The application of Fermat's principle for imaging anisotropic and inhomogeneous media with application to austenitic steel weld inspection

George D. Connolly*, Michael J. S. Lowe, and J. Andrew G. Temple

UK Research Centre in NDE
Imperial College
London SW7 2AZ
UK

Stanislav I. Rokhlin
The Ohio State University
Edison Joining Technology Center
1248 Arthur E Adams Drive
Columbus OH 43221
USA

Abstract

Ultrasonic inspection in anisotropic and inhomogeneous media has long presented a challenge because of the complex steering of ultrasonic paths. An approach is presented in which the true geometry of a previously-used austenitic steel weld is distorted so that, from one viewing location, all ultrasound travels in straight lines with a constant isotropic velocity. The mapping from the real space to this distorted space is accomplished using Fermat's theorem of least travel time applied through ray tracing. Applications specific to inspection design and data interpretation for manual ultrasonic inspection of welds in austenitic steel plates are given. Validation of some intermediate results is provided using finite element analysis.

*Corresponding author, electronic mail g.connolly05@imperial.ac.uk
1. Introduction

Ultrasonic inspection in anisotropic and inhomogeneous media presents a challenge because of the complex steering of ultrasonic paths in the material. By using Fermat’s principle of least travel time in such media, we present an approach in which true geometry of an inhomogeneous nature is distorted so that, from one viewing location, all ultrasound travels in straight lines with a constant isotropic velocity. Although the approach is of general applicability, in this paper we demonstrate it in the specific context of ultrasonic inspection of austenitic steel welds.

Historically, austenitic steel has been used in the construction of the pressure boundaries of nuclear reactors. These are typically sections of up to 250mm thick in civil nuclear plant. Austenitic steel is used because of its high fracture toughness and resistance to corrosion and is increasingly being considered for use in other industrial sectors: modern conventional fossil-fuelled plants and offshore oil and gas. In all these safety-critical industries, non-destructive inspection is used to ensure that the plant enters service without any defects above a certain size and during service, regular inspections are used to verify that no defects have grown to an unacceptable size. Cracks are the defect of most concern and particularly those having a sizeable through-wall extent. Such cracks, if present, in nuclear plant structures are almost always in the welds and associated heat affected zone and not in the bulk material. Radiography finds such cracks most efficiently only if the beam of radiation aligns well with the plane of the crack, whereas ultrasound is more effective when the beam aligns normal to the crack. In practice, both NDT methods are used but there is much interest in developing confidence in ultrasound NDT sufficiently to avoid the use of radiography, particularly for in-service inspection.

Welds of austenitic steel tend to solidify due to thermal gradients with large grains which can be millimetres across and centimetres long. The welding technique used determines the shape and pattern of these grains. Ultrasonic inspection makes a compromise between the resolution achievable and the amount of signal noise, or clutter, that arises from scattering within the grainy material. Higher frequencies give shorter wavelengths and better sizing accuracy but noise increases with increasing frequency. Typical ultrasonic frequencies used in nondestructive inspection are between 2 and 5 MHz, giving wavelengths of between about 1.5 - 0.6mm for shear
waves and 3 - 1.2mm for longitudinal waves respectively. The literature provides further information regarding grain boundary scattering\(^1\), a phenomenon with which this paper is not concerned.

At these frequencies, the wavelengths are smaller than the grain size found in the welds. Instead of behaving like waves in an isotropic homogeneous material, where the ultrasound travels in straight lines with an isotropic velocity until it encounters a change in elastic constants, the elastic waves in austenitic weld metal, at the typical frequencies used in inspection, follow curved paths dictated by the orientation of the grains and their elastic constants. This is known as beam steering. Although all metal crystals are intrinsically elastically anisotropic, this beam steering effect is not found in most fine grained metals such as ferritic steel or aluminium because the grains are randomly oriented and the ultrasonic wavelength, being much larger than the grain dimensions, averages over many grain orientations to yield effectively isotropic material constants.

Difficulties with ultrasonic inspection of welds in austenitic material have been explored since about 1976\(^2\). Much work over the following 15 years sought to understand this problem and achieve better inspection capability\(^3\)-\(^6\).

Despite some improvements in techniques, the problem of ultrasonic inspection of austenitic welds is far from solved. The advent of increasingly cheap computing power and new ultrasonic imaging techniques based on ultrasonic arrays provides the stimulus for a new look at this problem.

The microscopically detailed structure of any particular weld is unknown. It depends on parameters including the speed of the welding tool, the chemical composition of welding material compared to the base material, the heat input, the geometry of the weld and its orientation during welding. For a given type of welding, such as manual metal arc or tungsten inert gas welding, the overall pattern of grains is known, even if the microscopic detail is not. We use this knowledge together with general understanding and some simplification of the elastic constants to create models of the welds which can occur in practice.

The aim of this paper is to use ray tracing to predict the paths of ultrasound through model welds, then to use this capability to produce distorted maps of the weld. In
these distorted maps, the real weld geometry, with its inhomogeneous and anisotropic nature, is replaced with a distorted weld geometry in which, viewed from the origin of the map, all materials are homogeneous and isotropic. Ultrasound propagating from the map origin does so in straight lines with a constant isotropic velocity everywhere within the map. These are called Fermat maps because they are based on Fermat's principle of least travel time. The potential of this approach is shown in examples of how these can be used to aid inspection design.

2. Theory

The material in an austenitic weld is both anisotropic and inhomogeneous. Theory governing the propagation of elastic waves in such media is given together with the theory for reflection and refraction at material boundaries (abrupt changes in the elastic constants of the material).

(a) Anisotropy and slowness

Some basic theory governing elastic wave propagation in anisotropic and inhomogeneous media is presented in standard texts\textsuperscript{7,8}. The wave equation can be written in terms of the derivative of the stress and the acceleration of a particle as:

\[
\sigma_{ij,j} = C_{ijkl} u_{k,jj} = \rho \ddot{u}_i
\]  

where \(\rho\) is the material density. Here the usual summation convention of summation over repeated indices is implied, and differentiation with respect to time is denoted by a superscript dot, and differentiation with respect to a spatial coordinate of a quantity is denoted with a comma, thus:

\[
\sigma_{ij,j} \equiv \sum_{j=1}^{3} \frac{\partial \sigma_{ij}}{\partial x_j}
\]  

The relationship between stress \(\sigma_{ij}\) elastic constants \(C_{ijkl}\) and strains \(\varepsilon_{kl}\) is given by:

\[
\sigma_{ij} = C_{ijkl} \varepsilon_{kl}
\]  

The elastic constants \(C_{ijkl}\) have certain permutation symmetries:
\[ C_{ijkl} = C_{klij} = C_{jikl} = C_{ijlk} \]  \hspace{1cm} (2.4)

Plane waves can be written either in terms of the direction of their phase (wavevector) as

\[ u_k = A p_k \exp \left[ i \omega \left( \frac{n_j x_j}{V} - t \right) \right] \]  \hspace{1cm} (2.5)

or in terms of their slowness \( S \) as

\[ u_k = A p_k \exp \left( i \omega (S_j x_j - t) \right) \]  \hspace{1cm} (2.6)

Other terms in (2.5) and (2.6) are:

(a) Displacement in the medium \( u \)

(b) The amplitude of the wave \( A \)

(c) The polarization of the wave \( p \)

(d) The angular frequency of the wave \( \omega \)

(e) Complex number \( i = \sqrt{-1} \)

(f) Direction cosines of the phase (wave) vector \( n \)

(g) Spatial position \( x \)

(h) Phase velocity \( V \)

(i) Time \( t \)

Substitute (2.5) into (2.1) to get:

\[ (\rho V^2 \delta_{ik} - \Gamma_{ik}) p_k = 0 \hspace{1cm} \text{with} \hspace{1cm} \Gamma_{ik} = n_j n_l C_{ijkl} \]  \hspace{1cm} (2.7)

where \( \Gamma_{ik} \) is the Green Christoffel tensor. An alternative description is, from substituting (2.6) into (2.1) to get:
This description is useful for finding the sextic equation (see Appendix) governing the unknown component of the slowness normal to an interface in reflection and transmission problems, giving either (2.9) or (2.10) as the eigenvalue equation to solve:

\[ \rho \delta_{ik} - \alpha_{ik} \right) p_k = 0 \quad \text{with} \quad \alpha_{ik} = S_j S_i C_{ijkl} \]  

\[ |\rho V^2 \delta_{ik} - \Gamma_{ik}| = 0 \]  

\[ |\rho \delta_{ik} - \alpha_{ik}| = 0 \]  

(b) Voigt notation

It is usual to find elastic constants written with just two indices, which run from 1 to 6, rather than the four used in the tensor notation, each of which runs from 1 to 3. This contraction is called Voigt notation and is obtained by the following mapping of the first and second pairs of indices as \( C_{ijkl} \rightarrow C_{ij;kl} \), according to Table 1.

(c) Elastic constants

Austenitic steel comes in various types with different chemical compositions. The basic material often used in nuclear power vessels is denoted type 304 but can have a variety of carbon contents; ranging from less than 0.03% to about 0.1%; those with higher carbon contents have greater yield strengths. The typical alloy content of 304 is 18-20% Cr and 8-12% Ni. Other types of austenitic steel are the basic 304 with differing amounts of alloying elements. For example, another common type is 316 with 16-18% Cr, 10-14% Ni and 2-3% Mo. These different chemical compositions give different elastic wave speeds and differing intrinsic anisotropies. In some, but not all, cases the elastic constants in welds can be reasonably taken to have transversely isotropic crystal symmetry. In this case, elastic waves propagating in the basal plane (xy-crystal plane) have a velocity independent of direction whereas those propagating in any other direction exhibit velocities which depend on the direction\(^{10,11} \). The elastic constants used in this work are those used by Roberts\(^{12} \), appropriate to a transversely isotropic material (see Table 2).
As a first approximation, the elastic constant z-axis (rotational symmetry axis) of the transversely isotropic elastic constants is taken to lie in the plane perpendicular to the weld direction in the steel plate. This ensures that ultrasonic waves propagating in this plane, perpendicular to the weld direction, remain in that plane despite any beam steering effects in the plane. This allows results for application of the Fermat maps to be presented easily for interpretation of manual ultrasonic inspection data presented in section 6. However, real welds typically have a layback of about 10º or so, depending on the speed of the welding tool. This corresponds to the z-axis pointing about 10º out of the plane perpendicular to the weld direction.

3. Ray tracing model

The large grains in an austenitic weld tend to follow the heat flow during solidification, growing initially from the weld preparation faces\(^1\). This reference shows the sorts of grain patterns which are associated with different welding techniques. If there is significant epitaxial growth across the boundary between successive weld runs then the grains will grow into long, curved columnar grains. Ray tracing is a useful technique for studying the propagation of elastic waves in anisotropic and inhomogeneous media. Ogilvy created a computer ray tracing model to study ultrasonic wave propagation in austenitic steel welds of the manual metal arc type and used this to identify pulse-echo rays\(^1\), ultrasonic beam profiles\(^1\), focused beams\(^1\) and reflection from planar defects in austenitic welds\(^1\). Finally, the model was used to identify the ultrasonic paths connecting any two points in a three-dimensional weld\(^1\). The latter developments of this model are described by Hawker\(^1\).

In this paper, the principle of ray tracing is based on Fermat's principle, which applies to any wave propagation, electromagnetic or elastic. In a medium with spatially varying elastic constants, the ultrasonic equivalent of optical distance \(D\) is the travel time given by taking a curved path between A and B and:

\[
D = \int_A^B \frac{ds}{V_g(r)} \tag{3.1}
\]

Fermat's principle is that signals travel along a path with a travel time \(D\) that has a stationary value with respect to variations in the path A-B\(^2\) (usually a minimum,
although in principle a maximum is possible). To model this in a weld, steps are taken from an initial point along the direction of the group velocity for the desired propagation mode at that point, usually taken to be the transducer or a specific part of a defect, such as a crack tip. At each point the elastic constants are obtained from a lookup table or from a parametric formula (see section 3(c)). Between the new point and the previous point an artificial boundary is constructed across which the ray is refracted (see discussion below). A constant time step is used, giving steps of different length in different parts of the material. This is a key step in creating the maps described in section 5.

(a) Group velocity

The group velocity is derived as follows:

\[ V_g = \frac{\partial \omega}{\partial k} \]  \hspace{1cm} (3.2)

Suppose there is a general form for the equation of the wavevector surface:

\[ F(k_i, \omega) = 0 \]  \hspace{1cm} (3.3)

Then

\[ \frac{\partial F}{\partial k_i} + \frac{\partial F}{\partial \omega} \frac{\partial \omega}{\partial k_i} = 0 \]  \hspace{1cm} (3.4)

And components of group velocity are:

\[ V_{gi} = \frac{\partial F}{\partial k_i} \left(\frac{\partial F}{\partial \omega}\right) \]  \hspace{1cm} (3.5)

Taking (2.7) and multiplying it by \( p_i \) and expanding gives

\[ \rho \omega^2 = C_{ijkl} p_i p_j k_l k_i \]  \hspace{1cm} (3.6)

And differentiating gives
\[ 2\rho\omega \tilde{\omega} = C_{i\ell m}\ell_{j}P_{i}P_{m}(k_{j}\delta_{ij} + k_{i}\delta_{ij}) \]  

(3.7)

So:

\[ V_{i}^{e} = \frac{\rho_{i}p_{i}k_{m} + p_{m}k_{i}}{2\rho\omega} = \frac{L_{i}}{2\rho\nu} \]  

(3.8)

where \( \nu \) has been used to denote phase velocity and

\[ L_{i} = C_{i\ell m}p_{j}(p_{j}n_{m} + p_{m}n_{j}) \]  

(3.9)

Then the direction cosines of the group velocity are given by:

\[ \frac{L_{i}}{\sqrt{L_{1}^{2} + L_{2}^{2} + L_{3}^{2}}} \]  

(3.10)

(b) Refraction and reflection at real or artificial boundaries

When an elastic wave meets a planar boundary, the component of the slowness projected onto the boundary plane is preserved and this is used to solve for the components of slownesses perpendicular to the boundary. The equation governing this is sextic\(^9\) so there may be up to 6 reflected and 6 refracted waves at an interface between two materials I and II with differing elastic constants. The xy-plane is chosen as the boundary plane with the z-axis positive pointing into medium I and with wave incidence in medium I. The key points about the various roots are\(^9,22\):

(a) All slowness vectors are confined to the incidence plane. (Snell's law).

(b) The sextic has real coefficients and therefore has roots which are real or complex conjugates two by two, in one of the following forms:

(i) 3 +ve and 3 -ve real roots

(ii) 4 +ve and 2 -ve real roots

(iii) 2 +ve and 4 -ve real roots

(iv) 4 real and 2 complex roots
(v) 2 real and 4 complex roots

(c) In the case of 3 real roots for the sextic for material I then the +ve roots correspond to reflected waves and the -ve roots correspond to transmitted waves (though in some cases, see below, the -ve roots may also be required). For material II the -ve roots correspond to transmitted waves (though again some +ve roots may be required). Which modes to include in the solution of the boundary equation are chosen according to the group velocity (or energy vector). Those that point towards +ve z-values in material I and towards -ve z-values in material II are used.

(d) The net flow of energy parallel to the interface is zero. As the direction of the group velocity (energy vector) tends towards the parallel to the interface, so the wave amplitude tends to zero.

(e) Solutions of the sextic yield real or complex values. In cases of complex values the imaginary part is always perpendicular to the interface so damping is always exponential away from the interface.

The energy flux vector is given by

\[
F_i = \frac{1}{4} \left| A^2 \right| \omega C_{ijkl} \left\{ p_j^* p_k k_i + p_j p_k^* k_i^* \right\}
\]

where a real \( m_3 \) implies a real slowness vector and thus a homogeneous wave and a complex \( m_3 \) implies a complex slowness vector and an inhomogeneous wave with:

\[
m = m' + i m''
\]

\[
u_k = A p_k \exp \left\{ -\omega \sum_{q=1}^3 m_q'' x_q \right\} \exp \left[ i \omega \left( \sum_{r=1}^3 m_r' x_r - t \right) \right]
\]

which is a plane wave travelling toward

\[
n' = \frac{m'}{|m|}
\]

with phase velocity
\[ V' = \frac{1}{|m'|} \]  
(3.15)

and decaying exponentially in the direction

\[ n'' = \frac{m''}{|m''|} \]  
(3.16)

with decay constant

\[ \alpha = \frac{|m'|}{|m|} \]  
(3.17)

For arbitrary orientation of \( n' \) and \( n'' \) this is an inhomogeneous wave. For this type of wave:

(a) When \( n' \cdot n'' = 0 \), the wave is evanescent.

(b) When \( n' \) is parallel to \( n'' \), the wave is homogeneous and damped.

(c) From the six roots found in each material, only three are physically acceptable (see, for instance, Federov\textsuperscript{10}).

(d) For homogeneous waves (real roots) the group velocity (energy vector) is used to choose roots.

(e) For inhomogeneous waves the flux vector component normal to the interface is always zero\textsuperscript{23,24}, so the sign of the imaginary part of the slowness determines which are reflected or refracted waves: for reflected waves \( m_3'' > 0 \) and for transmitted waves \( m_3'' < 0 \).

(f) Complex polarization vectors correspond to elliptical polarization.

(c) The weld model

Although the detailed structure of welds can be predicted from parameters of the welding process\textsuperscript{25,26}, any weld encountered in the field will have a generally unknown microscopic grain structure. Hence the aim is to use as few parameters as possible to characterise welds of differing type. A useful parameterisation of the orientation of
elastic constants in a single V-butt weld between two plates was given by Ogilvy\textsuperscript{27} and used by Halkjær et al.\textsuperscript{28}. This form is given by:

\[
\theta = \arctan\left(\frac{T(D + z \tan \alpha)}{x^n}\right) \tag{3.18}
\]

where \(\theta\) is the angle of the crystallographic \(z\)-axis with respect to the through-wall direction of the weld, \(T\) is proportional to the tangents of the crystallographic \(z\)-axes at the weld preparation boundaries, \(D\) is the distance from the weld centre to the edge of the weld preparation on the narrowest side (this is half the weld root for a symmetric weld), \(z\) is the through wall location, \(\alpha\) is the angle of the weld preparation, \(x\) is the distance from the weld centre and \(\eta\) is a parameter with \(0 \leq \eta \leq 1\) governing the rate of change of angle with \(x\) (see Fig. 1). For \(\eta = 1\), the grains are normal to the weld preparation.

This weld model gives us a description of the spatial variations of the material properties and will eventually be used in sections V and VI in order to make the ray propagation calculations.

4. Boundary ray behaviour validation

A Finite Element (FE) simulation has been used as a validation of the prediction of generated wave properties at a single interface by the semi-analytical model described above. We first view the generalities of the FE modelling and then the processing of results.

(a) Discretisation

Here we are only concerned with the propagation of bulk waves. The FE models, generated using the finite element software package ABAQUS\textsuperscript{29}, use a standard two-dimensional spatial discretisation composed of square elements with linear shape functions and four nodes with each node having two degrees of freedom with respect to displacement. The motion of the nodes perpendicular to the plane of operation is prohibited, and the plane strain condition has been enforced.

The discretisation in the computations were such that the following condition was satisfied\textsuperscript{30}:
\[ \lambda_{\text{min}} \geq 8\Delta d \] (4.1)

where \( \lambda_{\text{min}} \) is the shortest of the wavelengths in the model, and \( \Delta d \) is the dimension of the element. This is minimise erroneous wave propagation distortions and to improve modelling accuracy. The time step \( \delta t \) is chosen such that\(^{30}\):

\[ \delta t \leq 0.8 \frac{\Delta d}{V} \] (4.2)

An absorbing region\(^{31}\) is placed around the interface to eliminate reflections so that results can be more easily extracted from the simulations and that a smaller FE model can be employed for our purposes. The absorbing layers are placed at the truncated boundary and dissipate energy according to a damping factor, whose magnitude is governed by a cubic asymptotic function:

\[ F_{\text{damping}} \propto \left( \frac{x}{w_{\text{absorb}}} \right)^3 ; 0 \geq x \geq w_{\text{absorb}} \] (4.3)

where \( w_{\text{absorb}} \) is the width of the absorbing layer and \( x \) is the distance through the layer, starting at the interior. The constant of proportionality depends on the size of the model. A schematic is shown in Fig. 2.

(b) Simulations of wave interaction with a single interface

The models were square of which the sides were 60mm in length and divided into 600 elements. Waves were introduced at a frequency of 800kHz. For the materials used, the number of elements per wavelength was safely in excess of the requirement outlined above, being no fewer than about 20 elements per wavelength.

For the simulations shown in this paper, the wave is introduced into the structure from a series of nodes along the top of the nonabsorbing region, in the form of a five cycle tone burst modified by a Hanning window. Each of the nodes oscillates with the same amplitude in the direction given by the polarisation vector, computed according to the phase vector as outlined in section II. For the production of a coherent wave, a time delay is prescribed as a function of the node index \( n \):
where $w$ is the width of the oscillating area and $\theta$ is the angle between the phase vector and the perpendicular. The method explained here holds true for both isotropic and anisotropic materials.

(c) Processing of results

The validation presented here will involve only the P and the SV waves due to the plane strain condition leading to the inability of supporting the SH wave. We also find that in a transversely isotropic material, the SH wave will not transfer energy to either of the other two wave modes as long as the active plane coincides with that of the anisotropy of the material, thus by such a definition, its polarisation vector will always be perpendicular to that of the corresponding P and SV waves.

We compare, where possible, the following properties of the P and SV waves reflected from the interface and of the P and SV waves transmitted past the interface: phase angle; phase velocity; polarisation angle and group angle. All angles are measured anti-clockwise from the positive horizontal axis. The results of the comparison are presented for the cases of an interface between transversely isotropic steel and isotropic mild steel, with the wave source in the isotropic mild steel (see Fig. 3), and an interface between two transversely isotropic steels at different orientations of elastic constants (Fig. 4).

To determine the polarisation angle of a wave in the FE simulation, the displacement histories from nodes at either side of the interface are extracted and can be processed to verify the properties of the generated waves. These nodes are positioned slightly away from the interface to allow the waves of different modes to separate from one another before passing the monitoring position. We observe the horizontal and the vertical displacement, apply the Hilbert transform to obtain the envelope of the signal, defined as

$$H(x) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{u(t)}{t-x} \, dt$$

(4.5)
to the received signals, and then compute the maximum magnitudes. The tangent of the ratio will yield the polarisation vector, and thus its angle, which can be compared to the theory.

The group angle is measured from the FE simulations by observing the change in the position of the part of the wave with the largest absolute amplitude between snapshots corresponding to known times.

The phase vector and the phase velocity are compared via the phase spectrum method, presented by Sachse and Pao\textsuperscript{32}. Here the vector is computed through measurements of displacement histories from three nodes close to one another arranged in the shape of a right-angle; the horizontal component of velocity being computed from the pair of nodes lying horizontal to one another and the vertical component of velocity being taken from the pair lying vertical to one another. Let us denote the received time signal functions as $a(t)$ and $b(t)$. Then, upon the assumption that there is no frequency-dependent damping, the Fourier transforms of the signals are:

$$F(a(t)) = |F(a(t))\exp[i\phi_a]| \quad \text{and} \quad F(b(t)) = |F(b(t))\exp[i\phi_b]|; \quad (4.6)$$

being the product of the spectra of amplitude and of the phase $\phi$. The difference in the phase spectra $\Delta\phi = \phi_a - \phi_b$ and thus the phase velocity can be expressed in terms of a unit, here denoted as $n$, that varies from 0 to $2\pi$ in a step size that is dependent on the number of data points originally recorded:

$$c_p(n) = \frac{2\pi n}{\Delta t(\exp[i\phi_a]\exp[i\phi_b])}; \quad 0 \leq n \leq 1 \quad (4.7)$$

The components in turn are calculated from the difference between the two phase spectra, and are combined to yield the phase velocity and the phase vector. A more thorough description of this method, and an overview of an alternative known as the amplitude spectrum method, may be found in Pialucha\textsuperscript{33}.

Good agreement has been found between the predicted reflected and transmitted wave properties for both cases, as indicated by the FE models across a range of different phase angles. Consistency in the methodologies of both techniques has been demonstrated and some of the small differences within the comparison can be
explained by the nature of the coarse spatial discretisation of the FE modelling as compared to the relatively fine temporal discretisation of the ray tracing. It has also been observed that the effectiveness of the absorbing region lessened with higher angles of incidence. Consequently, for input phase angles of higher than 315°, reliable results were difficult to extract and were more likely to be corrupted by other signals reflected from the absorbing boundary.

Given that the ray tracing model is but a succession of interfaces and that in the limiting case, ray propagation can be considered as interaction with an unbounded number of boundaries between pairs of homogeneous materials, it is concluded here that the validation is applicable to the whole ray tracing model.

5. Fermat mapping process

Fermat maps are drawn to allow the visualisation of the particular space as perceived from the transformation origin, which is usually either a ray source or a potential defect location. At first we shall consider the mapping process for a space containing a weld. A grid of horizontal and vertical lines is imposed upon the area. At each intersection lies a node, and the transformation process is applied to each node in turn. Here we observe the mapping process for a single point. Using an iterative process based on an educated initial guess, all the possible Fermat paths connecting this point to the transformation origin are found. It is noted that the anisotropy of the material can result in up to three eligible paths though in practice, computational efficiency is vastly improved if one assumes that there be only one path joining the two points.

For each Fermat path found, the ray tracing algorithm described previously will note the phase angle at which the ray left the source and the time taken for the ray to reach the target node (see Fig. 5). In transformed space, a ray is allowed to leave the transformation origin with this same phase angle and to propagate for this same length of time at a chosen phase velocity, which is usually that of the appropriate wave mode in the homogeneous material. At the end of this ray is the mapped target point. The points in mapped space are then joined together in order of increasing initial phase angle. This process is repeated to map ray paths to nodes, boundaries, cracks and any other points of interest as required. Examples of generated Fermat maps are shown in
Fig. 6, and for this section, the weld parameters are: $T=1.0$; $D=2.0$; $\eta=1.0$ and $\alpha=\arctan(0.4)$.

(a) Properties of transformed space

Some universal properties of rays in mapped space have been noted. It is already known by the definition of this Fermat space that every ray passing through the transformation origin travels at a constant velocity in a straight line. In Fig. 7, quasi-compression rays are projected from an omnidirectional source at equally spaced phase angles in a model consisting of a region composed of inhomogeneous austenite and two regions of isotropic ferrite. Upon transformation, the rays are seen to trace simpler paths. It is emphasised here that neither wave speed nor direction change at the weld boundaries in mapped space and thus they would not be apparent to the rays in Fermat space since the entire structure is now of a quasi-isotropic material.

Reciprocity has also been observed between any pair of points joined by a ray. If, for instance, there are $n$ Fermat paths leaving a transformation origin at phase angles $\phi_1, \ldots, \phi_n$ taking times $\tau_1, \ldots, \tau_n$ to reach the target and we were to reverse the roles of these points, we would find that there would be $n$ Fermat paths taking the same times $\tau_1, \ldots, \tau_n$ to reach the new target, having left the new source at angles $\phi_1+\pi, \ldots, \phi_n+\pi$.

(b) Inaccessible areas

It has often been noted (see for instance Hudgell and Gray\textsuperscript{34}) that certain areas of the weld are more difficult to inspect than others and that other areas are almost inaccessible from some transducer locations via ultrasonic waves. The transformation process offers a different interpretation as to the existence of such unobservable areas in the weld.

The ray tracing diagram in Fig. 8(a) shows the emission of omnidirectional SH waves (for illustration purposes) from a point source on the inspection surface. The lower right area of the weld, labelled $i$ on the figure, cannot be inspected because rays that are attempting to access that area meet an interface between the weld metal and the parent metal at such an angle such that no real transmitted ray (of the same mode) can be found that satisfies Snell’s law and thus the ray is terminated. Hence an absence of
grid in transformed space is indicative of the presence of an inaccessible area (see Fig. 8(b)).

The area labelled $\hat{u}$ in Fig. 8 is also inaccessible to the transducer since it falls in a shadow created by the geometry of the weld. In mapped space, it is given that rays must travel in straight lines from the transformation origin, and so it can clearly be seen that the convex protrusion at the top of the weld is responsible for the shadow. The grid, however, could still be generated by allowing rays to leave and subsequently re-enter the structure through extrapolation of material properties beyond the upper and lower boundaries. For this purpose, the ray tracing function does not enforce any physical boundaries to the edges of the structure with the exception of one parallel to the bottom face, passing through the point $p$, whose coordinates are defined thus:

\[
p(x) = D_L - (D_L + D_R) \sin(\alpha_L) \tag{37}
\]

\[
p(z) = D_L - (D_L + D_R) \cos(\alpha_L) \tag{38}
\]

The purpose of this boundary, where all rays terminate upon contact, is to prevent the ray from entering a region where (3.18) cannot be applied.

Due to material inhomogeneity, rays might also terminate before reaching certain areas of the weld not due to geometric reasons but due to their interaction with the nonphysical boundaries introduced by the ray tracing function, as described above in section III. They may encounter such a boundary past which they are unable to transmit, where there are no roots corresponding to the input wave mode. In reality, the greater part of the energy is reflected, thus deflecting the ray away from the blind region.

(c) Multiple paths

Earlier it was noted that there may be up to three Fermat paths joining a given pair of points. This situation would occur wherever the anisotropy of a material would allow waves with different phase vectors to travel with the same group vector. Here we examine the consequences with respect to the Fermat mapping.
The transformation process may result in the multiplication of a region in original space. Any object falling into this region would be seen as many times as there are rays accessing the position of the object. Any object falling within the darker shaded region in Fig. 9(a) would be seen once in each of the labelled regions in Fig. 9(b). The phenomenon is demonstrated once more in Fig. 10. There are three possible ray paths joining the source to the tip \(c_0\) of the crack-like defect within the weld due to beam-steering. After transformation, the crack tip splits into three images \(c_1\), \(c_2\) and \(c_3\) such that the aforementioned rays follow straight paths.

6. Simplification of ultrasonic inspection

This section presents an application of the transformation of space to the problem of ultrasonic inspection of austenitic welds in a conceptual form, using data gathered from FE simulations. An inspection device may record A-scans from a weld, various prominent signals will result and one is tasked with matching these signals to certain features within the structure to locate a potential defect. In this section, the weld parameters used were: \(T=1.0; D=2.0; \eta=1.0\) and \(\alpha=\arctan(0.32)\).

In mapped space, certain reflected signals may be correlated to the corners of the weld, as shown in Fig. 11. Isochrones are drawn on the mapped weld diagram with the knowledge that they are circles centred about the origin of transformation, as opposed to their complex shapes in unmapped space. In Fig. 11(a), it is thus established that three such corners (labelled \(i\), \(iii\) and \(iv\)) are responsible for several salient signals. The corner labelled \(ii\) lies within an inaccessible region of the weld and cannot explain the signal to the left of that caused by \(iii\). In Fig. 11(b), all the prominent signals are accounted for by the weld features save that similarly positioned to the left of the signal caused by \(vi\).

The same process is used to locate the potential defect in Fig. 12. The isochrones intersect the far weld boundary in two places in both Fig. 12(a) and Fig. 12(b). Since the angle of incidence of the interrogating wave is known, it can be decided which of the intersections is the more likely location of the defect. This method can be extended to include mode converted reflections by the substitution of the mapped welds of Fig. 6 into Fig. 11. Such distorted weld maps for specific transducer
locations in specific weld geometries could be supplied to manual ultrasonic inspectors to aid signal interpretation.

7. Conclusions and discussion

This paper puts forward a new method of space transformation using Fermat’s principle for ultrasonic imaging in a medium that is inhomogeneous and anisotropic. The method is of wide general applicability but it has here been demonstrated within the specific application of the improved inspection of austenitic steel welds. The paper has focused on the generation and the basic properties of the resulting transformed space from a previously developed weld model. The transformations also offer visual explanations and alternative visualisations of the weld interior that are inaccessible or that are accessed multiple numbers of times.

In addition, the paper has been particularly concerned with the validation of intermediate results through the comparison of certain properties of generated waves at a single interface within the FE model and within the ray tracing model. The comparisons have been very consistent for the phase vectors and polarisation vectors for various scenarios, to include anisotropic materials and mode-converted waves. We have also discussed that the successful validation can be extended to the entire ray tracing model on the grounds that the model is a repeated application of the single-boundary problem.

It is believed that the large body of knowledge of methods associated with conventional imaging, many of which are applicable to isotropic homogeneous materials, may eventually be exploited to improve material inspection by means of the production of a quasi-homogeneous material via the Fermat mapping process. The concept shown in section VI is one such example. Future work in this area might involve tailoring these methods more closely to specific needs of industry.
This research was undertaken as part of the research program of the UK Research Centre in NDE (RCNDE) with support from the EPSRC.
Appendix: The sextic equation

The allowed values of reflected and refracted slownesses at a planar boundary are found from a sextic equation. This is obtained following Rokhlin et al.\textsuperscript{9}, and our equation (8). The z-axis is taken normal to the planar boundary and incidence is taken to be in the yz-plane:

\[ G_{ik} = C_{ijkl} m_j m_i - \rho \delta_{ik} \quad (A.1) \]

Expanding and collecting terms yields:

\[ G_{11} = C_{31} m_1^2 + 2C_{41} m_m m_2 + 2C_{51} m_m m_3 + 2C_{61} m_m^2 + 2C_{71} m_2 m_3 + C_{81} m_3^2 - \rho \]

\[ (A.2) \]

\[ G_{12} = C_{32} m_1^2 + (C_{12} + C_{26}) m_m m_2 + (C_{14} + C_{56}) m_m m_3 + C_{62} m_m^2 + (C_{44} + C_{55}) m_2 m_3 + C_{81} m_3^2 \]

\[ (A.3) \]

\[ G_{13} = C_{33} m_1^2 + (C_{13} + C_{55}) m_m m_2 + (C_{14} + C_{45}) m_m m_3 + C_{63} m_m^2 + (C_{44} + C_{55}) m_2 m_3 + C_{81} m_3^2 \]

\[ (A.4) \]

\[ G_{22} = C_{66} m_1^2 + 2C_{66} m_m m_2 + 2C_{46} m_m m_3 + C_{22} m_2^2 + 2C_{22} m_2 m_3 + C_{44} m_3^2 - \rho \]

\[ (A.4) \]

\[ G_{23} = C_{65} m_1^2 + (C_{66} + C_{23}) m_m m_2 + (C_{36} + C_{45}) m_m m_3 + C_{23} m_2^2 + (C_{23} + C_{44}) m_2 m_3 + C_{44} m_3^2 \]

\[ (A.5) \]

\[ G_{33} = C_{55} m_1^2 + 2C_{45} m_m m_2 + 2C_{35} m_m m_3 + C_{44} m_m^2 + 2C_{35} m_2 m_3 + C_{55} m_3^2 - \rho \]

\[ (A.6) \]

For incidence in the yz-plane set \( m_1 = 0 \) then solve:

\[
\begin{bmatrix}
G_{11} & G_{12} & G_{13} \\
G_{12} & G_{22} & G_{23} \\
G_{13} & G_{23} & G_{33}
\end{bmatrix} = 0 \quad (A.7)
\]

which is a sextic in \( m_3 \) with six roots that occur in complex conjugate pairs. Writing \( \beta \) for \( m_3 \) to simplify the notation, (A.7) is first expanded as:

\[
G_{11} \begin{bmatrix} G_{22} & G_{23} \\ G_{23} & G_{33} \end{bmatrix} - G_{12} \begin{bmatrix} G_{13} & G_{23} \\ G_{13} & G_{33} \end{bmatrix} + G_{13} \begin{bmatrix} G_{12} & G_{22} \\ G_{12} & G_{22} \end{bmatrix} = 0 \quad (A.8)
\]

where \( G_{11} = a_0 + a_1 \beta + a_2 \beta^2 \) and \( a_0 = C_{66} m_2^2 - \rho \), \( a_1 = 2C_{56} m_2 \) and \( a_2 = C_{55} \).

Define:
\[
\begin{vmatrix}
G_{22} & G_{23} \\
G_{23} & G_{33}
\end{vmatrix}
= c_0 + c_1 \beta + c_2 \beta^2 + c_3 \beta^3 + c_4 \beta^4
\]  
(A.9)

with \( G_{12} = d_0 + d_1 \beta + d_2 \beta^2 \) and \( d_0 = C_{68} m_2^2 \), \( d_1 = (C_{46} + C_{25}) m_2 \) and \( d_2 = C_{45} \). Also define:

\[
\begin{vmatrix}
G_{12} & G_{23} \\
G_{13} & G_{33}
\end{vmatrix}
= h_0 + h_1 \beta + h_2 \beta^2 + h_3 \beta^3 + h_4 \beta^4
\]  
(A.10)

with \( G_{13} = f_0 + f_1 \beta + f_2 \beta^2 \) and \( f_0 = C_{64} m_2^2 \), \( f_1 = (C_{46} + C_{25}) m_2 \) and \( f_2 = C_{45} \). Also define:

\[
\begin{vmatrix}
G_{12} & G_{22} \\
G_{13} & G_{23}
\end{vmatrix}
= g_0 + g_1 \beta + g_2 \beta^2 + g_3 \beta^3 + g_4 \beta^4
\]  
(A.11)

Finally, if we also define:

\[
S_1 = \left\{ a_0 + a_1 \beta + a_2 \beta^2 \right\} \left\{ c_0 + c_1 \beta + c_2 \beta^2 + c_3 \beta^3 + c_4 \beta^4 \right\}
\]  
(A.12)

\[
S_2 = \left\{ d_0 + d_1 \beta + d_2 \beta^2 \right\} \left\{ h_0 + h_1 \beta + h_2 \beta^2 + h_3 \beta^3 + h_4 \beta^4 \right\}
\]  
(A.13)

\[
S_3 = \left\{ f_0 + f_1 \beta + f_2 \beta^2 \right\} \left\{ g_0 + g_1 \beta + g_2 \beta^2 + g_3 \beta^3 + g_4 \beta^4 \right\}
\]  
(A.14)

then the required sextic in \( \beta \) is given by:

\[
S_1 - S_2 + S_3 = 0
\]  
(A.15)
References


Ogilvy, J. A. (1992) “An iterative ray tracing model for ultrasonic nondestructive testing,” NDT&E Int. 25, 3-10


List of figures

FIG. 1. Weld parameters of the model

FIG. 2. Schematic of the FE validation model. Numbers refer to the thickness of that section in terms of the number of elements

FIG. 3. Validation of (a) phase vectors of reflected waves, (b) phase velocities of reflected waves, (c) polarisation vectors of reflected waves, (d) group vectors of reflected waves, (e) phase vectors of transmitted waves, (f) phase velocities of transmitted waves, (g) polarisation vectors of transmitted waves and (h) group vectors of transmitted waves plotted against phase angle of an incident SV wave from an isotropic mild steel to a transversely isotropic steel at an orientation of 24°. Dashed vertical lines indicate critical angles.

FIG. 4 Validation of (a) phase vectors of reflected waves, (b) phase velocities of reflected waves, (c) polarisation vectors of reflected waves, (d) group vector of reflected wave, (e) phase vectors of transmitted waves, (f) phase velocities of transmitted waves, (g) polarisation vectors of transmitted waves and (h) group vector of transmitted wave plotted against phase angle of an incident SV wave from a transversely isotropic steel at an orientation of 13° to a transversely isotropic steel at an orientation of 44°. Dashed vertical lines indicate critical angles.

FIG. 5 Illustration of the Fermat mapping process of a point \( b \) from (a) unmapped space to (b) Fermat space from the ray of time length \( \tau \) whose source is \( s \).

FIG. 6 Examples of generated Fermat maps for a ray source, whose position is indicated by the dark square, at (50,60) mm to the spatial origin for (a) P waves without mode conversion (b) P waves with mode conversion to shear (c) SV waves without mode conversion (d) SV waves with mode conversion to longitudinal and (e) SH waves without mode conversion. The reflected space is below the backwall.

FIG. 7 Ray tracing through a structure, showing paths in (a) unmapped space simplifying in (b) Fermat space. The ray source uses P waves and is (8,20) mm from the origin.
FIG. 8 The labelled areas are inaccessible using SH waves from the source at (50,60) mm relative to the weld origin; area $i$ is inaccessible due to ray interaction with the boundary and area $ii$ is inaccessible due to the weld geometry; (a) ray tracing diagram and (b) Fermat space diagram.

FIG. 9 The darker shaded area in (a) is seen thrice from the source emitting SV waves at (45,60) mm to the origin. In (b) mapped space, this area becomes three separate areas, labelled $a$, $b$ and $c$. White areas are inaccessible.

FIG. 10. The upper end of the crack in (a) is seen a multiple number of times, thus in (b) mapped space, the point $c_0$ becomes transformed to three images $c_1$, $c_2$ and $c_3$.

FIG. 11 Matching of prominent signals within time traces from FE simulations to known features within the weld for (a) P waves and (b) SV waves.

FIG. 12 Matching of prominent signals that were not accounted for by weld features in fig. 11 Isochrones are drawn on the mapped weld for (a) P waves and (b) SV waves to identify possible locations for the feature responsible for the reflected signals.
TABLE 1: Voigt notation for the contraction of indices: \( C_{ijkl} \rightarrow C_{ij;kl} \rightarrow C_{LM} \) The contraction for the first pair of indices \( ij \) is shown. That for the second pair, \( kl \) is similar. For example, \( C_{1231} \rightarrow C_{65} \)

<table>
<thead>
<tr>
<th>Original indices</th>
<th>Contracted indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>( i = j = 1, \text{L}\rightarrow 1 )</td>
<td>( i = 2, j = 3, \text{L}\rightarrow 4 \leftrightarrow \text{L}, i = 3, j = 2 )</td>
</tr>
<tr>
<td>( i = j = 2, \text{L}\rightarrow 2 )</td>
<td>( i = 1, j = 3, \text{L}\rightarrow 5 \leftrightarrow \text{L}, i = 3, j = 1 )</td>
</tr>
<tr>
<td>( i = j = 3, \text{L}\rightarrow 3 )</td>
<td>( i = 1, j = 2, \text{L}\rightarrow 6 \leftrightarrow \text{L}, i = 2, j = 1 )</td>
</tr>
</tbody>
</table>
TABLE 2: Material properties for the transversely isotropic material. Voigt notation applies

<table>
<thead>
<tr>
<th>Material parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{11}$</td>
<td>$249 \times 10^9 \text{Nm}^{-2}$</td>
</tr>
<tr>
<td>$C_{12}$</td>
<td>$124 \times 10^9 \text{Nm}^{-2}$</td>
</tr>
<tr>
<td>$C_{13}$</td>
<td>$133 \times 10^9 \text{Nm}^{-2}$</td>
</tr>
<tr>
<td>$C_{33}$</td>
<td>$205 \times 10^9 \text{Nm}^{-2}$</td>
</tr>
<tr>
<td>$C_{44}$</td>
<td>$125 \times 10^9 \text{Nm}^{-2}$</td>
</tr>
<tr>
<td>$C_{66}$</td>
<td>$62.5 \times 10^9 \text{Nm}^{-2}$</td>
</tr>
<tr>
<td>density $\rho$</td>
<td>$7.85 \times 10^3 \text{kgm}^{-3}$</td>
</tr>
</tbody>
</table>
FIG. 1. Weld parameters of the model
FIG. 2. Schematic of the FE validation model. Numbers refer to the thickness of that section in terms of the number of elements.
FIG. 3. Validation of (a) phase vectors of reflected waves, (b) phase velocities of reflected waves, (c) polarisation vectors of reflected waves, (d) group vectors of reflected waves, (e) phase vectors of transmitted waves, (f) phase velocities of transmitted waves, (g) polarisation vectors of transmitted waves and (h) group vectors of transmitted waves plotted against phase angle of an incident SV wave from an isotropic mild steel to a transversely isotropic steel at an orientation of 24°. Dashed vertical lines indicate critical angles.
FIG. 4 Validation of (a) phase vectors of reflected waves, (b) phase velocities of reflected waves, (c) polarisation vectors of reflected waves, (d) group vector of reflected wave, (e) phase vectors of transmitted waves, (f) phase velocities of transmitted waves, (g) polarisation vectors of transmitted waves and (h) group vector of transmitted wave plotted against phase angle of an incident SV wave from a transversely isotropic steel at an orientation of 13° to a transversely isotropic steel at an orientation of 44°. Dashed vertical lines indicate critical angles.
FIG. 5 Illustration of the Fermat mapping process of a point \( b \) from (a) unmapped space to (b) Fermat space from the ray of time length \( \tau \) whose source is \( s \).
FIG. 6 Examples of generated Fermat maps for a ray source, whose position is indicated by the dark square, at (50, 60) mm to the spatial origin for (a) P waves without mode conversion (b) P waves with mode conversion to shear (c) SV waves without mode conversion (d) SV waves with mode conversion to longitudinal and (e) SH waves without mode conversion. The reflected space is below the backwall.
FIG. 7 Ray tracing through a structure, showing paths in (a) unmapped space simplifying in (b) Fermat space. The ray source uses P waves and is (8,20) mm from the origin.
FIG. 8 The labelled areas are inaccessible using SH waves from the source at (50,60) mm relative to the weld origin; area $i$ is inaccessible due to ray interaction with the boundary and area $ii$ is inaccessible due to the weld geometry; (a) ray-tracing diagram and (b) Fermat space diagram.
FIG. 9 The darker shaded area in (a) is seen thrice from the source emitting SV waves at (45,60) mm to the origin. In (b) mapped space, this area becomes three separate areas, labelled $a$, $b$ and $c$. White areas are inaccessible.
FIG. 10. The upper end of the crack in (a) is seen a multiple number of times, thus in (b) mapped space, the point $c_0$ becomes transformed to three images $c_1$, $c_2$ and $c_3$. 
FIG. 11 Matching of prominent signals within time traces from FE simulations to known features within the weld for (a) P waves and (b) SV waves.
FIG. 12 Matching of prominent signals that were not accounted for by weld features in fig. 11. Isochrones are drawn on the mapped weld for (a) P waves and (b) SV waves to identify possible locations for the feature responsible for the reflected signals.