
\]

The reflected L response becomes an inhomogeneous wave on the void's periphery when the reflected L wavefront is parallel to the normal at the specular point. With \( \kappa = c_L/c_T = 1.83 \), this occurs at a critical opening angle \( \theta_{cr} = 123^\circ \) and hence eq. (3) is no longer valid beyond \( \theta_{cr} \). The L/T amplitude ratio (eq. (3)) and T→L displacement reflection coefficient are plotted vs. opening angle \( \theta \), with \( \theta \) varied from \( 0^\circ \) to \( \theta_{cr} \) in Fig. 8. We see that EKA prediction starts at zero at \( 0^\circ \), reaches a maximum at \( 62^\circ \) and drops to zero again at \( \theta_{cr} \). Also note
that the displacement reflection coefficient dominates the ratio value in the small angle region ($\theta < 15^\circ$) where the geometrical factor in eq. (3) is close to 1.

Half-Plane and the Wiener-Hopf Diffraction Solution [28]

The model scatterer we used in this work for a crack-like flaw is the semi-infinite crack (half-plane) as depicted in Fig. 9. This semi-infinite crack model is legitimate for a general flat crack if the crack edges are sufficiently far apart so that each edge can be considered to respond independently.

Solving mixed boundary value problems as encountered in diffraction elastodynamics, is usually very difficult. For the case of the semi-infinite crack, fortunately, an exact solution is available by employing the classical Wiener-Hopf technique. In the far field, results can then be derived asymptotically via the method of steepest descents as described in [29]. Its validity breaks down at boundaries of shadow (transmitted wave) zones and at boundaries of specularly reflected waves. In these regions, solutions can be obtained by using the uniform asymptotic expansion theory or EKA in specularly reflected directions. Solutions to this canonical problem are actually the "building blocks" for more generalized problems to include curved edges and curved incident waves as extended by Keller's geometrical theory of diffraction (GTD) [30]. The complete derivation is quite involved. For full accounts the reader is referred to [24] and [29]. Here we merely provide the expressions for T→ T and T→ L cases which are considered in this paper.

The far-field diffracted shear potential due to an incident plane shear wave (T→ T) measured at distance $y$ is given by

\[ \Psi^{\text{diff}} = D^{t}_{T}(\alpha, \beta) \left( \frac{\lambda_T}{y} \right)^2 \exp(ikTy) \]
and the T→L longitudinal potential is

\[ \Phi^{\text{diff}} = D^T_T(\alpha, \beta) \left( \frac{\lambda_T}{y} \right)^{\frac{1}{2}} \exp(ik_Ly). \]

Where \( D^T_T(\alpha, \beta) \) and \( D^L_L(\alpha, \beta) \) are the T→T and T→L "diffraction coefficients" respectively and \( \lambda_T = 2/k_T \) is the T wave length. \( \beta \) is the incident angle measured from the crack face and \( \theta \) is the opening angle as before. Owing to the symmetry, the range \( 0 \leq \beta \leq \pi \) is sufficient. Angle \( \alpha \) is the observation angle, which is equal to \( \beta \) for pulse-echo T→T scanning. For the pitch-catch L→T case, \( \alpha \) equals to \( \beta + \theta \) if \( \beta + \theta \leq \pi \) or \( 2\pi - \beta - \theta \) if \( \beta + \theta \geq \pi \). In the pulse-echo mode, \( D^T_T(\alpha, \beta) \) can be explicitly expressed as

\[ D^T_T(\beta, \beta) = \exp\left(\frac{i\pi}{4}\right) \frac{k_T^2 \sin^2\left(\frac{1}{2}\beta\right)}{4\pi(k^2_T - k^2_L) \cos \beta (k_o - k_T \cos \beta)^2} \frac{k_T^2 \sin \left(\frac{1}{2}\beta\right) [A - B] \frac{1}{C}}{4\pi(k^2_T - k^2_L) (k_L \cos \alpha + k_T \cos \beta)}, \]

and in pitch-catch mode

\[ D^L_L(\alpha, \beta) = \exp\left(\frac{i\pi}{4}\right) \frac{k^2_T \sin \left(\frac{1}{2}\beta\right) [D \pm E] \frac{1}{F}}{2\pi(k^2_T - k^2_L) (\cos \alpha + \cos \beta)}, \]

\[ D^T_T(\alpha, \beta) = \frac{k^3_T \sin \left(\frac{1}{2}\beta\right) [D \pm E] \frac{1}{F}}{2\pi(k^2_T - k^2_L) (\cos \alpha + \cos \beta)}, \]

where

\[ A = k^2_L (2k_T)^\frac{1}{2} \cos 2\beta \sin 2\alpha (k_T - k_L \cos \alpha)^\frac{1}{2}, \]

\[ B = (32k_L)^\frac{1}{2} \cos \left(\frac{1}{2}\beta\right) \cos \beta \sin \left(\frac{1}{2}\alpha\right) (2k^2_L \cos^2 \alpha - k^2_T) (k_L - k_T \cos \beta)^\frac{1}{2}, \]
\[ C = (k_o - k_L \cos \alpha) (k_o - k_T \cos \beta) K^+ (-k_L \cos \alpha) K^+ (-k_T \cos \beta), \]
\[ D = k_T \cos 2\beta \cos 2\alpha \sin \left( \frac{1}{2} \alpha \right), \]
\[ E = 2 \cos \left( \frac{1}{2} \beta \right) \cos \beta \sin 2\alpha \left( k_L - k_T \cos \alpha \right)^{\frac{1}{2}} \left( k_L - k_T \cos \beta \right)^{\frac{1}{2}}, \]
\[ F = (k_o - k_T \cos \alpha) (k_o - k_T \cos \beta) K^+ (-k_T \cos \alpha) K^+ (-k_T \cos \beta), \]
\[ K^+ (\gamma) = \exp \left\{ -\frac{1}{\pi} \int_{k_L}^{k_T} \tan^{-1} \left[ \frac{4x^2 (x^2 - k^2_L)^{\frac{1}{2}} (k^2_T - x^2)^{\frac{1}{2}}}{(2x^2 - k^2_T)^2} \right] \frac{dx}{x - \gamma} \right\} \]

and \( k_o \) is the Rayleigh wave number. The choices of sign in \( A \) and \( E \) depend on \( \alpha \) measurement being clockwise or counterclockwise [29]. \( D_L^+ (\alpha, \beta) \) and \( D_T^+ (\alpha, \beta) \) are plotted with respect to various opening angles in Fig. 10 for the incident angle \( \beta = 120^\circ, 135^\circ \) and \( 150^\circ \) with Poisson's ratio \( v = 0.29 \). These three incident angles correspond to the commonly used 30\(^\circ\), 45\(^\circ\) and 60\(^\circ\) angle beam transducer/wedge combinations. Note that in our quasi-pulse-echo setup, both transmitter and receiver will lie on the “same side” of the crack in a region which does not include any shadow or reflection boundaries. Thus both diffraction curves are generally smooth and finite. For a small opening angle \( \theta \), both curves are rather consistent and are of the same order of magnitude.

However, deep nulls in the \( D_T^+ (\alpha, \beta) \) curves can be seen around \( \theta = 45^\circ \) for each incident angle while the second nulls occur around \( \theta = 135^\circ \) for incident angles of 120\(^\circ\) and 135\(^\circ\). The \( D_L^+ (\alpha, \beta) \) curves also reach a minimum around 45\(^\circ\) but increase significantly afterwards as \( \theta \) increases. Physically, these null regions correspond to the receiving positions in the normal and edge-on directions relative to the crack. Poor crack detection probability would be expected in those regions.
For our inspection setup in Fig. 9, the L/T signal ratio can be obtained from eqs. (6) and (7) in the form

\[
\frac{L}{T} = \left| \frac{u_{\text{diff};TL}^{\alpha,\beta}}{u_{\text{diff};TT}^{\beta,\beta}} \right| \approx \left| \frac{\sigma_T}{\omega} \frac{D_T^{\alpha,\beta}}{D_T^{\beta,\beta}} \right|.
\]

To compare the L/T ratios for the two classes of crack-like vs. volumetric flaws, the crack L/T (eq. (9)) ratios for three incident angles are plotted in Fig. 11 against that for the circular cavity (eq. (3)) as predicted by EKA. It is seen that both crack and void curves are parabolic with crack ratios concave up and the void ratio concave down. At the small angle end, good separation between crack curves and the void curve is observed. In the limit as \( \theta \to 0^\circ \), the void ratio vanishes while crack ratios are all above 0.45. Even up to \( \theta = 10^\circ \) the crack ratios are at least two times higher than the void ratio. Thus, the classification concept presented earlier is quantitatively supported by the elastodynamic theories. For the middle region between \( \theta = 20^\circ \) and \( \theta = 75^\circ \), the void curve overlaps the crack curves and no classification is possible. At the high end (\( \theta > 100^\circ \)), crack ratios increase significantly whereas the void ratio decays gradually to zero at critical angle. If pitch-catch inspections are used, this would seemingly be the optimal range. However, for void flaws significant creeping wave components are present in this range, which are not predicted by EKA, so results based on Fig. 11 for these high angles cannot be trusted. Furthermore, pitch-catch inspections at such high angles would require a symmetrical arrangement which may not always be practical. Thus our quasi-pitch-catch configuration appears to be the optimal configuration.
EXPERIMENTAL VERIFICATIONS

In this section, we report the experiments conducted to verify the theoretical predictions obtained in the last section. Both immersion and contact experiments were undertaken. The theoretical model scatterers (cylindrical cavity and semi-infinite crack) were replaced in our experiments with side-drilled holes and through-thickness fatigue cracks respectively. In addition, several pieces of slag were used to represent a more realistic volumetric scatterer as might be found, for example, in weld problems.

Initial Immersion Testing

Since the crack tip diffraction signals were expected to be weak, a considerable amount of experimental effort was made to try to find an optimal testing setup to maximize their amplitudes. The initial testing was carried out in an immersion mode because the incident wave angle could then be easily adjusted. A specimen was made of a 6x4x1 inch aluminum plate, and a 0.0625 inch wide saw cut was made across 4 inch width to about a half inch depth. An ultrasonic inspection system consisting of a Panametrics 5052PR pulser/receiver, a Tektronix 7912AD digitizer and an immersion tank mounted with motorized scan arms was used. A Panametrics V312 10 MHz wideband transducer was used to generate an incident L pulse in the water. With the oblique water angle (tilted from normal) ranging from 10° to 25°, the L wave in the specimen ranged from 48° to 72° and the mode-converted T wave ranged from 23° to 65° in the aluminum specimen. When scanning along the 6 inch dimension of the plate on the opposite side of the cut, signals of L type from both the corner and tip were observed. However, no mode-converted tip signal could be traced.
Because of the cut width, it was also difficult to distinguish diffracted components from reflected tip responses. By simple calculation, we found that over the entire range of angles considered in the water at most 27% of incident energy was able to enter the specimen in L wave motion, and at most 45% in T wave motion. In the returning path, the energy is lost once more by the same percentages, and the mode-converted signal will not be normal to the transducer. Thus, we believe that a simple immersion test configuration of this type is not a practical setup for obtaining L/T ratios.

Contact Apparatus

The second approach to our proposed classification technique (Fig. 4), as discussed earlier, was implemented as shown in Figs. 5 and 6. Westinghouse Corp. provided us with two mild steel blocks of 43.2x10.2x7.3 cm dimensions (L wave speed = 0.59 cm/μs and T wave speed = 0.32 cm/μs), containing fabricated defects. In one block, a through-thickness fatigue crack (generated by repeated cyclic loading cycle) was located at the midspan of the 43.2 cm length. The crack depth was visually measured as 1.9 cm on the side face and had an average inside depth of 2.05 cm as measured later ultrasonically. Four side-drilled holes with the same depth as the crack tip were also placed in the specimen. The diameters of the holes were 2, 3, 4 and 5 mm, respectively. The other steel block contained a strip of slag which was welded (at the same crack depth as the previously mentioned flaws) across the block width. From further ultrasonic 6dB drop tests at normal incidence, the slag strip was deduced to be a uniform U-shape profile opening downward with the same width and depth across the steel block width.
Fig. 5 shows the dual wedge combination. Wedge I is a Panametrics M4011s snap-in miniature angle beam wedge made of polystyrene for low attenuation. Wedge II was made of Lucite and was specially designed to receive mode-converted signals from the crack-hole-slag targets below the top scan surface. To ensure alignment of both wedges and transducers, a V-shaped notch was machined into the back edge of the wedge II to match the leading edge of the wedge I. The distance between the entries of the probes’ center beam at the interface was calculated to be 1.6 cm, which corresponds to an angular opening $\theta$ of $9^\circ$ in a near pulse-echo scan geometry. A pair of Panametrics 535s 0.25 inch diameter miniature angle beam transducers of type A were chosen for higher sensitivity and better penetration in detecting the crack-tip diffraction signals. These transducers have a fairly narrow frequency spectrum centered at 5 MHz. Based on this center frequency, the smallest side-drilled hole diameter was at least 3.1 times as long as the T wave length. Thus the high frequency Kirchhoff postulate is satisfied. One transducer T/R$_1$ mounted on wedge I was driven by a Panametrics 5052PR pulser-receiver to generate a mode-converted $49^\circ$ T incident pulse in the steel block (Fig. 4). The same-mode T echo returning from the flaw converts to L’ in wedge I and is then received by T/R$_1$. The other transducer R$_2$ mounted on wedge II receives the L” signal that is refracted into the wedge from the returning mode-converted ($40^\circ$ L) response. The signals received by the transducers were displayed on a LeCroy 9400 dual trace digital oscilloscope and recorded by an on-site camera. Low viscosity mineral oil was used as couplant. The schematic diagram of the instrumental setup is shown in Fig. 12.
Contact Measurements

Fig. 13 illustrates different wedge positions and ray paths for scanning the fatigue crack. First wedge I was placed at position 1 aiming at the crack root (path a) to obtain a strong corner reflection echo. We used this echo as a reference to identify the same-mode crack-tip T diffracted echo (path b) as we moved wedge I to position 2. Once this diffracted echo was maximized, the mode-converted L diffracted signal can then be located from simple ray tracing along path c. A total of seven data sets was taken at different locations along block width from both sides of the crack. The photo copy of one scan is shown in Fig. 14, where the upper channel records the signals received by transducer T/R₁, and the lower channel records R₂ signals. Major signals have been identified and denoted in Arabic numerals as in Fig. 14. The ray tracing signal-paths (paths denoted in subscripts) are as follows:

signal 1: T₁ → L₁, signal 2: Tₑ → Lₑ, signal 3: T₁ → T₁,
signal 4: T₁ → T₁, signal 5: Tₑ → Tₑ, signal 6: Tₑ → Tₑ,

in which signal 1 is L' and signal 4 is L''. For this record, the measured L'/L'' ratio is therefore 0.34.

It is known that during the measurement process the incident pulse is subjected to various losses or distortions due to effects such as finite transducer bandwidth, transducer beam spreading, transmission differences at interfaces, and material attenuation. For contact testing, couplant pressure also needs to be controlled. However, a complete model for oblique contact testing is very complicated and is not available today. Since we are interested in signal ratio measurements obtained in a quasi-pulse-echo configuration, most losses or distortions can be considered to be "divided out". The remaining effects in our
case are then reduced to the attenuation difference between the two wedges and the path length difference from the flaw to the two wedges. For the wedge attenuation difference, Lucite wedge II was estimated to have 80% more attenuation than the Polystyrene wedge I at the probe center frequency of 5 MHz. This is the reason why the lower channel is "cleaner" than the upper channel in Fig. 14. We also noted that path b is 1.1 cm longer than path c. After taking these differences into account, we found a multiplicative factor of 1.65 should be applied to all data. The corrected amplitude ratio of the seven data then had a range of values from 0.31 to 0.70 whereas the corresponding theoretical prediction (eq. (9)) with $\theta = 9^\circ$ is 0.56.

The scanning procedure for the side-drilled holes was the same as for the fatigue crack except that the reference crack corner reflection was replaced by the direct specular reflected echo. One datum for each hole was taken. The photo record of the 5 mm hole data is given in Fig. 15. Major signals which follow paths as depicted in Fig. 16 are listed below:

- signal 1: $T_a \rightarrow L_b$, signal 2: $T_a \rightarrow L_c \rightarrow L_d$, signal 3: $T_e \rightarrow T_c \rightarrow L_d$,
- signal 4: $T_f \rightarrow R_g \rightarrow T_h \rightarrow T_r$, signal 5: $T_a \rightarrow T_b$, signal 6: $T_f \rightarrow R_g \rightarrow R_j \rightarrow T_k$,

where signal 1 is $L'$, 5 is $L''$ and R denotes the creeping wave. In Fig. 15 the backscattered creeping wave can also be clearly seen. The corrected $L'/L''$ ratios are from 0.08 to 0.14 and the corresponding Kirchhoff value (Fig. 3) is 0.1.

Similarly, we have obtained two slag data. These data were sufficient because of the uniform profile of the slag strip across the block depth. One photo example is shown in Fig. 17. The corrected amplitude ratio was calculated to be 0.12 for this case and a value of 0.18 for the other. These ratio values generally support what were expected from the slag's smooth edge.
Both theoretical and experimental amplitude ratios for the three types of defects are shown in Fig. 18. It is seen that all fatigue crack data have clearly separated from the group of side-drilled holes and slag data; even the lowest crack "outlier" (measurement number 7) is at least twice larger than the highest hole data. Due to the irregularities of the fatigue crack tip, significant scatter in the crack data is understandable. In contrast, the hole and slag data were more evenly distributed. Both theoretical lines in Fig. 18 are in reasonable agreement with the measurements.
SUMMARY AND DISCUSSION

As presented in this paper, the proposed classification concept has been verified in experimental testing on fatigue cracks, side-drilled holes, and pieces of slag in steel specimens. All data show good consistency in the separation of smooth vs. sharp-edged flaws provided that the signal-to-noise ratio is sufficient. Reasonable agreement is observed in comparison with elastodynamics theories.

Since the present technique requires only a direct pulse-echo scanning path, it is advantageous over previous methods in situations where

1. geometry limits multiple transducer (pitch-catch) configurations,
2. multiple reflecting routes are severely blocked or intervened,
3. auxiliary reflecting boundaries are not available or not reliable.

On the other hand, the present technique suffers, as all diffracted wave techniques do, from the inherently weak nature of the diffracted waves in comparison with specular responses. For material with high ultrasonic grain noise, this method may therefore become impractical. Another inherent limitation is the dependence of the amplitude ratio on the incident wave angle. For example, it can be shown that at an edge-on incident angle the mode-converted diffracted signals vanishes in the backscattered direction. Likewise, in several “dead zones” the direct diffracted signals also vanish. For these cases, the present method apparently leads to error indications.

Despite the above mentioned weakness, the present technique can be viewed as a very useful addition to existing techniques at fairly low cost of implementation. As demonstrated in this work, the equipment required
included a standard pulser-receiver, oscilloscope and probe-wedge combination. Only the wedge needs to be specially manufactured. Although we have shown the applicability of this technique, more development work is necessary in order to transfer it from a research concept to a reliable method for general field inspections. For further development of the present technique, we propose the following possible directions:

(1) Continue to search for a better wedge material for implementing a single wedge in contact mode testing or obtain a high energy pulser-receiver-probe combination to make immersion testing practical;

(2) Incorporate recently developed concentric dual-probes [31] on a single wedge to reduce the angle \( \theta \) to zero and maximize the response;

(3) Develop a complete analytical model for angle contact inspection to help optimize the detection device design so that other sensors such as focused transducers may be considered.
ACKNOWLEDGEMENTS

We would like to thank Messrs. A. Jusino and B. Clark from Westinghouse Corporation for providing us with the flawed steel blocks used in this work. Similarly, we would like to thank Dr. R. K. Chapman of CEGB NDT Application Centre, U. K. for many valuable discussions and for providing us the numerical data from his work on the diffraction of a semi-infinite crack. Thanks are also due to Mr. S.-J. Song for his experimental assistance in the final data measurements. This research was supported by the Center for Nondestructive Evaluation at Iowa State University and was performed at Ames Laboratory. Ames Laboratory is operated for the U.S. Department of Energy by Iowa State University under Contract No. W-7405-ENG-82.
In this appendix, we examine some details of the theoretical model used in this paper, i.e., the scattering problem of scattering by a void and the special case of a circular cavity.

We consider the steady-state, time-harmonic wave scattering problem in a two-dimensional, elastic medium which is assumed linearly isotropic and homogeneous. As depicted in Fig. A-1, a smooth, stress-free void of arbitrarily convex shape bounded by closed boundary \(s\) is embedded in an unbounded region \(V\) with elastic constants \(c_{ijkl}\) and density \(\rho\). Clearly, \(c_{ijkl} = \rho = 0\) inside boundary \(s\).

The total field displacement solution to the scattering problem is formally defined by

\[
\mathbf{u}^{\text{to}} = \mathbf{u}^{\text{in}} + \mathbf{u}^{\text{sc}}
\]

as the sum of incident and scattered fields. An incident plane wave of \(\alpha\) type and amplitude \(A\) is denoted by

\[
\mathbf{u}^{\text{in} \alpha} = A \mathbf{D}^{\text{in} \alpha} \exp (i k_{\alpha} \mathbf{P}^{\text{in} \alpha} \cdot \mathbf{x}) \quad (\alpha = L, T).
\]

Where time dependence \(\exp (-i\omega t)\) is understood. The Greek letter \(\alpha\) in the superscripts indicates the wave type: \(L\) as longitudinal wave and \(T\) as transversely shear wave. The unit vectors in the wave propagation and polarization directions are denoted by \(\mathbf{D}\) and \(\mathbf{P}\), respectively. The wave number is \(k_{\alpha} = \omega / c_{\alpha}\) where the wave speed \(c_{\alpha}\) is given in terms of Lame constants \(\lambda\) and \(\mu\) by
(A-3) \( c_L = \left( \frac{\lambda + 2\mu}{\rho} \right)^{\frac{1}{2}}; \)

Also note that in the following, repeated Roman indices mean summation and commas indicate partial differentiation with respect to the succeeding indices.

We now express the scattered displacement field \( u^{sc} \) through the exterior surface integral formulation of Helmholtz type by employing the elastodynamic Betti-Rayleigh reciprocal theorem and the Sommerfeld radiation conditions to obtain

\[
(A-4) \quad u^{sc}_m(y) = \int_s c_{ijkl} n_j(x) \left[ u_l(x) \frac{\partial G_{km}(x,y)}{\partial x_i} - \frac{\partial u_k(x)}{\partial x_i} G_{lm}(x,y) \right] \, ds(x).
\]

Where \( u^{sc}_m(y) \) is the m-th component of the scattered displacement observed at field point \( y \) in \( V \), \( x \) is the boundary point on \( s \) and \( n_j(x) \) is the j-th component of the outward unit normal at \( x \). \( G(x,y) \) is known as the two-dimensional free-space fundamental solution

\[
(A-5) \quad G_{ij}(x,y) = \frac{i}{4\rho\omega^2} \left\{ \delta_{ij} k^2 H_0^{(1)}(kR) + \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} [H_0^{(1)}(kR) - H_0^{(1)}(kL)] \right\}
\]

that satisfies the equation

\[
\frac{\partial G_{km}(x,y)}{\partial x_i \partial x_j} c_{ijkl} + \rho \omega^2 G_{lm}(x,y) + \delta_{im} \delta(x-y) = 0
\]

where \( H_0^{(1)}(kR) \) is the zero-th order Hankel function of first kind and

\( R = |x-y| \).

The following approximations are valid in the far field as \( y \) approaches infinity:
(A-6) \( R \sim y - x \); \( \frac{1}{R} \sim \frac{1}{y} \),

(A-7) \( H_0^{(1)}(kR) \sim \left( \frac{2}{\pi ky} \right)^{\frac{1}{2}} \exp \left[ i \left( ky - \frac{\pi}{4} \right) \right] \exp(-ikx) \)

and

(A-8) \( \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j}[H_0^{(1)}(kR)] \sim \left( \frac{2}{\pi ky} \right)^{\frac{1}{2}} (ik)^2 \exp \left[ i \left( ky - \frac{\pi}{4} \right) \right] \exp(-ikx) \gamma_i \gamma_j .\)

Placing (A-6,7,8) into (A-5) and subsequently into (A-4), we find the far-field displacement is of the form

(A-9) \[ u_m^{(y)} = \left( \frac{2}{\pi kL y} \right)^{\frac{1}{2}} \exp \left[ i \left( k_L y - \frac{\pi}{4} \right) \right] \gamma_i \gamma_m f_i(k_L) \]

+ \( \left( \frac{2}{\pi k_T y} \right)^{\frac{1}{2}} \exp \left[ i \left( k_T y - \frac{\pi}{4} \right) \right] (\delta_{mi} - \gamma_i \gamma_m) f_i(k_T) \)

where \( f_i(k_p) = - \frac{ik_p^2}{4\rho\omega^2} c_{ijkl} \left( \int_s ik_p\gamma_j \gamma_l n_k \exp(-ik_p x) \right) ds \)

+ \( \int_s u_{k,i} \gamma_j \exp(-ik_p x) \) ds).

For a void with stress-free boundary, we have the traction \( t_i = c_{ijkl} u_{k,l} n_j = 0 \) on \( s \). Thus \( f_i(k_p) \) in (A-9) reduces to

(A-10) \[ f_i(k_p) = \frac{k_p^2}{4\rho c_p^2} \int_s c_{ijkl} \gamma_j \gamma_l n_k \exp(-ik_p x) \) ds.

In the high frequency limit, boundary \( s \) can be sharply divided into a lit side \( s_i \) and a dark side \( s_d \). The Kirchhoff assumptions we then make are:
(1) \( s_i \) is locally considered as a perfectly planar reflector so that

\[
A-11 \quad u_i = u_i^{in} + u_i^r \quad \text{on} \quad s_i
\]

where \( u_i^r \) is the \( i \)-th component of displacement as predicted by the reflection of the incident plane wave from an infinitely planar free surface.

(2) The dark side \( s_d \) is in deep shadow so that \( u = 0 \) on \( s_d \). This reduces the integration range in (A-10) to side \( s_i \) only.

Noting that all waves are in-phase on \( s_i \) and recalling (A-2), we obtain

\[
\begin{align*}
  u^{\alpha \alpha} &= A \left[ R_L^\alpha q^L + R_T^\alpha q^T \right] \exp \left( i k_a \varphi^{in} \right) \nu \nu \\
  u &= u^{in} + u^{\alpha \alpha} \\
  &= A \left[ D^{in} + R_L^\alpha q^L + R_T^\alpha q^T \right] \exp \left( i k_a \varphi^{in} \right) \nu \nu
\end{align*}
\]

where \( R_\beta^\alpha (\alpha, \beta = L, T) \) are the displacement reflection coefficients known explicitly as:

\[
\begin{align*}
  &A-12 \quad R_L^\alpha (\theta) = \frac{\sin^2 \theta \sin 2\theta - \kappa^2 \cos^2 2\theta}{\sin^2 \theta \sin 2\theta + \kappa^2 \cos^2 2\theta} \\
  &A-13 \quad R_T^\alpha (\theta) = -\frac{2 \kappa \sin \theta \cos 2\theta}{\sin^2 \theta \sin 2\theta + \kappa^2 \cos^2 2\theta}
\end{align*}
\]

where \( \kappa = \frac{c_L}{c_T} \) and \( \cos \phi = \frac{\cos \theta}{\kappa} \) (see Fig. A-2); 

\[
\begin{align*}
  &A-14 \quad R_L^\alpha (\theta) = \frac{\kappa \sin 4\theta}{\sin^2 \theta + \kappa^2 \cos^2 2\theta} \\
  &A-15 \quad R_T^\alpha (\theta) = \frac{\sin^2 \theta \sin 2\theta - \kappa^2 \cos^2 2\theta}{\sin^2 \theta \sin 2\theta + \kappa^2 \cos^2 2\theta}
\end{align*}
\]

where \( \cos \phi = \kappa \cos \theta \) (see Fig. A-3).

If we further define
\( (A-16) \quad \mathbf{a}^\alpha = D_\text{in}^{\alpha \alpha} + R_L^{\alpha \alpha} R_T^{\alpha L} + R_T^{\alpha T} \)

then

\( (A-17) \quad \mathbf{u} = A \mathbf{a}^\alpha \exp (i k_o \mathbf{a}^\alpha) \) on \( S_L \).

For isotropic materials, the elastic constant is given explicitly by

\( (A-18) \quad c_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu \delta_{ik} \delta_{jl} + \mu \delta_{ij} \delta_{jk}. \)

With the results of (A-17,18) and the Kirchhoff assumption (2), we arrive at

\( (A-19) \quad f_1 (k_p) = f_i^\alpha (k_p) \)

\[ \frac{k_p^2 A}{4 \rho c_p^2} \int_S \left[ \lambda (a^\alpha a^\beta \gamma)_{ij} + \mu (a^\alpha a^\beta n_i + \mu (a^\alpha a^\beta a_i\gamma) \exp [-ik_p q_{\alpha \beta} \psi] \right] ds \]

where \( k_p = k_p \gamma \) and

\( (A-20) \quad q_{\alpha \beta} = \frac{k_{\alpha \beta}}{k_p} p^{\alpha \beta} - \gamma. \)

We now re-define (A-9) as

\( (A-21) \quad u_{m}^{sc} (\gamma) = u_{m}^{sc, \alpha \beta} \)

\[ = u_{m}^{sc, \alpha L} + u_{m}^{sc, \alpha T} \]

\[ = \left( \frac{2}{\pi k_L Y} \right)^{\frac{1}{2}} \exp \left[ i \left( k_L Y - \frac{\pi}{4} \right) \right] A_{m}^\alpha \]

\[ + \left( \frac{2}{\pi k_T Y} \right)^{\frac{1}{2}} \exp \left[ i \left( k_T Y - \frac{\pi}{4} \right) \right] B_{m}^\alpha \]

where

\[ A_{m}^\alpha = Y_1 Y_m f_i^\alpha (k_L) \]

\[ B_{m}^\alpha = (\delta_{mi} - Y_1 Y_m) f_i^\alpha (k_T) \]
are the scattering amplitudes of type L and T respectively.

For a circular cavity of radius \( r \), the line integral of (A-19) can be converted to a one-dimensional integral by changing the integration variable \( ds \) to \( rd\phi \). Then this one-dimensional integral may be evaluated by the method of stationary phase formula to give

\[
\int_a^b f(\phi) \exp[ik g(\phi)] \, d\phi \\
\sim f(\phi_s) \exp[ik g(\phi_s)] \left( \frac{2\pi}{k|g''(\phi_s)|} \right)^{1/2} \exp\left\{ i \frac{\text{sgn}[g''(\phi_s)] \pi}{4} \right\}
\]

provided that the second derivative \( g''(\phi_s) \) does not vanish (hereafter the subscript \( s \) denotes the stationary point). From the correspondence \( g(\phi) = q_{\alpha\beta} x(\phi) \) we obtain \( q_{\alpha\beta} = -q_{\alpha\beta} n_s \), \( g''(\phi_s) = q_{\alpha\beta} r(\phi_s) \) and \( g(\phi_s) = q_{\alpha\beta} r \cos \psi \). Where \( r = |x_s| \) and \( \psi \) is the angle between \( q_{\alpha\beta} \) and \( x_s \). \( \psi \) can be eliminated by re-orienting the coordinates so that \( q_{\alpha\beta} \) and \( x_s \) are aligned. In the following \( \cos \psi = -1 \) is taken.

By applying the above stationary phase terms to (A-19) and then (A-21), we have the far-field scattered displacements as

\[
(A-22) \quad u_{m}^{\text{sc}_{\alpha\beta}} \sim A \left( \frac{r}{4q_{\alpha\beta} y} \right)^{1/2} \exp[ik (y - rq_{\alpha\beta})] I_{\alpha\beta,m} \quad (\alpha, \beta = \text{L or T})
\]

where

\[
I_{\alpha L,m} = \left[ \frac{\lambda}{\lambda + 2\mu} (a^\alpha n_{s}) + \frac{2\mu}{\lambda + 2\mu} (Y^\alpha a^\alpha)(Y n_{s}) \right] Y_m
\]

and

\[
I_{\alpha T,m} = (Y^\alpha a^\alpha)(n_{s}) Y_m + (Y n_{s}) a^\alpha_m - 2(Y^\alpha a^\alpha)(Y n_{s}) Y_m
\]
The explicit $I_{\alpha\beta}$ expressions can be obtained by properly evaluating the vector products in $I_{\alpha\beta}$ at the stationary phase point. Due to limited space, we only summarize these expressions below:

\[(A-23)\]

1. $\alpha = L, \beta = L$: $q_{\alpha\beta} = 2\cos\frac{\theta}{2}; I_{\alpha\beta} = 2 R_L^L \left(\frac{\theta}{2}\right) \cos\frac{\theta}{2}$,

2. $\alpha = L, \beta = T$: $q_{\alpha\beta} = \frac{\sin\theta}{\sin\phi}, \phi = \tan^{-1} \frac{k\sin\theta}{1 + \kappa\cos\theta}; I_{\alpha\beta} = 2 R_L^T (\phi) \cos(\theta - \phi)$,

3. $\alpha = T, \beta = T$: $q_{\alpha\beta} = 2\cos\frac{\theta}{2}; I_{\alpha\beta} = 2 R_T^T \left(\frac{\theta}{2}\right) \cos\frac{\theta}{2}$,

4. $\alpha = T, \beta = L$: $q_{\alpha\beta} = \frac{\sin\theta}{\sin\phi}, \phi = \tan^{-1} \frac{\sin\theta}{\kappa + \cos\theta}; I_{\alpha\beta} = 2 R_L^T (\phi) \cos(\theta - \phi)$,

in which $\kappa = \frac{c_L}{c_T}$ and $R$'s are the displacement reflection coefficients as given in $(A-12,13,14,15)$.

The results of cases (3) and (4) agree with eqs. (27) and (28) in [22] as obtained from an asymptotic ray series solution. For the limiting condition $\theta = 0^\circ$ (backscattered direction) in cases (2) and (4), we have $\phi = 0^\circ$ and $q_{LT} = 1 + \kappa^{-1}, q_{TL} = 1 + \kappa$ instead. Note that $(A-23)$ is not valid when the reflected $L$ wave goes beyond critical angle. Also note that this approach can be easily extended to three-dimensional scatterers by applying the two-dimensional stationary phase formula for the surface integral in the three-dimensional case. However, the three-dimensional extension and the situations above the critical angle are beyond the scope of this paper and will not be included here.
REFERENCES


10. M. F. Whalen and A. N. Mucciardi. "Inversion of Physically


28. Numerical data for diffraction coefficient in semi-infinite crack scattering are provided by Dr. R. K. Chapman, NDT Application centre, CEGB, UK.


Figure 1a. Ultrasonic signals scattered from a planar crack-like flaw

Figure 1b. Ultrasonic signals scattered from a volumetric void-like flaw
Figure 2a. Modified Delta Technique for volumetric flaw:
path A signal > path B signal

Figure 2b. Modified Delta Technique for planar flaw:
path A signal < path B signal
Gain difference:
\[ \Delta_i = V^i_{LLL} \text{ (dB)} - V^i_{LLT} \text{ (dB)} \]
i = F (flaw) or C (calibration reflector)

Classification Factor:
\[ CF = \Delta_C - \Delta_F \]

Figure 3. Flaw identification by LLT technique (adapted from [15])

Figure 4. Proposed flaw classification concept
(other signals at interface are not shown)
Figure 5. Dual wedge compromise

Figure 6. Dual-wedge inspection geometry
Figure 7. Scattering geometry of void-like flaw (Kirchhoff theory)
Figure 8. Comparison between Kirchhoff solution and T -> L displacement reflection coefficient for circular cavity.
Figure 9. Scattering geometry of semi-infinite crack (Wiener-Hopf diffraction theory)
Figure 10a. $L$ and $T$ potential diffraction coefficients for semi-infinite crack ($T$ incident angle = 120 deg.)
Figure 10b. L and T potential diffraction coefficients for semi-infinite crack
(T incident angle = 135 deg.)
Figure 10c. L and T potential diffraction coefficients for semi-infinite crack (T incident angle = 150 deg.)
Figure 11. Comparisons of L/T signal ratio between circular cavity and semi-infinite crack (half plane)
Panametrics 535S 5MHz/.25\" miniature angle beam transducers

dual wedge combination

on-site camera

Panametrics 5052PR pulser / receiver

LeCroy 9400 dual digital oscilloscope

Figure 12. The experimental apparatus

Figure 13. Probe positions and signal paths in scanning fatigue crack
Figure 14. Example of photo record of fatigue crack scan
(vertical scale: 10mv/div., horizontal scale: 5μs/div.)
Figure 15. Photo record of 5 mm side-drilled hole scan
(vertical scale: 0.2v/div. in upper channel, 0.1v/div. in lower channel; horizontal scale: 2µs/div.)
Figure 16. Probe positions and signal paths in scanning side-drilled holes
Figure 17. Example of photo record of slag strip scan
(vertical scale: 0.2v/div. in upper channel, 50mv/div. in lower channel; horizontal scale: 2µs/div.)
Figure 18. Data distribution of crack, hole and slag measurements
Figure A-1. Flaw scattering geometry
Figure A-2. Reflection geometry for L incident wave on stress-free boundary

Figure A-3. Reflection geometry for T incident wave on stress-free boundary
PART II. NEW APPROACHES TO MODEL-BASED ULTRASONIC FLAW SIZING
NEW APPROACHES TO MODEL-BASED ULTRASONIC FLAW SIZING

Chien-Ping Chiou and Lester W. Schmerr

Center for NDE and Department of Engineering Science & Mechanics Iowa State University Ames, Iowa 50011
ABSTRACT

Current ultrasonic equivalent flaw sizing methods have a number of important limitations: 1) they are both iterative and highly nonlinear in nature, 2) they require the availability of flaw classification information, and 3) they are restricted to finding equivalent flaw sizes in terms of very restrictive shapes (ellipsoids and ellipses).

Here, a series of approaches are outlined that provide complete or partial solutions to all of the above problems. Both numerical and experimental results that validate these new approaches are also given. First, we describe a new linear least squares/eigenvalue method that can replace existing nonlinear routines and provide a computationally fast, simple, and robust sizing procedure. Second, we demonstrate a way to do equivalent flaw sizing even when flaw classification information is not present. When such classification information is available, we also demonstrate how it can be used to improve sizing estimates for cracks, by the elimination of systematic measurement errors due to finite transducer bandwidth effects. Third, we discuss an alternative spherical harmonics expansions sizing algorithm that is not strictly limited to simple ellipsoids or elliptical shapes.
INTRODUCTION

Ultrasonic flaw sizing has long been a challenging task in nondestructive evaluation (NDE). For flaws that are many wavelengths in diameter, simple scanning and dB drop criteria can often be used effectively [1]. For flaws whose sizes are comparable to a wavelength, however, such methods fail. Detailed imaging techniques, such as three-dimensional computerized tomography [2] can resolve small flaws, but they typically require massive amounts of data and long processing time.

Another way to deal with small flaws is a model-based approach called equivalent flaw sizing [3-5]. In this method, one reconstructs the flaw in terms of a suitably general “best-fit” equivalent flaw shape that matches the scattering data. The advantages of this technique are 1) it does not require extensive scanning data so that it is fast and efficient; 2) it constrains the sizing problem to finding a few “major” sizing parameters thereby avoiding many of the non-uniqueness and ill-posedness problems of more general formulations [6-7], and 3) it provides information directly usable by modern structural design methods, such as fracture mechanics.

Previous work on the equivalent flaw sizing approach [3-5] has shown that it is a viable sizing method. However, there are three major disadvantages of the technique as used to date: 1) it involves the solution of a difficult nonlinear optimization problem, which can be ill-behaved and unpredictable when significant experimental errors are present in the scattering data; 2) the method relies on having flaw classification information (i.e., if the flaw is volumetric or crack-like) which may be as difficult to obtain as the sizing estimates themselves
and 3) the equivalent flaw shapes currently allowed by the method are very restrictive (ellipsoids for volumetric flaws and ellipses for cracks).

Here, we describe a series of new approaches to equivalent flaw sizing problems that eliminate or minimize the problems mentioned above. Specifically, the nonlinear least squares algorithm used previously is shown to be replacable by a much simpler linear least squares/eigenvalue problem. Also, a new method for performing equivalent flaw sizing without having classification information available is given. If such classification is available, however, we also show how it can be used, together with a model-based signal processing approach, to eliminate some of the systematic sizing errors present for cracks due to finite transducer bandwidth effects. Finally, we present a spherical harmonics expansion algorithm that uses exactly the same data as the equivalent ellipsoid/ellipse sizing methods but is not restricted to just those constrained shapes.
THE INVERSE SCATTERING MODEL

Born and Kirchhoff Approximations

The Born and Kirchhoff approximations serve as the foundation of the equivalent flaw sizing methods. Many previous works have studied in detail these approximations (see, for example, [3-5, 8-11]) so here we only briefly outline their key features which can be exploited for sizing purposes.

First, consider a general volumetric flaw subject to a plane incident wavefront as shown in Fig. 1a. As the wavefront continuously sweeps across the flaw, a plane area A is profiled at each intersection with the flaw. If the incident waveform is a unit impulse function of time, the signal received far away from the flaw is called the far-field scattering impulse response $u^\infty(t)$. For an general volumetric flaw the Born approximation predicts that $u^\infty(t)$ is given by the relationship [9]

$$u^\infty(t) \propto \frac{d^2 A(t)}{dt^2}.$$  

Fig. 1b illustrates the Born approximation result for the far-field scattering impulse response of a weakly scattering ellipsoidal inclusion. In contrast, Fig. 1c depicts an actual response from a 200x400 μm spheroidal void embedded in titanium. Both the Born and actual responses (Figs. 1b and 1c) show that the first major feature of the scattering is the large specular reflection produced at the leading point $E_1$ of the flaw. For the weakly scattering inclusion, this specular reflection is followed by a region of constant response and finally, by a large back surface response, $E_1$ (Fig. 1b). In contrast, the void response is non-uniform after the first specular reflection, followed by a (usually small) creeping
wave contribution. However, between the center of the flaw (point C, located at the zero-of-time, \( t = 0 \) in Figs. 1b and 1c) and \( E_{\nu} \), the Born approximation (when suitably bandlimited) and actual results do not differ appreciably. Thus, if a method can be found for estimating the location of the flaw center, C, the Born approximation should be able to be used to estimate the time interval, \( \Delta t \), between \( E_{\nu} \) and C, or equivalently the effective radius, \( r_{e} \) (Fig. 1a) of the flaw in the incident wave direction, since \( \Delta t \) is just proportional to \( r_{e} \) [4]. In practice, this \( \Delta t (r_{e}) \) estimate can be made using either time data directly [9] or equivalent frequency domain values of the scattering response [8]. In either case, however, first locating the center of the flaw is essential. Fortunately, a solution to this problem is also available entirely within the Born approximation, since, as eq. (1) indicates, C should be located where the area function, \( A(t) \), is a maximum for volumetric flaws with center of inversion symmetry. Thus, by integrating the far-field scattering impulse response twice and locating the maximum of the resulting waveform, an estimate of the location of C can be made. This method, termed as the "area function" approach, provides one solution to the "zero-of-time" problem for volumetric flaws with center of inversion symmetry. Discussion of other schemes can be found in, for example [8].

Now, consider instead an isolated smooth planar crack. In this case, the Kirchhoff approximation predicts in general [12] that for an incident L or T-wave

\[
\mathbf{u}^{sc}(t) \propto \frac{dD}{dt} + \frac{dH[D]}{dt}
\]

where \( D \) is the length of the line swept out by the intersection of the incident wavefront with the crack (Fig. 2a) and \( H[] \) is the Hilbert transform operator. For
an L-wave incident at any angle or an T-wave not beyond the critical angle with respect to the crack surface, the Hilbert transform term disappears. A comparison between the Kirchhoff model of the scattering of a L-wave by an elliptical crack (Fig. 2b) and an actual signal (Fig. 2c), shows that Kirchhoff approximation locates the two "flashpoints" (A and B) responses correctly, but misses the later arriving waves such as the Rayleigh wave (Fig. 2c). Since the time separation, Δt, between flashpoints now is just proportional to 2r_e (Fig. 2a), a measurement of Δt yields a measurement of the effective radius of the crack in the incident wave direction. Note, however, in the crack case, it is not necessary to locate the center of the flaw, C, to find r_e.

The above discussions show that the Born and Kirchhoff approximations can be used to provide estimates of an effective radius parameter, r_e, in a given incident wave direction. In the next section, we will describe how such a parameter, when measured from a number of different directions, can be used to estimate the major dimensions of a flaw and its directions.
EQUIVALENT FLAW SIZING

Equivalent flaw sizing using ultrasonic waves is an approach whereby shape and orientation information of a defect are obtained in terms of a “best fit” simple geometry that is able to represent the major aspects of the flaw. Examples of this approach for volumetric flaws have been developed by Hsu et al. [5] based on the Born approximation and by Sedov and Schmerr [3] for cracks using the Kirchhoff approximation. A previous paper [4] describes how these separate algorithms can be unified into a single flaw sizing algorithm that can determine the size and orientation of an isolated defect in terms of a best-fit ellipsoid (for volumetric flaws) or ellipse (for cracks). We will briefly review this unified algorithm below. Then, we will show how an equivalent but much simpler procedure can be used instead.

Unified Equivalent Sizing Algorithm

Symbolically, the equivalent radius, \( r_e \), can be related to the flaw parameters and the transducer orientation by a function \( r_e = f(x, y_i) \) where the components of \( x \) are the flaw orientation and size parameters, and \( y_i \) denotes the \( i \)-th transducer orientation. For an ellipsoid in spherical coordinates, \( x \) consists of three semi-radii and three angles for the semi-axes, while \( y_i \) contains two scanning angles (Fig. 3). The elliptical flat crack flaw shape is then a special case with one semi-radius equal to zero.

In coordinate-invariant form, this functional relationship can be expressed as:

\[
(3) \quad r_e = \left[ a^2 (e_q \cdot e_a)^2 + b^2 (e_q \cdot e_b)^2 + c^2 (e_q \cdot e_c)^2 \right]^{1/2}
\]
where \( a, b, c \) are the size parameters and \( \mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c \) are orientation unit vectors. \( \mathbf{e}_q \) is determined by the known transducer setup; for example, in a L to L case (both incident and scattering waves are of L type) \( \mathbf{e}_q = \mathbf{p} - \mathbf{e}_o \) where \( \mathbf{p} \) and \( \mathbf{e}_o \) are the unit vectors in the incident and scattering directions, respectively. If we write all of our unit vectors \( \mathbf{e}_i \) (i=a, b, c, q) in terms of spherical coordinates, eq. (3) reduces to a very complicated functional form [4] in terms of six unknowns from which a complete description of the size and orientation of the equivalent flaw can be given.

To turn eq. (3) into a practical sizing method, we assume \( N \) measurements of \( r_e \) are made of a given flaw from different directions. Previously, the unified flaw sizing algorithm assumed that the flaw type (crack or volumetric) was known so that \( r_e \) could be obtained from either the correct center of flaw or flashpoint locations, respectively [4]. However, as we will show shortly, \( r_e \) can be obtained even if such classification information is absent.

Once \( N \) values have been experimentally determined, we can then form up the function

\[
I (a, b, c, ...) = \sum_{i=1}^{N} [(r_{e, \text{exp}})_i - (r_e)_i (a, b, c, ...)]^2
\]

and search for the best fit parameters \((a, b, c, ...)\) that match the measurements, i.e., that minimize \( I \). Thus, we can convert the equivalent flaw sizing scheme to a standard nonlinear optimization (minimization) problem. Nonlinear optimization problems such as this one are notoriously difficult to deal with and no completely general scheme is available for their solution. By comparing a number of available direct search and gradient methods, we found that a variation of the Levenberg-Marquadt method [13] gave the best overall results.
However, some modifications were found to be necessary to produce a robust algorithm. The strategies used in [4] included a multistart technique and a bisection of the feasible region. With these enhancements, we found the optimization procedure was both fast and robust when tested on a wide variety of synthetic and experimental examples.

For further refinement of this algorithm, one approach is to put "reasonable" error bounds around the central values of the flaw parameters [14]. These central values for the size parameters in some situations may be estimated apriori from the time-of-flight $\Delta t$ measurements. However, applying these error bounds would introduce additional numerical difficulties associated with solving constrained least squares problems. Fortunately, such an extension of the algorithm of [4] is not necessary. In fact, the unified sizing algorithm can be reformulated as a two-step procedure which avoids solving the difficult nonlinear optimization problem entirely.

**Equivalent Sizing as a Linear Least Squares/Eigenvalue Problem**

As the first step in this reformulated sizing algorithm we express the equivalent flaw radius for an ellipsoid as

\[ r_e^2 = C_{xx}L_x^2 + 2C_{xy}L_xL_y + 2C_{xz}L_xL_z + C_{yy}L_y^2 + C_{yz}L_yL_z + C_{zz}L_z^2 \]

where

\[ L_x = \cos\theta \cos\phi \]
\[ L_y = \cos\theta \sin\phi \]
\[ L_z = \sin\theta \]

are the Cartesian components of \( \mathbf{e}_q \) in terms of spherical coordinates angles \((\theta, \phi)\). The C-coefficients are explicitly given by
in which $a$, $b$, $c$ are the three semi-axis sizes and $m_x$, $m_y$, $m_z$ ($m = a, b, c$) are three sets of unit vector components along the three semi-axes. As these expressions for the $C$s show, obtaining the flaw parameters ($a, b, c, ...$) directly is a very nonlinear process. However, if we instead form up the function

$$J(C_{xx}, C_{yy}, ...) = \sum_{i=1}^{N} \left[ (x_{\text{exp}}^i)^2 - C_{xx} L_{x,i}^2 - 2 C_{xy} L_{x,i} L_{y,i} - ... \right]^2$$

and find the best-fit coefficients ($C_{xx}, C_{xy}, ...$) that fit the measured data, we only have to solve a much simpler (and well-behaved) linear least squares problem.

Once the $C$s are obtained in this manner, the next step is to minimize the functional form

$$F = r_e^2 + \lambda (1 - L_x^2 - L_y^2 - L_z^2)$$

where $\lambda$ is a Lagrange multiplier, or equivalently, solve the eigenvalue problem

$$C \bar{x} = \lambda \bar{x}, \quad C = \text{real symmetric matrix of } C\text{-coefficients}$$

with the Lagrange multipliers being the eigenvalues. For the real symmetric $C$ matrix, we see that the three eigenvalue solutions are real and equal to the square of the three semi-axis sizes ($a^2, b^2, c^2$). Likewise, the corresponding
normalized eigenvectors are identical to the three orientation unit vectors \( \mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c \), respectively. This can be easily seen by drawing an analog to the principal stress analysis in elasto-statics [15].

This algorithm has been tested on both synthetic and experimental data and compared with the nonlinear least squares algorithm as shown in Table 1. When tested on synthetic data, like the nonlinear algorithm, all results reach machine accuracy even for very degenerate flaw shapes so that no error estimates are given in Table 1 for these cases. When tested on experimental data, this algorithm again has similar performance to the previous nonlinear algorithm [4], but is computationally more efficient.
CORRECTION OF BANDWIDTH EFFECT IN CRACK DATA

All experiments are subject to various errors in the data collection process. In the ultrasonic NDE measurement process, flaw signals are often contaminated by material and electronic noise sources, transducer diffraction effects, transducer bandwidth limitations, material attenuation, etc. Since our sizing algorithms rely on ideal error-free theoretical models of the flaw signals, it is important to eliminate as many of these sources of variability as possible.

One way to remove attenuation and transducer diffraction effects is through the use of deconvolution procedures such as the measurement model [16]. However, these methods do not compensate for material and electronic noise or finite transducer bandwidth. Here, we will show it is possible, in some cases, to attack the finite bandwidth problem by other means.

All transducers used for NDE defect characterization introduce bandwidth limitations. If an input signal has wider frequency range than the transducer (as is often the case), its spectrum is always cutoff to some extent, and consequently useful information is lost. Some improvement of bandwidth can be accomplished by spectrum extrapolation methods [17-18]. However, these techniques generally reduce to solving another type of inverse problem and currently there are no generally acceptable approaches. Thus, the study of errors caused by the absence of frequency information remains a very open area of research.

In sizing cracks by our equivalent flaw sizing algorithms, the error in measuring $\Delta t (r_e)$ due to limited transducer bandwidth can be reduced significantly by having classification information. This is possible because the
knowledge that the flaw is a crack allows one to model, and hence remove, a consistent error present in the measurement process. The manner in which this is done is as follows. Consider the ideal (infinite bandwidth) crack response in the Kirchhoff approximation (Fig. 2b). The frequency components of this response can be obtained exactly and then multiplied by a bandpass function that represents the effective bandwidth of the ultrasonic transducer(s). If this result is then Fourier transformed back into a time waveform and the bandlimited time interval \( \Delta t \) measured between the two major antisymmetric pulses in that waveform, an error term (defined as \( \frac{\text{exact } \Delta t - \text{bandlimited } \Delta t}{\text{exact } \Delta t} \)) can be plotted versus the bandlimited \( \Delta t \). This plot can then be used as a "look-up curve", for any measured \( \Delta t \) value, to correct for the systematic effects of the transducer bandwidth. Fig. 4 shows such results for the case where the Kirchhoff response has been bandlimited between 2 MHz to 20 MHz.

The general shape of the error curve in Fig. 4 can be understood if we consider the transducer frequency response to be a simple lowpass rectangular window so that the time domain transducer impulse response is just proportional to a sinc \( \frac{\sin(t)}{t} \) function. When the \( \Delta t \) between flashpoints in Fig. 2c is much larger than the main lobe of the sinc function, then the crack response will be the convolution of the sinc function with the Kirchhoff response of Fig. 2b, leading to a "smoothed" version of the ideal (infinite bandwidth) Kirchhoff signal, where the antisymmetric peaks at the flashpoints will be closer together than in the infinite bandwidth case. This leads to positive errors as shown in Fig. 4. If, however, the \( \Delta t \) between flashpoints is much smaller than the main lobe of the sinc function, the dipole-like Kirchhoff signal acts like a "differentiator" when it is convolved with the sinc function,
leading to an antisymmetric response, but where the $\Delta t$ between peaks of opposite sign in this response really are only a measure of the time interval between peaks in the derivative of the sinc function, and do not represent a time between flashpoints. Since the sinc function or its derivative are only functions of the assumed bandwidth of the transducer, we expect this erroneous bandlimited $\Delta t$ measurement to remain eventually constant as the real $\Delta t$ between flashpoints decreases, leading to increasingly larger negative errors as shown in Fig. 4.

The importance of this result is that in an experimental setup, where multiple look-angles are taken of a given flaw with a fixed frequency bandwidth, the crack will appear to be larger or smaller depending on the separation of flashpoints. Fig. 4 shows that we can not only make $\Delta t$ corrections to the data when $\Delta t$ is relatively large, but also we can estimate apriori, from the given bandpass characteristics of the transducer used, the minimum acceptable $\Delta t$ measurement, below which the results are suspect. The look-up curve was then tested on a set of ultrasonic pulse-echo data from a 400$\mu$m artificial "crack" (actually a circular pill box of 1:10 ratio) in titanium. These data were measured starting at normal incidence and then proceeding with 5° inclination increments from the crack normal and were obtained after transducer diffraction effects and attenuation errors were eliminated through the use of the measurement model mentioned previously [16]. By assuming the remaining data errors were primary due to the bandwidth effect, we corrected the $\Delta t$ measurement for each look-angle by a certain percentage based on the curve of Fig. 4. Any $\Delta t$ value below a preset criteria of 50 ns was rejected. A comparison of corrected $\Delta t$ data with the Kirchhoff prediction is shown in Fig. 5, where we
note that the look-up curve does help to remove the systematic error from the original data; the data error is now quite small and more evenly distributed around 0%. Table 2 shows that better sizing estimates are also obtained for the crack when using the corrected data, although the differences are not particularly large.
SIZING AND CLASSIFICATION

The problem of flaw characterization can be viewed as a multi-step process where decisions are made as to flaw type (a classification process) and flaw geometry (a sizing process). Previously, we have described the use of an expert system, FLEX [19], for determining if an unknown flaw is a volumetric flaw or a crack and the subsequent use of equivalent flaw sizing algorithms whereby the flaw is sized in terms of a best fit ellipsoid (for volumetric flaws) or ellipse (for cracks). Here, we will show how equivalent flaw sizing can be done without having prior classification information and demonstrate how the quality of the sizing estimates depends on such information.

Fig. 6 describes an ultrasonic flaw characterization methodology in the form of a process tree. In that tree, the flaw is first classified as to type. If the flaw is of a volumetric type (void or inclusion), the 1-D Inverse Born [8] is used to estimate the equivalent radius, \( r_e \), from the center of the flaw to the front surface in the viewing direction. For a crack, the same distance is estimated by measuring the time separation between "flashpoints", i.e., when the incident wavefront strikes the front and back flaw edge. For either case, a single unified inversion algorithm then sizes the flaw in terms of an equivalent ellipsoid (volumetric case) or ellipse (crack case) that best matches the distances measured, at multiple viewing angles, in a least squares sense. The classification step is essential in this method so that the data can be pre-processed appropriately (1-D Inverse Born or time between flashpoints) to extract the radius \( r_e \).

Even if classification information is not available, it may be possible to extract \( r_e \) from the ultrasonic scattering data. Consider, for example, our
previous ideal impulse responses of a volumetric flaw and a crack in the Born and Kirchhoff approximations respectively in Fig. 7. In the volumetric case, the first part of the time domain response exhibits a large specular reflection followed by a step-like response, whereas in the crack case, we see two distinct antisymmetrical "flashpoints".

Although these responses are quite different, if we integrate each response, the time, $\Delta t$, between the first extremum of the signal and the first following zero can be used in both cases to estimate the distance $r_e$ required in the sizing algorithm. In this case, a general ellipsoid is used to fit the data for both volumetric flaws and cracks.

Table 3 gives the results of sizing both a 400 $\times$ 400 $\times$ 200 nm spheroidal void and a 400 $\mu$m radius circular crack in titanium based on experimental data taken with a multiviewing transducer system [5]. For the volumetric flaw (cases 1-3), the use of classification information (to allow the 1-D Inverse Born approximation to be used) results in a reduction of the error in the largest size estimates from 88% to 16%. For the crack (cases 4-6) errors in the maximum size were reduced from 6% to 2.8%. Thus, classification information does also allow improved sizing, at least for the algorithms employed in this comparison.

For the crack example shown in Table 3, the crack parameters ($\theta_c$, $\phi_c$, a, b, c) would be estimated, using classification information and no frequency bandlimit corrections, as (93°, $\ldots$, 374, 348, 0) resulting in a maximum sizing error of 6.5%, which is very similar to the results without classification. However, knowing the flaw is a crack and applying the corrections shown in Fig. 4 results in the size parameters shown in Table 3 (case 4) and reduces the maximum size errors to the 2.8% value mentioned previously.
As the results of Table 3 demonstrated, the unified sizing algorithm can be used, with and without classification information, to provide reasonably accurate sizing and orientation information based on experimental measurements of the $\Delta t (r_e)$ parameter. Some of the disadvantages of this method are: a) it is iterative, leading at times to long computations b) it uses a nonlinear least squares approach where convergence to the correct answer is difficult to guarantee, particularly in the case of noisy data and finally c) it is restricted to obtaining shape information only in terms of best-fit ellipsoids. The new linear least squares/eigenvalue method overcomes the first two disadvantages, but does not eliminate the constrained ellipsoidal shape assumption.

Recently, we have developed an alternate sizing algorithm that uses exactly the same scattering $\Delta t (r_e)$ data but instead fits that data to a set of spherical harmonics with unknown coefficients. This new approach essentially increases the degrees of freedom in parameterizing the flaw shape, and leads to a noniterative linear least squares problem. In principle, it is not restricted to simple shapes such as ellipsoids. The method works as follows.

As mentioned previously, the measured times $\Delta t$ can be used to estimate the equivalent flaw radius, $r_e$, in a given direction. This equivalent radius geometrically is merely the tangent-plane distance between the center of the flaw and the front surface in a given viewing direction (Fig. 1a). If we let $(\theta_i, \phi_i)$ be the angles which denote the i-th viewing direction in a fixed spherical
coordinate system, then \( r_e(\theta, \phi) \) can be expanded in spherical harmonics \( S_{mn}(\theta, \phi) \) as

\[
(6) \quad r_e(\theta, \phi) = \sum \sum C_{mn} S_{mn}(\theta, \phi)
\]

or more precisely

\[
K_n \quad r_e(\theta, \phi) = \sum \sum P^m_n(\sin\theta) [A_{mn} \sin m\phi + B_{mn} \cos m\phi]
\]

where \( P^m_n(\sin\theta) \) are known normalized associated Legendre functions and \( C_{mn} \) (in terms of \( A_{mn} \) and \( B_{mn} \)) are unknown coefficients. Using the expansion and the \( N \) measured values of \( r_e, r_\text{exp} \), an error measure \( E \) can be formed as

\[
(7) \quad E = \sum_{i=1}^{N} [(r_e)_i - (r_\text{exp})_i]^2
\]

Minimizing \( E \) then leads to a linear least squares problem for the determination of the \( C_{mn} \). Once the \( C_{mn} \) are found, the value of \( r_e(\theta, \phi) \) for any \( (\theta, \phi) \) is known through eq. (6). From this function the actual surface Cartesian coordinates \((X, Y, Z)\) of the flaw can be found [20] (note the change in definition of our spherical coordinates):
The use of spherical harmonics for representing an arbitrary smooth function in spherical coordinates is well justified (see, for example, [21]). The combination of spherical harmonics and Minkowski's support function [22] has also been discussed in [23]. Similar expressions to eq. (10) have appeared in [20] for electromagnetic inverse scattering problem, and in [24] with different coordinates.

Following the same procedures as before, we have tested this algorithm for a number of synthetic and real flaws. In Figs. 8a to 8e simulation results involving a synthetic $100 \times 200 \times 300$ $\mu$m ellipsoid are presented. Fig. 8a shows four perspective views of the original geometry. In Fig. 8b, the ellipsoid was reconstructed by using 54 error-free data in a full $4\pi$ aperture with 49 terms taken in the series expansion. Although some distortions appear due to the truncation effect of the series, very good agreements on individual projected profiles are observed. Since this geometry is a simple ellipsoid, one would expect that only a few of the lower order terms of the series are sufficient to represent the major aspects of the flaw shape (in fact, just the leading term alone already gives the best-fit sphere). As Fig. 8c illustrates, the first 6 non-zero terms can indeed reconstruct the bulk flaw shape quite well. When we examined the corresponding tangent-plane distance envelope (Fig. 8d), we found that the reconstructed envelope has identical match with the exact
envelope at those 54 nodal points as given by the data. Between the known
nodal points, the flaw surface are completed by least squares curve fitting
through the spherical harmonics polynomials. This can be seen from Fig. 8e
where the periodic oscillations on the error surface (the values plotted on the
vertical axis) of the tangent-plane distance are clearly due to the modulation by
the trigonometrical function contained in spherical harmonics.

In the case of real flaw testing, Fig. 9a shows a perspective view of the 400 x
400 x 200 spheroidal void in titanium that was used in our previous
sizing/classification discussions. Fig. 9b depicts the reconstruction by using 49
series terms and 54 synthetic error-free data. Applying the method to 9 actual
experimental measurements of the spheroid taken in a 2π/3 conical aperture,
and using only 6 non-zero terms in the expansion of eq. (6), a reconstructed
shape was obtained as shown in Fig. 9c. Although the scales of Figs. 9a and 9c
are somewhat different, a comparison of the two does show that the general
flaw shape and orientation were captured quite well by the new algorithm. The
size was also estimated well since along the (x, y, z) coordinate axes the exact
flaw dimensions were (360, 400, 260) respectively, while in the reconstruction
the corresponding values were (375, 497, 260), leading to a maximum size error
of approximately 24%, which is very comparable to that of the unified flaw
sizing algorithm. We also applied this new method to the 400μm crack
considered previously. The results are given in Fig. 10. The reconstructed
shape and orientation were both remarkably good considering the fact that
spherical harmonic functions are not a good set of basis functions for
reconstructing "disk"-like objects. The measured dimensions along the (x,y,z)
axes were (512, 548, 160) respectively, compared to the exact values of (400, 400,
The error in the largest dimension was therefore, 37%. This result was not nearly as good as obtained by the linear least squares/eigenvalue method or the unified sizing algorithm, but still quite acceptable we feel, given the limitations of the method for handling such singular geometries. (Note that the apparent "holes" shown in these figures are merely plotting artifacts and not real features.) We feel that the accuracy of this new sizing method is competitive with the unified sizing algorithm for volumetric flaws and still quite acceptable for cracks. Furthermore, it is fast because it is noniterative, and stable. However, this new method is sensitive to data "outliers" since it, in effect, directly fits the data to the surface location. For cases such as shown above, where the average measurement errors are on the order of 12%, the method performs well. To illustrate the sensitivity of the method to large errors, we reran the test using a data set of the previous 9 measurements of the spheroid plus two more measurements which contained errors of 40-60%. As shown in Fig. 11, the reconstruction quality degraded substantially due to the presence of these outliers. In contrast, the unified sizing algorithm has been shown to be able to handle such outliers considerably better. This is to be expected since the constraint of fitting to a particular (ellipsoidal) shape causes large individual errors to be smeared out, leading to simply a degraded overall sizing estimate.

Finally, we propose two possible improvements to the present algorithm. First, it is observed that while the exact tangent-plane envelope has a very close resemblance to the reconstructed one (Fig. 8d), after the actual surface coordinates are calculated by eq. (8), the resulting flaw shape reconstruction has much larger deviations from the exact shape. This is understandable since the differentiation process in eq. (8) also enhances the "noise". Thus, some
stabilization scheme is needed here. One possibility is to apply a smoothing technique to those derivatives such as curve fitting along both \( \theta \) and \( \phi \) directions. Second, we note that truncation, one of the major drawbacks of series expansion methods, has a great influence on the reconstruction ability of this algorithm. Thus, it would be desirable to be able to take such "incompleteness" into consideration. Converting the ordinary least-squares of eq. (7) into a total least squares [25] to include the modeling of the truncated series is seemingly a promising direction for future implementation.
CONCLUSIONS

In this paper, we presented several new approaches in doing ultrasonic equivalent flaw sizing. First, we considered a linear least squares/eigenvalue scheme which is both computationally efficient and stable. We have also shown that although classification information is not necessary to do equivalent flaw sizing, such information does improve flaw sizing estimates in two ways. First, it allows the use of what are currently the most robust equivalent flaw sizing algorithms that are available. Second, in the case of cracks, it allows systematic reduction of bandlimited-induced errors in the measurements and hence improved results. In addition, we demonstrated another new sizing method, based on an expansion of spherical harmonics, that is noniterative and linear in nature, that can be used as an alternate sizing method, with or without classification information, provided that large individual measurement errors are not present.

We should point out that it is crucial to solve the "zero-of-time" problem [8] for volumetric flaws, since otherwise the necessary $\Delta t (r_e)$ values used by the sizing methods discussed here cannot be obtained. Unfortunately, no completely practical and general solution to this problem currently exists. However, once the effective radius of the flaw is obtained, the methods presented here give a very complete set of tools for efficiently performing equivalent flaw sizing.
ACKNOWLEDGMENTS

We would especially like to thank Professor A. Sedov for his contribution to prior work in the past years. We would also like to thank Professors D. O. Thompson and D. K. Hsu and Mr. S. J. Wormley for supplying us with the experimental volumetric flaw data used here. Similarly, we thank Dr. T. A. Gray for the corresponding crack data. Thanks are also due to Dr. J.-S. Chen for his valuable discussions. This research was supported by the Center for Nondestructive Evaluation at Iowa State University and was performed at Ames Laboratory. Ames Laboratory is operated for the U.S. Department of Energy by Iowa State University under Contract No. W-7405-ENG-82.
REFERENCES


Figure 1a. Volumetric flaw geometry
Figure 1b. Born model for an ellipsoidal inclusion (see Fig. 1a for correspondences)

Figure 1c. Actual time response from an ellipsoidal void (see Fig. 1a for correspondences)
Figure 2a. Crack flaw geometry
Figure 2b. Kirchhoff model

Figure 2c. Actual time response from an artificial circular crack (refer Fig. 2a for A, B and C correspondences)
Spherical Coordinates

$\mathbf{x} = (a, b, c, \theta, \phi, \phi_a)$

$\mathbf{x}_I = (\theta_I, \phi_I)$

Figure 3. Equivalent flaw geometry
Figure 4. Look-up curve for the elimination of finite-bandwidth effects on time measurements $\Delta t$.
Figure 5. Bandwidth effect on Kirchhoff model: Δt correction
Classification of Flaw as Volumetric or Crack-like

+ 

Measurement of $\Delta t$

Born Model:
1-D Inversion determines $\Delta t$ for volumetric flaws

Kirchhoff Model:
Flashpoints determine $\Delta t$ for cracks

Algorithm to Determine:

the Best-Fit Ellipsoid for Volumetric Flaw
or
the Best-Fit Ellipse for Crack Flaw

Equivalent Flaw Shape and Orientation

Figure 6. Flaw sizing scheme with classification information
Volumetric Flaw

Crack Flaw

Impulse Response

Integrated Impulse Response

Measurement of $\Delta t$ from the Integrated Impulse Response (Step Response)

Algorithm to Determine the Best Fit Ellipsoid

Equivalent Flaw Shape and Orientation

Figure 7. Flaw sizing scheme without classification information
Figure 8a. Spherical harmonics algorithm: exact geometry of a 100×200×300 μm ellipsoid
Figure 8b. Spherical harmonics algorithm: reconstruction of a 100x200x300 μm ellipsoid by 54 synthetic data in 4π aperture with 49 terms in expansion series
Figure 8c. Spherical harmonics algorithm:
reconstruction of a 100x200x300 μm ellipsoid
by 54 synthetic data in 4π aperture with 49 terms in
expansion series (first 6 non-zero terms used for
reconstruction)
Figure 8d. Spherical harmonics algorithm: tangent-plane envelope of a 100x200x300 μm ellipsoid
Figure 8e. Spherical harmonics algorithm: error distribution of the tangent-plane envelope of a 100x200x300 μm ellipsoid.
Figure 9a. Spherical harmonics algorithm: exact geometry of a 200x400x400 μm spheroid in titanium.
Figure 9b. Spherical harmonics algorithm:
reconstruction of a 200x400x400 μm spheroid in
titanium by 54 synthetic data in 4π aperture with 49
terms in expansion series
Figure 9c. Spherical harmonics algorithm: reconstruction of a 200x400x400 μm spheroid in titanium by 9 experimental data within limited aperture (6 non-zero terms in expansion series)
Figure 10. Spherical harmonics algorithm: reconstruction of a 400x400 μm disk-shaped cavity in titanium by limited experimental data within limited aperture
Figure 11. Spherical harmonics algorithm: reconstruction of a 200x400x400 μm spheroid in titanium by 11 experimental data within limited aperture (6 non-zero terms in expansion series)
Table 1. Test results of linear least squares/eigenvalue algorithm on synthetic data and experimental data

a. Tested on synthetic data

<table>
<thead>
<tr>
<th>Case</th>
<th>Size (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a</td>
</tr>
<tr>
<td>Thin strip</td>
<td>1</td>
</tr>
<tr>
<td>Long rod</td>
<td>1</td>
</tr>
<tr>
<td>Ellipsoid</td>
<td>100</td>
</tr>
<tr>
<td>Crack</td>
<td>200</td>
</tr>
</tbody>
</table>

b. Tested on experimental data and compared with nonlinear algorithm

<table>
<thead>
<tr>
<th>200x400 µm Spheroid in Titanium, 10 data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Axis</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>Linear Least Squares/Eigenvalue Algorithm</td>
</tr>
<tr>
<td>a</td>
</tr>
<tr>
<td>b</td>
</tr>
<tr>
<td>c</td>
</tr>
<tr>
<td>Non-linear Least Squares Algorithm</td>
</tr>
<tr>
<td>a</td>
</tr>
<tr>
<td>b</td>
</tr>
<tr>
<td>c</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>400 µm circular crack in titanium, 12 data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Axis</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>Linear Least Squares/Eigenvalue Algorithm</td>
</tr>
<tr>
<td>a</td>
</tr>
<tr>
<td>b</td>
</tr>
<tr>
<td>c</td>
</tr>
<tr>
<td>Non-linear Least Squares Algorithm</td>
</tr>
<tr>
<td>a</td>
</tr>
<tr>
<td>b</td>
</tr>
<tr>
<td>c</td>
</tr>
</tbody>
</table>
Table 2. Performance comparisons of linear least squares/eigenvalue algorithm on experimental data with vs. without finite transducer bandwidth error corrections

<table>
<thead>
<tr>
<th>Axis</th>
<th>Size</th>
<th>Size % Error</th>
<th>Angle Deviation in Deg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>400 μm Circular Crack in Titanium, 12 Data without Correction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>400</td>
<td>-8</td>
<td>23</td>
</tr>
<tr>
<td>b</td>
<td>400</td>
<td>-15</td>
<td>23</td>
</tr>
<tr>
<td>c</td>
<td>0</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>400 μm Circular Crack in Titanium, 9 Data with Correction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>400</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>b</td>
<td>400</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>c</td>
<td>0</td>
<td>-</td>
<td>3</td>
</tr>
</tbody>
</table>
Table 3. Comparison of sizing of volumetric flaws and cracks with and without classification information

<table>
<thead>
<tr>
<th>Case</th>
<th>Orientation</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\theta_c$</td>
<td>$\phi_c$</td>
</tr>
<tr>
<td>1. Vol. with class</td>
<td>36</td>
<td>165</td>
</tr>
<tr>
<td>2. Vol. without class</td>
<td>58</td>
<td>175</td>
</tr>
<tr>
<td>3. Exact</td>
<td>(60)</td>
<td>(180)</td>
</tr>
<tr>
<td>4. Crack with class</td>
<td>92</td>
<td>—</td>
</tr>
<tr>
<td>5. Crack without class</td>
<td>102</td>
<td>—</td>
</tr>
<tr>
<td>6. Exact</td>
<td>(90)</td>
<td>(—)</td>
</tr>
</tbody>
</table>
PART III. A NEURAL NETWORK MODEL FOR ULTRASONIC FLAW SIZING
A NEURAL NETWORK MODEL FOR ULTRASONIC FLAW SIZING

Chien-Ping Chiou and Lester W. Schmerr

Center for NDE and Department of Engineering Science and Mechanics

Iowa State University

Ames, Iowa 50011
A multilayered neural network model is used here for solving an inverse sizing problem in the field of ultrasonic nondestructive evaluation. In particular, a feed-forward network trained via the error backpropagation algorithm is shown to be able to invert size and orientation information for circular cracks from time domain ultrasonic data. Test results of the network's performance on both theoretical and experimental data are presented. A new adaptive learning scheme for improving the training speed of such methods is also presented.
INTRODUCTION

One important task in ultrasonic nondestructive evaluation (NDE) is to obtain quantitatively the size, shape and orientation of flaws from ultrasonic measurements. Most such tasks, referred to as inverse flaw sizing problems, are very difficult to solve even under ideal conditions [1-2]. These sizing problems become even more difficult when facing complicated geometries and noisy or incomplete data. Recently, artificial neural networks have emerged as powerful tools that are capable of handling problems of this type. These networks, loosely model the human brain's massive layers of connections, and possess the ability to approximate arbitrary mappings from sets of input-output patterns presentations [3-5]. Many physical phenomena which are fuzzy or are otherwise impossible to be modeled by theory can often be translated into input-output patterns and then solved by neural networks. A priori restrictions on the data can also be implicitly cast in these patterns. Furthermore, neural networks are able to accommodate noise as well as deal with incomplete data. Thus, they are particularly suitable for problems such as the inverse sizing tasks.

Although neural network have been actively pursued for over thirty years, in the late '60s Minsky and Papert [6] proved that the simple networks then under consideration were severely limited in their capability to learn. This limitation resulted in the virtual abandonment of neural network research for many years. However, the recent discovery of new training algorithms such as backpropagation [7-9] has attracted renewed interest in the development of neural networks. Today many successful applications have used this algorithm (see, for example [10] and the references therein). Other neural network models
have also received considerable attention including Hopfield's continuous model for combinatorial optimization [11], Kohonen's associative memory model [12] and the Boltzmann machine [13].

In the field of nondestructive evaluation, the application of neural networks has a relatively long history. The use of a neural network-like model was pioneered in the '70s by Mucciardi and his associates [14-15]. More recently, several authors have investigated the use of neural networks in solving NDE problems such as eddy current crack inversion and classification problems [16-17].

In this work, we first review the nature of multilayered feed-forward networks as trained by the backpropagation algorithm. We then demonstrate the applicability of this neural network approach for solving an ultrasonic sizing problem. In particular, time domain A-scan data are used here to obtain the radius and orientation parameters of an isolated circular crack. One major drawback of using this type of neural network is the extremely long learning (iteration) time consumed in the training process. Thus, we also show that a new learning scheme that incorporates a heuristic adaptive learning method can improve the training speed.
NEURAL NETWORK MODEL

Multilayered Feed-Forward Network

Here we shall use standard multilayered feed-forward networks which are parallel, distributed information structures consisting of layers of processing units (nodes) interconnected together. Each processing unit receives inputs from the outside or from other processing units, processes these inputs through some functional operations, and thereby produces its output value. Associated with each input connection, there is a weight factor which determines the amount of interaction. A typical "bottom-up" multilayered feed-forward network used in this work is shown in Fig. 1a. Note that only the units in the input layer obtain inputs from the outside. Fig. 1b describes two basic components in a processing unit: input weight sum and output activation functions. From presentations of the input-output patterns, the network is trained by a learning algorithm (in our case we apply the backpropagation learning algorithm via the generalized delta rule) so that the weights can be properly adjusted to reproduce the input-output relations.

The first component in a processing unit is to compute the net input \( F \) as the total weighted sum of inputs \( I_j \)

\[
(1) \quad F = \sum_j w_{ij} I_j
\]

in which the output of the \( j \)-th processing node in a previous layer becomes the input \( I_j \) to the \( i \)-th node in the current layer. The weight factor \( w_{ij} \) is associated with the connection between these two nodes. After the summation process an
activation function is applied to F to produce an output O from this i-th node. In principle, any activation function could be chosen. As will discussed later, the backpropagation learning scheme requires evaluating the gradient of the activation function throughout the training process. Thus, it is natural to choose a function that is everywhere differentiable. A popular activation function satisfying this condition is a sigmoid of the form

$$A(F, H, L, K) = \frac{H - L}{1 + \exp(-F \cdot K)} + L$$

Function A performs essentially as a gain gate to restrict the activation (hence the output O) between L and H. Its derivative is equal to

$$\frac{dA(F, H, L, K)}{dF} = \frac{K}{H - L} (H - A)(A - L)$$

which can be easily shown to be bounded between 0 and (H - L)K/4. As we will see, the change of weights rely on the derivative of A, so these bounds also help stabilize the training process. In many situations, a trainable bias term is added to the net node input F in eq. (1) to produce a favorable shift in the activation function value. In this case, the weight factor connected to the bias term is always unity. Also note that K is a scaling factor for the slope of eq. (3) to allow more activation flexibility between the upper bound H and lower bound L. Still, other activation functions may be considered. For example, in [8] a more generalized sigma-pi unit was introduced and has shown some advantages over the simple sigmoid.
Error Backpropagation Learning Algorithm

The learning rule used in this work is essentially the same as that described in [8]. A good account of the derivation of the generalized delta-rule was given there. Here, for completeness we briefly outline it.

The "learning" of the network takes place when a set of, say M, input-output patterns (the training set) are presented to it for weight adjustments. Based on a least squares criteria, the backpropagation algorithm processes each input-output pair in the following two steps. First, the input patterns are presented at the input layer and are fanned out to the processing units in the first "hidden" layer to compute the outputs. These outputs are then passed forward through every layer in the network to calculate the subsequent outputs (note that the nodes at the input layer only serve as distribution nodes and do not involve any processing activities). At the final output layer, the network "learning" ability with respect to the l-th pattern is measured by the least squares error sum

\[
E_l = \frac{1}{2} \sum_i (T_{l,i} - O_i)^2
\]

where \(T_{l,i}\) is the j-th output pattern (target) for i-th node and \(O_i\) is the i-th node output.

In the second step, the error \(E_l\) is propagated backward throughout the network to adjust the weight of the connections. The change of weight for the l-th pattern (we subsequently drop the subscript l explicitly for simplicity) can be shown equal to
where the subscript $i$ denotes the $i$-th processing unit of the "upper" layer and $j$ denotes the $j$-th node in the "lower" adjacent layer. $\lambda$ is a scaling factor called the learning rate. For the connections between the output layer and the last hidden layer, the gradient factor $\delta_i$ is calculated by

$$\delta_i = A'(F_i) (T_i - O_i)$$

in which $A'(F_i)$ is given in eq. (3). For connections between all subsequent layers, $\delta_i$ is

$$\delta_i = A'(F_i) \sum_k \delta_k \Delta w_{kj}$$

where $\delta$ and $\Delta w$ in the summation are for the connections between layers that are one level above the one currently being considered. These forward and backward processing steps are repeated for all $M$ input-output pairs and thereby constitute one complete training circle called an epoch. The weight can be updated every several pattern presentations or can be accumulated and then updated after one epoch is completed. In practice, a large number of epoch iterations are often required to minimize the total cost function

$$E = \sum_{l=1}^{M} E_l$$
The learning quality is usually judged by the level of the cost function $E$ of the training set or of a different testing set. In [8], a simple scheme was proposed to include the influence of previous learning by adding a fraction of the previous weight change, $\eta \Delta w(n-1)$, to the current weight update $\Delta w(n)$. This extra term called "momentum" (in which $\eta$ is called the momentum rate) has shown to improve the training speed and to stabilize the training process on many problems. We will apply this momentum term to all our implementations in this work.

**Neural network as a Function Approximator**

Multilayered neural networks are often viewed as approximators of multidimensional functions. This can be seen by considering the following least squares problem. For a known function $f$ satisfying $t_i = f(x, y_i)$, given $t_i, y_i, i = 1, 2, \ldots, N$ the least-squares solution $x$ is obtained such that

$$F = \sum_i (t_i - f(x, y_i))^2$$

is minimized. Now guess another function $g$. Given $M$ sets of $(t_i, y_i, i = 1, 2, \ldots, N)$, we hope to find an optimal $x$ such that $f \approx g$ and

$$F' = \sum_{j=1}^{M} \left( \sum_{i=1}^{N} [t_i - f(x, y_i)]^2 \right)$$

is also minimized with respect to $t_i, y_i$ and $x$ defined in some domains. Here in our case, $g$ is the layered neural network trained by backpropagation (which is equivalent to the least squares minimization above), and $x$ is the desired weights and biases in the network. The rigorous proof that a neural network is
such a universal function approximator was given by Hecht-Nielsen [18], using Kolmogorov's convergence theorem. Others have also considered this functional approximation from other aspects [19-20]. The beauty of the neural network approach is that even if the function $f$ cannot be obtained precisely (such as through physical laws), one can still approximate it by a neural network model provided that the training input-output patterns are available.

Following this line of approach, we see that backpropagation is actually a gradient descent search scheme (eqs. (6-7)) for locating the least squares solutions in the weight space defined by the input-output patterns of the training set. During an epoch iteration, if weights are updated at each presentation for $N$ patterns, it is equivalent to solving an underdetermined subproblem (eq. (4)) $N$ times and overall solving an overdetermined problem for the whole epoch. A different view point was stated in [5], where the entire network approximation is considered as sequential nonlinear transformations from input pattern space (the input layer) to successive spaces (the hidden layers) and finally into output pattern space.
THE ULTRASONIC FLAW SIZING PROBLEM IN NDE

To apply the neural network model as an inverse problem solver, there are three major steps that must be taken. First, the inverse problem must be converted to a mapping relationship suitable for input-output "sampling". Second, the mapping obtained in the first step must be discretized to generate a training set consisting of input-output patterns. Then the training procedure can be conducted, after by selecting a neural network configuration, by iteratively feeding the network with the input-output patterns. In the following sections we will describe these steps for a particular ultrasonic flaw sizing problem.

Ultrasonic Equivalent Flaw Sizing

The sizing problem considered here is a model-based approach for sizing an isolated crack. In this approach one matches the ultrasonic data with a best-fit "equivalent" simple shape that is able to represent the major aspects of the crack. The simple shape that we will use here is that of a planar, circular crack. A typical ultrasonic signal for such a crack is shown in Fig. 2. One of the prominent features in this crack response is the existence of two large antisymmetrical peaks (labelled A and B in Fig. 2) called "flashpoints". Through the use of the elastodynamic Kirchhoff approximation [21] the time separation, $\Delta t$, between these flashpoints can be directly related to the radius and orientation of the crack [21] as:

$$\Delta t = 4a \cos \eta / c_L$$
in which \( a \) is the crack radius, and \( c_L \) is the incident (longitudinal) wave speed. \( \mathbf{q} \) and \( \mathbf{n} \) are the unit vectors in the incident wave and in the crack normal directions, respectively. If spherical coordinates is used, \( \cos \eta \) can be further expressed as

\[
\cos \eta_i = \sqrt{1 - \left[ \cos \theta_i \cos \theta_n \cos (\phi_i - \phi_n) + \sin \theta_i \sin \theta_n \right]^2}
\]

where \((\theta, \phi)\) are the angular parameters for the above mentioned unit vectors; the subscript \( i \) denotes the \( i \)-th scan direction, and again \( n \) denotes the crack normal (Fig. 3). For a real crack, which is not circular, a measurement of the time interval, \( \Delta t \), between flashpoints can be made at different transducer orientations \((\theta_i, \phi_i)\). Then eqs. (9) and (10) can be used to find the best-fit values of the circular crack parameters \((a, \theta_n, \phi_n)\). Previously, this equivalent flaw sizing approach was carried out via a nonlinear least squares optimization (minimization) procedure [21]. However, nonlinear problems of this type are notoriously difficult to solve, particularly when there are significant experimental errors in the underlying measurements. Here, we will demonstrate that a neural network can provide an effective alternative sizing procedure.

**Functional Decomposition**

Although eqs. (9-10) could be used to construct a set of input-output pairs for training a neural network, there is a serious disadvantage in using these functional forms directly. This is because the incident direction \((\theta_i, \phi_i)\) appear explicitly in these relations so that we would have to give these values when
training the network. This is equivalent to specifying the "scan plan" of the transducer in advance and so the resulting network would only good for a single specific scan plan. However, as we will show in the this section, it is possible to separate the functional dependency on the scan parameters \((\theta_i, \phi_i)\) in \(\Delta t\) from the flaw parameters \((a, \theta_n, \phi_n)\) and train the network on a new set of inputs (derived directly from \(\Delta t\) measurements) that are independent of the scan plan. Thus, a single neural network needs only be trained for all possible scan plans. To see how this simplification can be made, consider the \(\Delta t\) function again. It can be written as

\[
\Delta t_i = \frac{4 a \cos \eta_i}{c_L} = f(a, \theta_n, \phi_n, \theta_i, \phi_i)
\]

where

\[
\cos \eta_i = \sqrt{1 - [\cos \theta_i \cos \theta_n \cos (\phi_i - \phi_n) + \sin \theta_i \sin \theta_n]^2},
\]

or we can alternatively choose to consider

\[
(11) \quad T_i = \left( \frac{c_L \Delta t_i}{4} \right)^2 = a^2 \{1 - [\cos \theta_i \cos \theta_n \cos (\phi_i - \phi_n) + \sin \theta_i \sin \theta_n]\}
\]

to eliminate the dependence of \(c_L\). After some expansion and re-arrangement, the \(\theta_i\) and \(\phi_i\) dependences can also be separated by decomposing eq. (11) in the form

\[
(12) \quad T_i = \sum_j C_j (a, \theta_n, \phi_n) f_j (\theta_i, \phi_i).
\]

Although we could consider the general case given by eq. (12), in our later comparison with experiments, the scan plan is fixed in the y-z plane i.e. \(\phi_i = 90^\circ\)
in eq. (12). Therefore, we will only consider this special case here where eq. (12) reduces to

\[ T_i = \sum_j C_j(a, \theta_n, \phi_n) f_j(\theta_i, \phi_i) \]

in which

\[ C_1 = \frac{a^2}{4} \left( 2 + \cos^2 \theta_n + \cos^2 \theta_n \cos 2\phi_n \right), \]

\[ C_2 = \frac{a^2}{4} \left( 2 - 3\cos^2 \theta_n + \cos^2 \theta_n \cos 2\phi_n \right) = C_1 - a^2 \cos^2 \theta_n, \]

\[ C_3 = -\frac{a^2}{2} \sin \phi_n \sin 2\theta_n, \]

\[ f_1 = 1, \]

\[ f_2 = \cos 2\theta_i, \]

and

\[ f_3 = \sin 2\theta_i. \]

It should be pointed out that this specialization is for convenience only and all the procedures outlined here can be carried out for more general situations.

Eqs. (13-14) form the basis for our simplified sizing method. The steps in this method are as follows. First, measurements of \( \Delta t_i (T_i) \) would be made experimentally. Then, we would solve a linear least squares problem consisting of minimizing the function \( I \) given by

\[ \sum_{i=1}^{N} \left[ T_i - C_j(a, \theta_n, \phi_n) f_j(\theta_i) \right]^2 \]
for the best fit unknown coefficients \( C_j \) \((j = 1, 2, 3)\). Note that unlike earlier procedures \([21]\), this involves solving only a linear least squares problem which can be done in general by standard, efficient algorithms.

The \( C_j \) then became the inputs to the neural network with the desired flaw parameters \((a, \theta_n, \phi_n)\) as the output. If the network has been previously configured and trained on known \( C_j \) and flaw parameters test cases, then to determine the size and orientation of an unknown flaw, we need only to feed the \( C_j \) values obtained, through eq. (15), from the measurements, and record the subsequent output values. Properly configuring and training the network, therefore, is essential to the success of this method. The manner in which we performed this task for our particular sizing problems is given in the next section.

**Network Configuration and Training**

Determining a neural network architecture and making an explicit choice of a training process are tasks for which in general there are no explicit guidelines. Thus, these tasks are problem dependent and involve the use of considerable "trial and error" numerical experiments.

To control our network to a proper size, the orientation angles were restricted in the first quadrant (positive \( x, y \) and \( z \) coordinates). The range for the crack radius was taken as \((0, 1000\mu m)\). We then generated a training set that contained over 300 input-output normalized combinations with uniform sampling in the desired regions. An example of the parameters used is shown in Table 1. Note that when \( \theta_n = 90^\circ \), \( \phi_n \) is arbitrary and is set to \( 0^\circ \) here. The rest of the patterns associated with \( \theta_n = 90^\circ \) are redundant and can be removed from the training set. Thus, the total number of patterns was \( 10 \times 6 \times 6 - 1 \times 5 \times 10 = 310 \).
The network was simulated by C [16] and Fortran programs sequentially executed on a Apollo DN10040 (Prism Architecture) workstation under Domain/OS SR10.1P. After initial trials, one successful network configuration was found which contained three hidden layers with 12 nodes in each layer. The total number of weights in this network was $3 \times 12 \times 2 + 12 \times 2 = 360$. The total number of biases was $12 \times 3 + 3 + 3 = 42$. So the total adjustable parameters were 402. Within the traditional gradient search context, this network would be an objective function of too large a size to be handled in a feasible matrix form.

The selection of the starting weights was purely random. The reason for this is to avoid symmetry, as pointed out by [8]. Also from previous discussions, extremely large initial weight values may pre-saturate the training by causing numerical over-flow in the exponential of the sigmoidal activation function (eq. (2)). This is the so-called "paralysis" problem. Our computational experience shows that a good range for starting weights was from -0.5 to 0.5 as was used here. The scaling factor used in the sigmoidal activation function is 3 so that considerable flexibility of output level is allowed during the iterations. In viewing the behavior of the activation function (eq. (2)), we note it would take infinite steps for the node response to reach either 0 or 1. Unfortunately, these values coincide with some output pattern values. One practical solution to this problem is to relax the bounds for the activation function. We found empirically that -0.1 for a lower bound and +1.1 for a higher bound were good choices. Another important setting that needs to be determined is the training batch size; that is the number of pattern presentations before weights are updated. As pointed out in [8], the true gradient descent method that backpropagation implements requires weight updates taking place at each
complete epoch presentation. However, from our own experience and other reports in the literature, we found that training by individual patterns generally give faster convergence.

The initial learning rate and momentum rate were chosen as the "standard" values of 0.7 and 0.9, respectively [8]. From further trials, it was seen that fixed learning and momentum rates in standard backpropagation often causes severe oscillations in convergence and can result in no learning. A simple strategy was used to overcome this difficulty. In this strategy, the learning history was traced. Once the magnitude of oscillation over some preset tolerance was detected, the learning and momentum rates were then decreased by another preset value. In the final successful run, the percentage of allowed overshooting is 20%, and the decreasing percentage is also 20% for both learning and momentum rates. The training process was terminated after about 14 CPU days, which corresponds to an extremely large 1.7 million iterations. The final network least squares error $E$ in eq. (8) was sufficiently low at 0.0001367.

**Sizing Results**

The trained network was then tested on two cases. The first case consisted of an input-output set of 1920 perfect patterns different from those in the training set. In the second case, the inputs were two sets of experimental data. These data consisted of 12 $\Delta t$ values taken from a 400x400 \( \mu m \) pill box of 1:10 ratio that was used to simulate a circular crack.

In the perfect testing, the first 90 patterns where the radius $a$ was beyond the training range, produced errors of the angular parameter $\phi_n$ which were constantly large while the estimates of the radius $a$ and the other angle $\theta_n$ were stable. It was observed that as the crack size increased, the absolute error in the
size estimate also increased. We found that 84% of the output values of node 1 (radius $a$) had absolute errors less than 100 $\mu$m, the "size sampling rate". In terms of percent error, there were 91% output values within 20% error. For the orientation angle, we measured the error by the angle deviation between the desired crack normal and the output normal. An encouraging 94% of all the patterns fell within $18^\circ$ deviation (which was equal to the "angular sampling rate"). The statistical histograms of these results were shown in Fig. 4a and 4b.

For the testing on experimental data, the results are shown in Table 2. In the first case the actual $\Delta t$ values obtained experimentally were used. In the second case the $\Delta t$ values were corrected to eliminate a systematic error due to finite transducer bandwidth [22]. In both cases, eq. (15) was used to obtain the three C-coefficients of eq. (14). By using the C's values, the network produced 17% size error in crack radius for case one and 30% for case two. The corresponding errors in orientation were $15^\circ$ and $0^\circ$, respectively. While both of these results are acceptable, it is surprising that the "corrected" data of case two gave a perfect orientation estimate but a poorer size estimate than case one. Although we examined the C's values in both cases, no obvious reason for this result was apparent. Of course, it is possible that the least squares procedure "accidentally" gave a better size fitting estimate for the C's in case one than in case two because of the distribution of errors in the data of this particular example.
Although backpropagation coupled with the delta rule has earned the reputation as a very powerful learning algorithm, there are two major disadvantages associated with it. One is the accuracy of learning: the network does not always learn well. The other is the learning speed: for larger networks, the learning process usually is very slow. As shown in previous sizing results, the iterations can take an extremely long time counted in CPU days even in mainframe computers. Here we are concerned with the second issue. We will show that an accelerated training scheme can be implemented with relatively little additional computation effort over "standard" backpropagation.

Adaptive Strategies

We start by drawing an analogy of the generalized delta rule to the method of steepest descent in the optimization field. Recalling eqs. (4,8) we wish to minimize the error function

\[
E(w) = \frac{1}{2M} \sum_{j} \sum_{i} \{[T_i - O_i(w)]^2 \}
\]

in which \( T_i \) is the target, \( O_i \) is the output at output layer of the network, and \( w \) is the matrix of weights in the network. By a multi-dimensional Taylor's expansion, we have

\[
E(w) = E(w_0) + G^T(w)\Delta w + [\text{high order terms}]; \ G = \frac{\partial E}{\partial w}.
\]

For a steepest descent direction, we choose \( \Delta w = -\lambda G \), where \( \lambda \) is a properly adjusted factor in each iteration such that \( E(w) \) is minimized. Then
which is equivalent to the generalized delta rule (eq. (5)), except that in the standard delta rule the quantity $\lambda$ is fixed. The steepest descent method is known to perform rather sluggish (the convergence rate is only linear) near the minimum due to step size vanishing as $G$ goes to zero. However, it and its backpropagation analog can be practically used if we do not require that the absolute global minimum be found but instead place a criteria like $E < \varepsilon$ on the process. Even so, methods to produce a faster convergence rate are desirable for many applications.

Within the context of backpropagation, there are two basic approaches to improve the training speed. The first approach is to go for higher order corrections (Eq. (17)) such as the direct use of second order derivatives in implementing the Newton method [23], or alternatively by approximating the network Hessian as in, for example, the Levenberg-Marquardt method [24]. The other approach is to develop various means for adjusting the learning rate. This includes using line search along the gradient to determine the optimal learning rate [25], or applying a heuristic trace-back algorithm to adjust learning rates for individual patterns [26]. The first approach is supported on solid theoretical ground but requires computationally expensive evaluations of second derivatives or products of first derivatives which is likely to compromise the speed gained from using this approach. Likewise, the line search of [25] also takes a fair amount of time in trials. The heuristic scheme in [26] consumes less extra computational effort than other methods. However, it needs additional storage to keep track of the past state for each patterns. In this work, we shall
propose another heuristic variation of the learning rate update that requires even fewer numerical steps than [26], and also gives good improvement of training time.

Our adaptive strategy consists of the following two components: (i) adaptive "trace-back" leaning update: similar to [26], the past learning history is recorded as a guide line to adjusting learning rates in future steps of learning. The difference here is that we only trace the total root mean square (RMS) error $E$ in eq. (16), not every input-output pattern. Thus, the present scheme relies on the global least squares prediction rather than local fluctuations from individual pattern errors. If $E$ is found to be constantly decreasing in the past $N$ iterations, then in the $N+1$-th iteration both the learning rate and momentum rate are raised up by a multiplicative factor $\alpha^N$, where $\alpha$ is a real power base greater than 1. On the other hand, if $E$ in the $N$-th iteration was higher than that in the $N-1$-th iteration, then in the $N+1$-th step both rates are dropped by another power factor $\beta^N$ with power base $\beta$ less than 1. All three parameters ($N$, $\alpha$ and $\beta$) need to be empirically determined and, unfortunately, are problem dependent. The initial values can be obtained from other "calibrator" problems if such problems are available or from trial runs of a standard backpropagation. (ii) multi-stage safe bounds: as shown in Fig. 5, a typical training path starts with a rapid learning state (region A) and followed by a long tail of slow convergence (region B). Usually region B consumes most of the learning time. Exceedingly high learning rates often cause constant oscillation in this region while rates that are too small may elongate the slow-learning tail even more. Thus, it would be advantageous to estimate (such as from the trial runs mentioned above) the learning path and subdivide the possible path into several stages. In
each stage, we set up higher and lower safe bounds for the learning and
momentum rates to prevent the search step from overshooting.

Since these two strategies require only a small amount of logical reasoning
in a computer program, the extra computation effort spent here is negligible
when compared with the entire iteration process. Thus, the training time saved
by this method is almost "free".

**Benchmark Testing**

This adaptive scheme has been tested by three mapping problems: The 2 to
1 exclusive-or (XOR, parity-2) problem (Table 3a), 3 to 1 parity-3 problem (Table
3b), and a 3 to 2 random mapping set (Table 3c). The parity problems, consisting
of binary input-output pairs, are classical benchmark testers. The random set
simulates a partial example of a more realistic mapping of physical problems.
This set contains 20 discrete patterns between 0 and 1. The parity problems are
known solvable by using as many nodes in one hidden layer as the input nodes.
Hence, for XOR problem we use a 2-2-1 (2 input nodes, 1 hidden layer with 2
nodes and 1 output node) paradigm and 3-3-1 for parity-3 problem. For the
random set, we found 3-12-2 is sufficient and also solvable by 3-5-5-2.

We ran each problem several times with different initial weights by using
both the adaptive and standard (fixed rates) versions. The three adaptive
parameters N, α and β were fixed after the first few trials. In the parity
problems, for only 2 cases did the adaptive scheme fail to converge faster than
standard backpropagation. The training time reduction varied from 3% to 220%
for the remaining successful cases. We believe that the optimal parameter
settings can always be found if an exhaustive trial is conducted. The test results
of the random set showed even more encouraging results; there the adaptive
learning update and multi-stage safe bound reduced training iterations from 2 to 5 times. A typical performance comparison is given in Fig. 6.

A general trend of learning was observed. For simpler problem such as XOR, the accelerating $\alpha$ and decreasing $\beta$ power bases could be relatively high values. Also, the training speed was highly dependent on the distribution of the initial weights. In the more difficult random set problem $\alpha$ and $\beta$ had smaller values and the problem was less dependent on the initial weights. This can be understood as follows. In a simple problem, the weight space is also of simple structure so that the search step length can be extended longer with less chance missing a hidden valley and a corresponding minimum. The distribution of initial weights is important since a "good" choice can likely get one very close to the solution. The situation is just the reverse for complicated problems such as the case of the random mapping.

It is well known that the solution domain of an inverse problem is better defined if more data describing the system's characteristics are available. Similarly, a mapping should be easier to be approximated if the input pattern number exceeds the output pattern number. We justified this intuitive thought by another benchmark experiment. We deleted the last output pattern (i.e., the last column in Fig. 3c) in the random set and treated it as a 2-2 mapping. The new mapping then became much more difficult to solve. In all trials the iteration time increased significantly (by 4 times in average).

**The Sizing Problem Revisited**

Since in modern safety predictions via fracture mechanics the flaw size is the most crucial factor, the crack radius output of the network is the most significant parameter. Also motivated by the findings in the last section, by
excluding the output patterns associated with the orientation parameters $\theta_n$ and $\phi_n$ from the training set we should be able to produce a network that can learn the size parameter more efficiently. In order to regularize the function approximation by the network in the situation when the crack size is particular small, we added 20 more patterns sampled in that region. The network configuration then became 3-12-12-12-1.

From the initial trials, it was found that the the accelerating factor $\alpha$ has to be kept low to prevent network paralysis. The final $\alpha$ was set to 1.05 along with decreasing factor $\beta = 0.9$ and back-tracing step $N = 3$. The starting learning rate was 0.7 and momentum rate was 0.5. Two safe bounds were set at 0.005 and 0.001 respectively. The training status was monitored every 100 iterations. The adaptive algorithm converged quite fast to reach an absolute accumulated error of 10 at the 40100-th iteration, while the standard algorithm required 35200 more iterations. After the 70000-th iteration, error reductions in both algorithms became rather slow although they were still monotonically decreasing. Since the RMS error reached the same low level as in the previous single-step adaptive result, we terminated the iteration process at RMS error = 0.000012 on the 81000-th iteration.

We tested the adaptively trained networks by the 1920 testing sets mentioned previously. The comparison between the present adaptive method and previous single-step adaptive scheme (Fig. 4a) is shown in Fig. 7a. We see that the adaptive network responses are much better where almost all the errors are in the 0% to 10% error category. In Fig. 7b we show a comparison between the performance of the adaptive method with the standard backpropagation results for a network that was configured identically (using only size as an
output). In this case the performance of both cases was nearly identical. However, the standard backpropagation convergence was 35200 iterations slower than the adaptive scheme.
CONCLUSION

We have demonstrated, for the first time, how neural networks can be used to solve equivalent flaw sizing problems. A general method for training such networks in a manner independent of the scan plan used for the interrogating transducer has been given. Results with this approach for sizing cracks are encouraging, based on tests with both synthetic and experimental data for plane circular cracks. Methods for improving the quality and speed of the training process were also discussed and have been shown to prove effective for the circular crack sizing problem.
ACKNOWLEDGEMENTS

One of the authors (CPC) would especially like to thank Mr. Jim Mann for his general assistance and valuable discussions. This research was supported by the Center for Nondestructive Evaluation at Iowa State University and was performed at Ames Laboratory. Ames Laboratory is operated for the U.S. Department of Energy by Iowa State University under Contract No. W-7405-ENG-82.
REFERENCES


Figure 1a. A multilayered feed-forward network with full connections

\[ O = \frac{1}{1 + e^{-F}} \]

Figure 1b. A typical processing unit
Figure 2. Time response from a planar crack
Figure 3. Equivalent flaw geometry
91% patterns have error below 20%  

94% Patterns have error below 18 deg.  

Figure 4a. Size error distribution on testing set  

Figure 4b. Orientation error distribution on testing set
Figure 5. A typical training history
Figure 6. Performance comparisons of adaptive step length scheme vs. standard (fixed step length) scheme
Figure 7a. Size error distribution comparisons between adaptive scheme (size output only) and single-step adaptive scheme (size and orientations outputs)
Figure 7b. Size error distribution comparisons between adaptive scheme (size output only) and standard scheme (size only)
Table 1. Function sampling for training set

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Radius a</th>
<th>$\theta_n$</th>
<th>$\phi_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampling Point</td>
<td>10</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Actual Range</td>
<td>100 – 1000 (µm)</td>
<td>0 – 90 (deg.)</td>
<td>0 – 90 (deg.)</td>
</tr>
<tr>
<td>Normalized Range</td>
<td>0.1 – 1</td>
<td>0 – 1</td>
<td>0 – 1</td>
</tr>
<tr>
<td>Actual sampling</td>
<td>100 (µm)</td>
<td>18 (deg.)</td>
<td>18 (deg.)</td>
</tr>
<tr>
<td>Increment</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
</tr>
</tbody>
</table>
Table 2. Test results of trained network on experimental data

<table>
<thead>
<tr>
<th>Case</th>
<th>Size %Error</th>
<th>Angle Deviation in Deg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 data without error correction</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td>9 data without error correction</td>
<td>18</td>
<td>3</td>
</tr>
<tr>
<td>9 data with error correction</td>
<td>30</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 3a. Parity-2 (XOR) training set for benchmark testing

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 0</td>
<td>0</td>
</tr>
<tr>
<td>0 1</td>
<td>1</td>
</tr>
<tr>
<td>1 0</td>
<td>1</td>
</tr>
<tr>
<td>1 1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3b. Parity-3 training set for benchmark testing

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 0 0</td>
<td>1</td>
</tr>
<tr>
<td>0 0 1</td>
<td>0</td>
</tr>
<tr>
<td>0 1 0</td>
<td>0</td>
</tr>
<tr>
<td>0 1 1</td>
<td>1</td>
</tr>
<tr>
<td>1 0 0</td>
<td>0</td>
</tr>
<tr>
<td>1 0 1</td>
<td>1</td>
</tr>
<tr>
<td>1 1 0</td>
<td>1</td>
</tr>
<tr>
<td>1 1 1</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 3c. A random training set for benchmark testing

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>0. 0. 0.</td>
<td>1. 0.2</td>
</tr>
<tr>
<td>0. 2 1.</td>
<td>0. 0.6</td>
</tr>
<tr>
<td>0. 1 0.</td>
<td>0. 0.0</td>
</tr>
<tr>
<td>0. 1. 1.</td>
<td>1. 0.1</td>
</tr>
<tr>
<td>1. 0.6 0.</td>
<td>0. 0.7</td>
</tr>
<tr>
<td>1. 0. 1.</td>
<td>1. 0.5</td>
</tr>
<tr>
<td>1. 0.5 0.</td>
<td>1. 0.3</td>
</tr>
<tr>
<td>1. 1. 1.</td>
<td>0. 0.2</td>
</tr>
<tr>
<td>0.1 0.9 0.1</td>
<td>0.5 0.9</td>
</tr>
<tr>
<td>0.8 0.2 0.8</td>
<td>0.2 0.5</td>
</tr>
<tr>
<td>0.6 0.1 0.3</td>
<td>0.4 0.0</td>
</tr>
<tr>
<td>0.5 0.5 0.1</td>
<td>0.3 0.4</td>
</tr>
<tr>
<td>0.9 0.3 0.6</td>
<td>0.6 0.6</td>
</tr>
<tr>
<td>0.3 0.2 0.1</td>
<td>0.0 0.3</td>
</tr>
<tr>
<td>0.4 0.8 0.6</td>
<td>0.2 0.1</td>
</tr>
<tr>
<td>0.5 0.3 0.2</td>
<td>0.8 0.8</td>
</tr>
<tr>
<td>0.4 0.5 0.2</td>
<td>0.1 0.5</td>
</tr>
<tr>
<td>0.0 0.3 0.0</td>
<td>0.4 0.3</td>
</tr>
<tr>
<td>0.9 0.9 0.9</td>
<td>0.0 0.5</td>
</tr>
<tr>
<td>0.1 0.1 0.5</td>
<td>0.4 0.7</td>
</tr>
</tbody>
</table>
GENERAL CONCLUSION

In this work, we have presented a number of new approaches for performing the tasks of ultrasonic flaw classification and sizing. These approaches make extensive use of a wide range of "tools", including elastic wave scattering models, experimental instrumentation, signal processing techniques, numerical optimization methods and neural computing.

In Part I, we showed that mode-converted diffracted signals can be used to develop a simple classification technique in a contact quasi-pulse-echo configuration. The classification concept was verified both by experimental data and by rigorous elastodynamics theories. With further development, this technique could be employed as a practical classification tool in real field inspection problems.

A series of investigations conducted in Part II have resulted in several new sizing methods and related results within the general framework of equivalent flaw sizing. First, we showed that a new linear least squares/eigenvalue method could be used for equivalent flaw sizing estimates, eliminating the difficulties associated with previous non-linear approaches. Second, we demonstrated how equivalent flaw sizing could be done without having flaw classification information and evaluated the "leverage" that such classification information provides in the sizing process. For a crack, we showed that classification information is particularly useful since one can use this information and a simple "look-up curve" to eliminate systematic measurement errors due to finite transducer bandwidth effects. Finally, we considered a new sizing technique, based on a spherical harmonics expansion,
which can potentially give more detailed flaw shape information than the highly constrained sizing methods considered previously.

Even with these advances, a number of problems and opportunities remain for equivalent flaw sizing methods. The "zero-of-time" problem for volumetric flaws is still incompletely resolved and it is a problem where advances could have significant impact in other areas such as imaging. The extension of equivalent flaw sizing methods to anisotropic materials, such as composites, and to surface-breaking or near-surface flaws are also places where future work could produce significant rewards.

In the last part of this dissertation, we have shown that neural networks can be employed to solve ultrasonic sizing problems. A functional decomposition procedure was illustrated to obtain a scan-plan-independent mapping suitable for neural network training. We demonstrated how the training process could be carried out practically for a circular crack, using a neural network's nature as an universal approximator. Then we showed that the training could be accelerated by utilizing a heuristic adaptive learning procedure. The success in using this method for sizing a planar crack from ultrasonic data should be considered as only a first step towards treating more general sizing problems with this approach. Also, considerable challenges remain in choosing efficient network architectures and in speeding up the training process. The extension of neural networks to classification problems is also an area where these methods should have considerable promise.
ACKNOWLEDGMENTS

I would like express my sincere gratitude to my major professor, Dr. Lester W. Schmerr, for his guidance and support in this work as well as in my entire graduate study; particularly in the "final rush".

Thanks are also due to Professors Donald O. Thompson, David K. Hsu, William Lord and Hsien-S. Hung for serving as my committee members; to Messrs. Steve Nugen, James Mann, Sung-J. Song, Samuel J. Wormley, Dr. R. K. Chapman, and my senior colleagues Drs. Jer-S. Chen and Lat-S. Koo for their generous assistance and discussions. I am also indebted to many of my friends for their help and encouragement. The support from both the Department of Engineering Science and Mechanics and Center for Nondestructive Evaluation at Iowa State University is acknowledged.

Finally, my deepest appreciation goes to my parents and my wife, to whom this dissertation is dedicated.


