NDT&E International 142 (2024) 103026

Contents lists available at ScienceDirect

NDT and E International

journal homepage: www.elsevier.com/locate/ndteint

Guided waves in long range nondestructive testing and structural health monitoring: Principles, history of applications and prospects

Peter Cawley

Imperial College, London, SW7 2AZ, UK

A B S T R A C T

The early work on long range guided wave nondestructive testing (NDT) was largely done in industry, and multiple applications were developed before a surge in academic interest starting in the early 1990s. This paper reviews the history of the field, covering the topics that must be addressed in the design of a test procedure, as well as applications in both NDT and Structural Health Monitoring (SHM). The use of guided waves in long range NDT is relatively mature, and largely limited to simple structures in applications that are difficult or impossible to address by conventional point inspection methods. Point methods are impractical in most large area SHM applications where it is necessary to detect isolated defects since it is not possible to cover the whole area of interest with transducers. Guided waves therefore have great potential in SHM as they enable a large area of structure to be monitored from a limited number of transduction locations, so reducing the added weight, cost and complexity of the attachment to the structure. It is also possible to improve the sensitivity of the test relative to that obtained in one-off inspection by monitoring the change in signals with time using baseline subtraction. Current SHM deployments are largely in the same types of structure as those where guided wave NDT is used, but there is substantial potential for increasing the range of applications. The paper concludes with a review of the research and development needs for this potential to be realised.

1. Introduction

There has been burgeoning interest in the application of guided waves to nondestructive testing (NDT) and structural health monitoring (SHM) since early work in the 1960s. Fig. 1 shows the number of papers output by a Scopus search for 'Guided waves OR Lamb waves' and 'nondestructive testing OR structural health monitoring OR inspection' by year. The red squares show the total number of papers published in each year, while the black circles show the number yielded by a similar search but with 'structural health monitoring' omitted. Interest in the area increased rapidly, starting in the 1990s, and the fraction concerned with structural health monitoring has increased to approaching half the total in recent years. Publications now show signs of plateauing at around 500–600 per annum; a Google Scholar search using the same search terms yields around double the number of papers.

Worlton at General Electric [1,2] was probably the first to recognise the potential of guided waves in nondestructive testing. Subsequent early work in the field was largely done in industry and public research labs in USA and Europe and was targeted at specific industrial applications; this is reviewed in later sections.

Later, Rose at Drexel University, and latterly Penn State University, led academic work and popularised the subject, identifying the huge potential given by the multiple modes and measurement possibilities. His book, first published in 1999 [3] and extended in a second edition under a modified title in 2014 [4] has become a standard reference text. However, the potential of guided waves comes at a cost of complexity as multiple modes can lead to very complicated signals that are difficult to interpret reliably, and this must be managed [5].

Guided waves can be used for defect detection in both short- and long-range regimes, each of which has been extensively researched. In the short-range methods, the region inspected is essentially confined to the volume below the transducer(s); methods in this category include leaky Lamb wave inspection of composite materials [6] and acoustic microscopy [7] in which a leaky surface wave is generated by the lens. This paper reviews long-range methods in which the region inspected extends away from the transducer by distances from ~0.3 m in the case of typical Rayleigh wave tests to up to ~1000 m in the case of some rail inspections.

The main attraction of long-range guided wave inspection is that it enables a large area of structure to be tested from a single transducer position, so avoiding the time-consuming scanning required by conventional ultrasonic or eddy current methods. As shown in Fig. 2, it essentially converts a point inspection to a line inspection, so that in a plate-like structure it is necessary to scan in one direction, rather than two, to obtain full-volume coverage, and in a one-dimensional structure such as a pipe or rail, the whole volume is covered from a single transducer location. There have also been many studies of the use of a sparse array of omni-directional transducers to monitor an area of a

E-mail address: p.cawley@imperial.ac.uk.

https://doi.org/10.1016/j.ndteint.2023.103026
Received 9 October 2023; Received in revised form 13 December 2023; Accepted 17 December 2023
Available online 22 December 2023

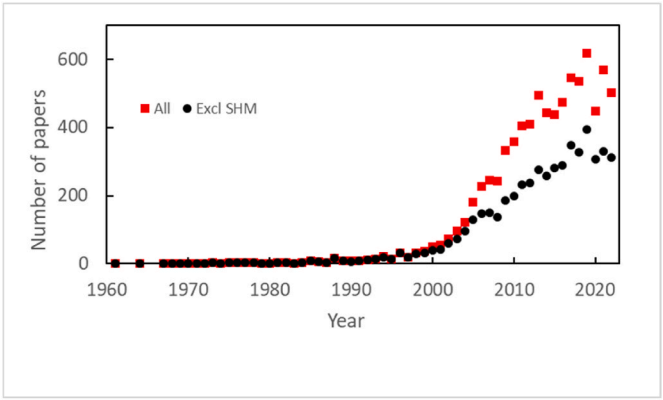

Fig. 1. Number of papers output by a Scopus search on Jan 10, 2023 for 'Guided waves OR Lamb waves' and 'nondestructive testing OR structural health monitoring OR inspection' by year (red squares) and number yielded by a similar search but with 'structural health monitoring' omitted (black circles). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

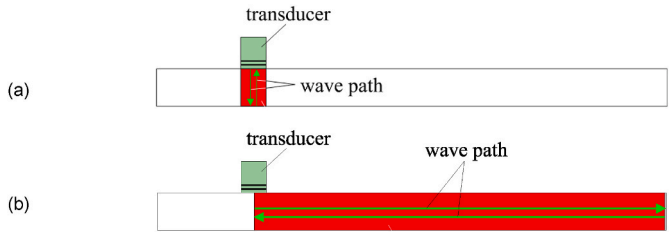

Fig. 2. Contrast between (a) conventional bulk wave ultrasonic testing and (b) long-range guided wave inspection.

plate or shell structure. The technique becomes even more attractive if part of the structure to be tested is inaccessible, for example a pipe passing under a road. The test is usually done in pulse-echo mode, the transducer transmitting the guided wave along the structure, and returning echoes indicating the presence of defects or other structural features. It is also possible to use pitch-catch transduction between a transmitter and a remote receiver; this has mainly been applied in sparse array monitoring of plate and shell structures.

Guided waves can also be used for material property measurement, for example, texture in metals [8,9], the properties of viscous liquids and soft solids [10–12] and the elastic constants of composite materials [13–15]; they are also the main transmission mechanism of the energy that is detected in acoustic emission testing [16] and leak detection [17].

This paper reviews the use of guided waves in active, long range guided wave inspection and structural health monitoring ie defect detection in NDT and SHM applications in which the guided waves are both generated and received by the transduction system. Passive applications such as acoustic emission and leak detection, in which only receiver transducers are employed with the aim of detecting the source of the guided waves, are not covered.

Section 2 reviews the theory of guided waves and discusses the prediction of the properties of the guided waves that can propagate in different geometries. This is followed in section 3 by a discussion of the prediction of the interaction of different guided wave modes with different types of defect and in section 4 by a discussion of guided wave mode selection for different applications. Section 5 then discusses transduction methods to obtain clean guided wave signals, while section 6 introduces the signal processing methods commonly applied to guided wave signals. Sections 7 and 8 then review applications in nondestructive testing and structural health monitoring respectively. Section 9 concludes the paper with a review of the keys to success in the different applications and current research needs.

The paper aims to give an overview of the history and current state of the field and to suggest fruitful areas for future work; it does not attempt to reference all the work on each topic, though there are several references to more detailed reviews of specific areas. Other general reviews of guided wave inspection and monitoring have been presented by, for example, Fromme [18], Abbas and Shafiee [19], Mitra and Gopalakrishnan [20] and Olisa et al. [21], while Giurgiutiu [22] and Ricci et al. [23] have looked specifically at the health monitoring of composite structures.

2. Guided wave theory

2.1. Original work

Lord Rayleigh [24] was the first to derive the solution for a surface wave travelling along the free surface of an elastic half-space, and these waves became known as Rayleigh waves. Stoneley [25] generalized this to show the conditions for waves, now known as Stoneley waves, to propagate at the interface between two half spaces, while Lamb [26] was the first to add a second interface and to solve the problem of waves travelling along an elastic plate of finite thickness, the waves being known as Lamb waves. His derivation was for plates in vacuum and the roots of his two equations, one for symmetric modes and one for antisymmetric modes, yield the well-known Lamb wave dispersion curves.

Cylindrical wave propagation problems were studied slightly earlier than the work on half spaces and plates, Pochhammer [27] and Chree [28] being the first to study waves in a free bar, and their names are still associated with the equation that describes the guided modes of a solid cylinder. Gazis [29] presents a clear derivation of the conditions for the propagation of waves in elastic cylinders, though with some typographical errors noted by Lowe [30].

Rayleigh, Lamb and Stoneley waves all describe waves whose particle motion is in the plane of the wave propagation direction and the normal to the free surface(s). Waves can also exist in which the particle motion is normal to this plane; in the case of a layer on a half space, these are Love waves, named after Love [31], while in plates they are termed shear horizontal (SH) waves and in cylindrical geometries they form the torsional mode family [4].

2.2. Analytical solution for dispersion relationships

Much of the original interest in the area came from the geophysics community since guided modes are important in understanding earthquakes, as well as in mineral prospecting, and the various solution methods developed are reviewed by Lowe [30]. It is possible to predict the dispersion relationships for each new geometry of interest from first principles, starting from the original work of section 2.1 or the many textbooks eg Refs. [4,32–34]. However, this is time consuming as it involves not only deriving the relationships but also solving them, and this is difficult in some cases, particularly when the attenuation is high. Lowe and Pavlakovic have produced the general purpose software, *Disperse* [35,36], based on the original work of [30] and extended to cylindrical geometries in Ref. [37]. It predicts wave propagation in an infinite body of uniform geometry, dealing with bodies comprising.

- an arbitrary number of layers;
- flat or cylindrical geometry;
- in free space, immersed in fluid or embedded in solid;
- undamped or with material damping;
- isotropic or anisotropic layers.

It outputs dispersion curves in various formats including phase velocity, group velocity, wavenumber and attenuation, together with the corresponding mode shapes of displacements, stresses, strains and energy density. The software is based on the global matrix method [30]

which starts with the four (in the case of a flat, isotropic layer) partial waves that can exist in each layer, shear and longitudinal waves propagating up and down, as shown in the three-layer system example of Fig. 3. Stress equilibrium and displacement continuity conditions are applied at the internal layer boundaries, while stress-free conditions are applied at the outer boundaries when the body is in free space. If the body is embedded in an inviscid liquid then there is no incoming wave at the outer boundaries and only longitudinal outgoing waves, while if it is embedded in a solid or a viscous fluid then there can be outgoing longitudinal and shear waves but again, no incoming wave. This analysis yields the characteristic equation of the form

$$f(frequency, phase\ velocity) = 0 \tag{1}$$

if there is no attenuation in the system due to either material attenuation or leakage of energy into surrounding media, and

$$f(frequency, phase\ velocity, attenuation) = 0 \tag{2}$$

if the system is attenuative. The roots of equation (1) or (2) give the conditions under which waves can propagate in the system without the need for continuous excitation; graphs of the roots in (frequency, phase velocity) space give the most commonly plotted dispersion curves, but the axes can be converted to (frequency, wavenumber) or (frequency, group velocity), among many other combinations. In the attenuative case, a three-dimensional plot in (frequency, phase velocity, attenuation) space is possible, but it is often easier to interpret a pair of two dimensional (frequency, phase velocity) and (frequency, attenuation) plots. In general, the curves are functions of frequency, geometry and material properties. However, for a single layer of a given material it can be shown that the same curves are obtained for all thicknesses if plotted as functions of frequency-thickness product.

For example, Fig. 4a and b shows the phase and group velocity dispersion curves respectively for an elastic steel (density, $\rho = 7900$ kg/m^3, longitudinal wave speed, $c_L = 5.96$ km/s, shear wave speed, $c_s = 3.26$ km/s) plate in vacuum using a frequency-thickness abscissa, the bulk longitudinal and shear wave velocities being marked as horizontal lines The fundamental symmetric (S0) and anti-symmetric (A0) modes exist at all frequencies, while the higher order modes have cut-off frequencies at which the phase velocity is infinite (corresponding to infinite wavelength) and the group velocity is zero, indicating that there is no energy propagation along the plate. At high frequency, the velocities of the two fundamental modes asymptote to the Rayleigh wave velocity, while those of the higher order modes asymptote to the bulk shear velocity. Fig. 4c and d shows the displacement of the fundamental modes at low frequency as a deformed grid, the A0 mode being essentially a bending wave and the S0 mode an extensional wave. Fig. 4e shows the S0 mode at a frequency-thickness of 10 MHz-mm, the motion now being largely confined to the surfaces. Fig. 4f shows the in-plane and out-of-plane displacements and direct stresses in the S1 mode at its maximum group velocity point (~4.2 MHz-mm) in line graph form, the coordinate system being shown in Fig. 3. The number of modes that can propagate increases with frequency and so higher frequency signals tend to be more complex. All the Lamb modes have frequency ranges where they are dispersive (their velocity changes with frequency). When the mode is dispersive, the different frequency components in a wave packet travel at different speeds and so the wave packet spreads out in time and its peak decreases. This tends to decrease the signal to noise ratio and reduces resolution (the ability to separate the reflections from closely spaced features). The dispersion is minimised at points of maximum group velocity and these are popular operating points for inspection [5, 38].

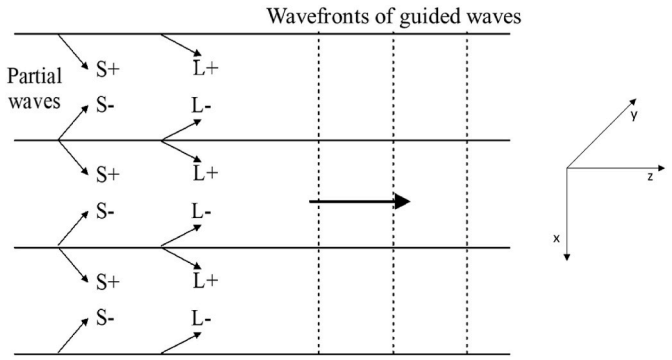

Fig. 3. Partial waves in layers combining to form guided waves in a system with three isotropic layers. L^+/L^-, S^+/S^- are longitudinal waves propagating upwards/downwards and shear waves propagating upwards/downwards in a layer respectively. After Lowe [30].

If, instead of being in vacuum, the plate of Fig. 4 is immersed in water, the system is attenuative and Fig. 5a and b shows the phase velocity and attenuation dispersion curves respectively for a plate thickness of 1 mm. The phase velocity dispersion relationships are very similar to those of Fig. 4a except that the A0 mode splits at the longitudinal wave speed in water, 1.5 km/s. One branch stays below this velocity and asymptotes towards it at high frequencies; this asymptote corresponds to the Scholte mode [39], while at lower frequencies it is termed the quasi-Scholte mode [11]. The mode shape of the quasi-Scholte mode at a frequency of 0.5 MHz is given in Fig. 5c, showing bending motion of the plate and primarily longitudinal motion in the water that decays exponentially away from the plate surface. Its energy is therefore trapped adjacent to the plate surface and does not leak energy away into the bulk of the (assumed inviscid) liquid; its attenuation is therefore zero (its attenuation dispersion curve in Fig. 5b is on the horizontal axis). By contrast, the A0-like mode above the longitudinal velocity in water leaks energy very strongly into the water, as shown at a frequency of 0.5 MHz in Fig. 5d by the waves radiating away from the plate, and the attenuation is very high at ~740 dB/m as shown at point X in Fig. 5b. The attenuation in each mode tends to a minimum at its maximum group velocity point (see Fig. 4b) where the surface motion of the plate is primarily in the in-plane direction and so does not couple to the (inviscid) water.

The particle motion of Lamb waves is in the z-x plane of Fig. 3; there is another family of modes, the shear horizonal modes, in which the particle motion is in the y direction, in the plane of the plate and normal to the propagation in the z direction. In a flat, uniform, isotropic (or anisotropic in a principal plane) plate, the SH modes are uncoupled from the Lamb mode family and Fig. 6 shows the phase velocity dispersion curves for the SH modes in a steel plate. The fundamental, SH0, mode exists at all frequencies and propagates at the shear wave bulk velocity; its particle motion is uniform throughout the plate thickness. The higher order modes have cut-off frequencies increasing with mode order, and at high frequencies their phase velocity asymptotes to the bulk shear wave velocity.

Torsional (T), longitudinal (L) and flexural (F) modes can propagate in the axial direction along cylindrical geometries; the axially symmetric torsional modes are analogous to the SH modes in a plate, while the axially symmetric longitudinal modes are analogous to the Lamb modes in a plate. The flexural modes are not axially symmetric and in general couple axial, radial and circumferential motion. Fig. 7 shows the group velocity dispersion curves for a 6 inch diameter schedule 40 pipe (wall thickness 7 mm). The modes are named X(n,m) where X can be L, T or F, n is the circumferential order and m is a counter. The particle motion is axially symmetric in the $n = 0$ modes, while it varies harmonically with n cycles around the circumference when $n > 0$. Flexural modes with $m = 3$ are in the longitudinal wave family and the phase and group velocities asymptote to the velocity of the L(0,2) mode at high frequency; $m = 2$ corresponds to the torsional mode family whose velocities asymptote to the torsional, T(0,1), mode velocity, which is also the bulk shear velocity. The $m = 1$ family corresponds to bending waves in the pipe wall and the velocities asymptote to the Rayleigh wave velocity at high

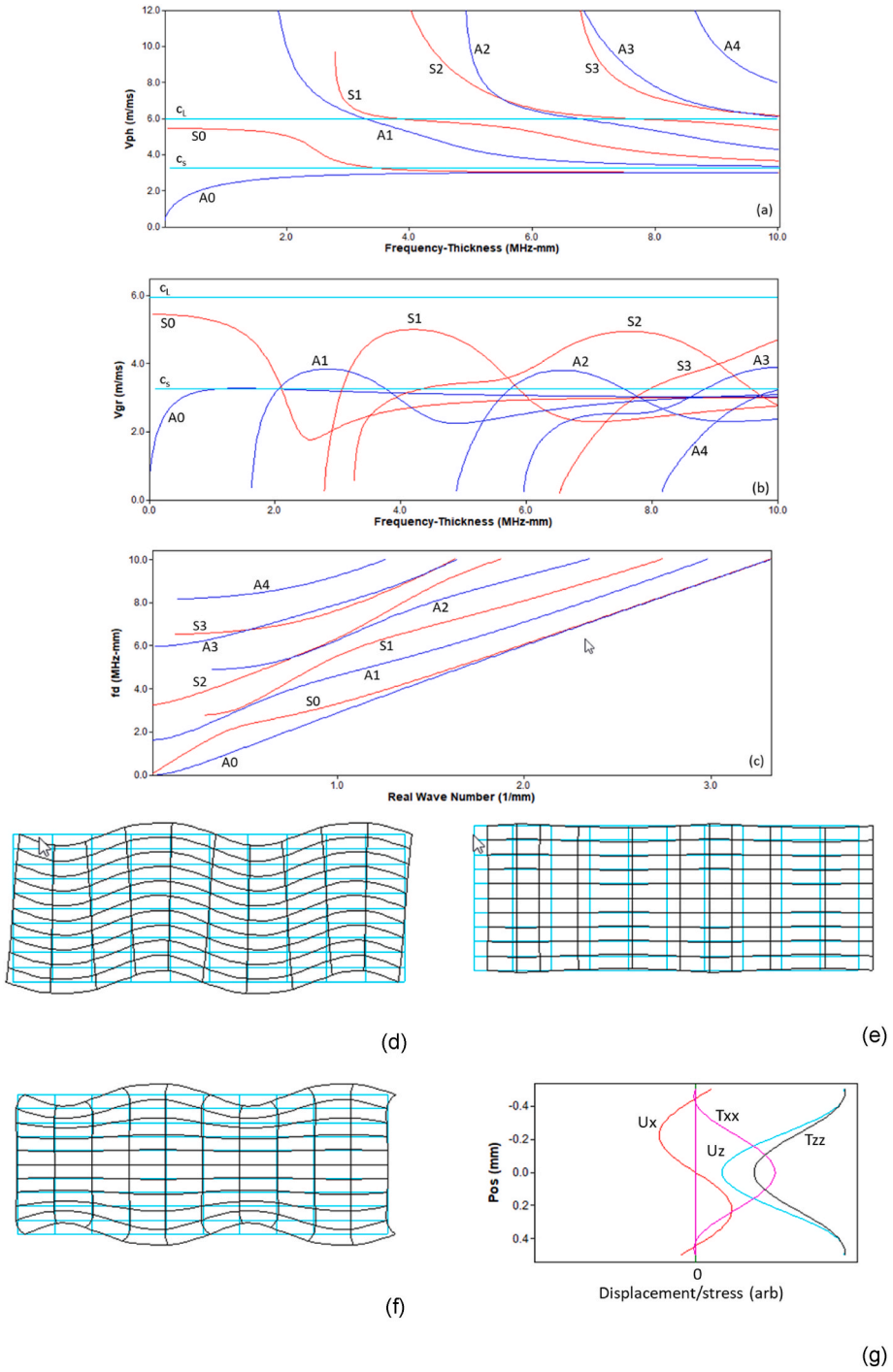

Fig. 4. Dispersion relationships for steel plate (a) phase, (b) group velocity and (c) wavenumber plotted with frequency-thickness axis; corresponding mode shapes (d) A0 at ~0.5 MHz-mm; (e) S0 at ~0.5 MHz-mm; (f) S0 at ~10 MHz-mm; (g) S1 at maximum group velocity (~4.2 MHz-mm) showing in-plane (Uz), out-of-plane (Ux) displacements and in-plane (Tzz) and out-of-plane (Txx) direct stresses. In grid plots (c, d, e), cyan shows un-deformed grid and black shows deformed shape. Line plots of (g) are for 1 mm thick plate with origin at midplane. Modes of (a)–(c) plotted using automatic tracing option in *Disperse* that for numerical reasons does not always plot as far as cut-off eg several modes in (c) not plotted back to zero wavenumber. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

frequency. The T(0,1) and L(0,2) modes are the most attractive for long range inspection as they are non-dispersive over wide frequency ranges.

It is also possible to model embedded cylindrical geometries. For example, Fig. 8 shows the attenuation dispersion curves for the axially symmetric modes of a 21.7 mm diameter steel bar bonded in limestone using epoxy. The attenuation is generally very high but reaches a series of minima at which the particle motion on the bar surface is very small and so minimises coupling to the surrounding rock; the lowest attenuation is reached at around 2 MHz in the L(0,11) mode, above which successive minima have higher attenuation since material attenuation becomes more significant with frequency. This suggests that the inspection range would be maximised by testing at ~2 MHz) rather than at lower frequencies as would usually be expected.

There has been a great deal of interest in the use of guided waves for testing composite materials and this has increased in recent years with the increasing emphasis on SHM. Rose [4] sets out the derivation of the

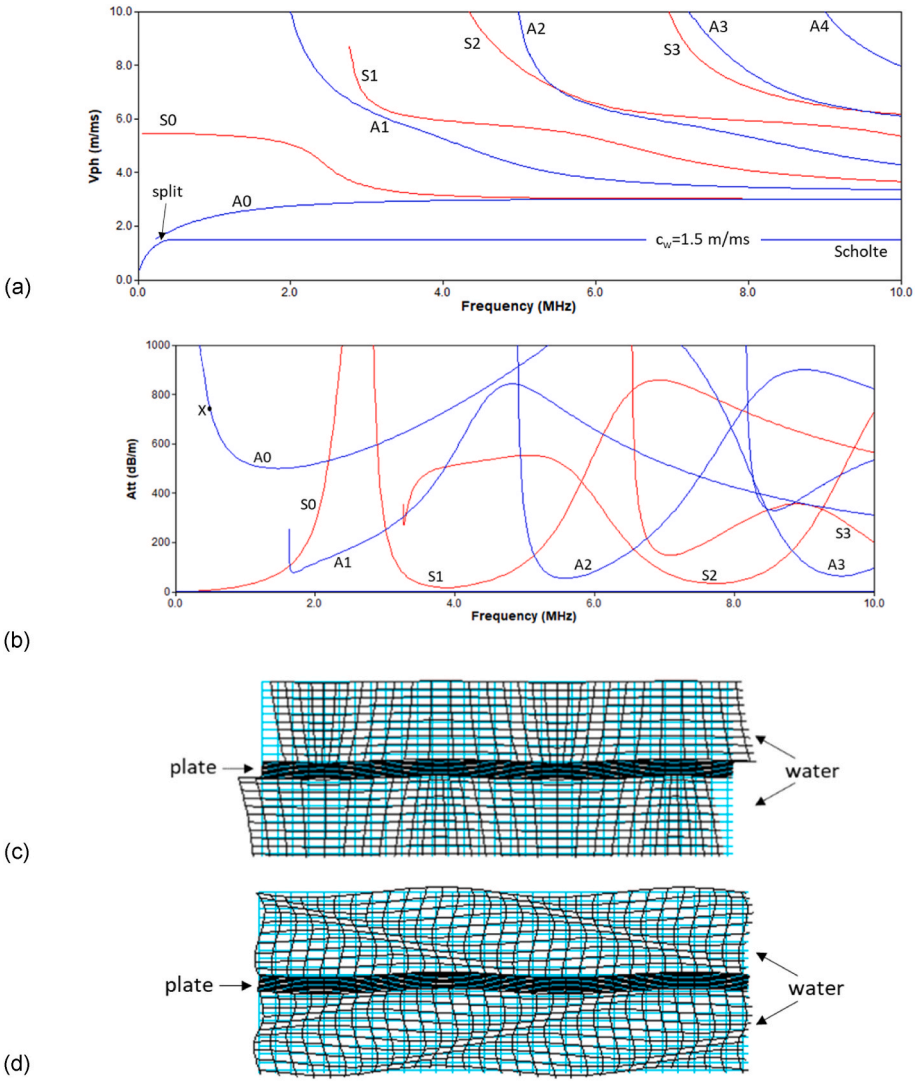

Fig. 5. Dispersion relationships for 1 mm thick steel plate immersed in water (a) phase and (b) attenuation; corresponding mode shapes (c) quasi-Scholte at ~0.5 MHz-mm; (d) A0 at ~0.5 MHz-mm (point X in (b)). In grid plots (c, d), cyan shows un-deformed grid and black shows deformed shape. The grid shows 5 mm of water on either side of the 1 mm thick plate; the water extends to infinity on either side – there is no top or bottom boundary. For numerical reasons, some modes in (b) not plotted where attenuation rising rapidly eg A1 mode at ~1.8 MHz. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

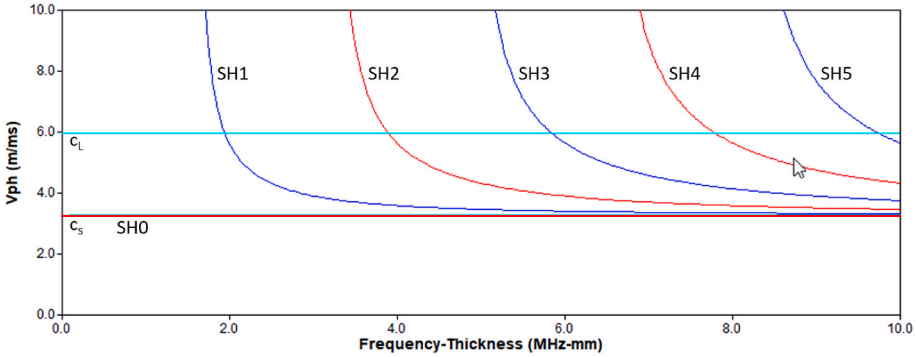

Fig. 6. Phase velocity dispersion curves for shear horizontal modes in steel plate.

dispersion relationships for composites and gives examples of propagation in different directions in both single and multi-layer constructions. The appearance of the dispersion curves is similar to that for isotropic materials and the fundamental modes at low frequencies have similar displacement mode shapes to those in the isotropic case; however, the stress distributions can be very different. For example, Fig. 9 shows the interlaminar shear stress distribution across the plate thickness in the S0 mode at ~0.5 MHz for a 1 mm thick, 8 layer, cross-ply [(0,

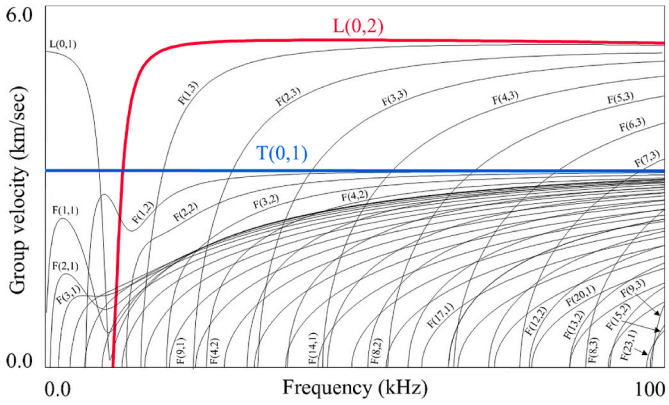

Fig. 7. Group velocity dispersion curves for 6 inch diameter, schedule 40 steel pipe (wall thickness 7 mm). The L(0,2) and T(0,1) modes are the modes most used in pipe inspection and monitoring.

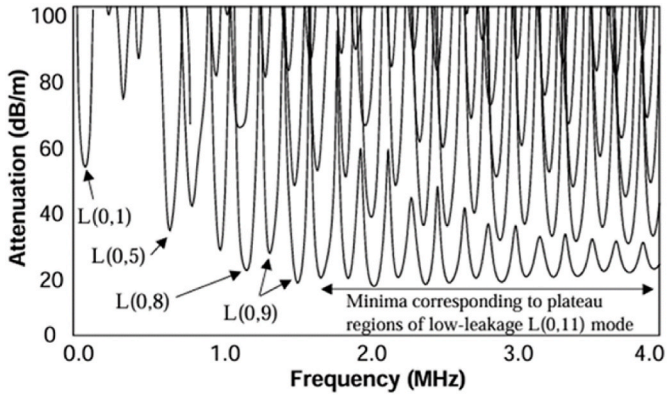

Fig. 8. Attenuation dispersion curves for the axially symmetric modes in a 21.7 mm diameter steel bar, surrounded by a 4 mm layer of epoxy and embedded in limestone, showing the low-leakage modes [46].

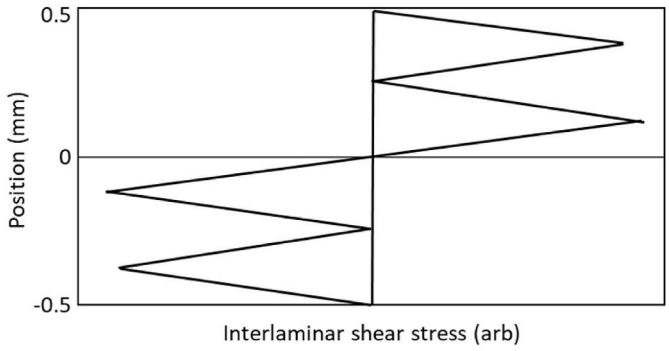

Fig. 9. Shear (σ_{zx}) stress distribution in S0 mode at 0.5 MHz in an 8 layer, cross-ply $[(0,90)_2]_s$ CRRP laminate (After [40]).

$90)_2]_s$ CRRP plate. There are clear discontinuities at the ply boundaries and the interlaminar shear stress is zero at 25 %, 50 % and 75 % of the way through the thickness (between plies 2&3, 4&5 and 6&7); this has significant implications for delamination detectability [40]. An important difference between guided wave propagation in isotropic and anisotropic materials is that when the propagation is not in a principal plane of an anisotropic material, the phase and group (energy) velocities are not necessarily in the same direction, the angle between their two directions being termed the skew angle [4]. This has important implications for testing and monitoring since when, for example, a piezoelectric transducer on a wedge (see section 5.2) is used, the transducer will be aligned with the phase velocity direction, but the energy will propagate in a different direction; therefore, the region of highest defect sensitivity will not be aligned with the transducer. Many authors have studied propagation in composite materials, for example the book by Datta and Shah [41]; in recent years, research has been concerned with issues such as energy propagation and skew [42,43] and the prediction of attenuation [44,45].

2.3. Semi analytical finite element method

Analytical solutions for the dispersion relationships are available for simple plates, bars and tubes/pipes, but not for complex geometries such as rail or T-shaped beams. Here a numerical approach, such as the finite element method, is needed. However, since it is known that the waveguide is uniform in the propagation direction and that the waves of interest are periodic in this direction, it is only necessary to mesh the cross section of the waveguide. The method used is known as the Semi-Analytical Finite Element (SAFE) method, which is also called the spectral element, strip element, or waveguide finite element method [47].

The SAFE method was first demonstrated in 1972 to produce the dispersion solutions of solid waveguides of arbitrary geometries [48]. Later, Damljanovic and Weaver [49] extended the idea to calculate both propagating modes and non-propagating, evanescent modes (complex wavenumbers) for anisotropic cylinders. The SAFE method confined to obtaining the propagating solutions has been applied to thin-walled waveguides [50], rails [51,52], wedges [53], nonhomogeneous anisotropic beams [54] and rods [52]. Bartoli et al. [55] extended the method to allow for viscoelastic material damping, so that the solutions are necessarily complex, the damping of the guided waves being represented by the imaginary part of the wavenumber.

A drawback of these models is that they all need to be developed in specific FE codes. Wilcox et al. [56] have implemented an approximation of the SAFE method in a standard finite element package by imposing a cyclic axial symmetry condition. It works by defining an axially-symmetric model with a very large radius compared to the dimensions of the cross section. For a specific cyclic order, and using axisymmetric harmonic elements, the finite element eigensolver generates a chosen number of eigenfrequencies and eigenvectors, i.e. the vibration solutions of standing waves in the ring. The cyclic order of the standing wave corresponds to the number of wavelengths of a guided wave mode around the axisymmetric body, with the eigenvector being its displacement distribution or mode shape at the corresponding eigen frequency. However, this model can only describe a waveguide with stress-free exterior boundaries.

Predoi et al. [57] have implemented the SAFE method in a commercial finite element package (COMSOL [58]) by reorganizing the SAFE equations into a standard Finite Element eigenvalue formalism which can be solved by commercial software. They also introduced periodic boundary conditions in the SAFE method, which allow the modelling of infinitely wide waveguides with periodic changes in geometry or material properties along the width. Castaings and Lowe [59] further developed this model to address the leaky wave problem in which guided waves propagate along an elastic waveguide with arbitrary cross-section and radiate into a solid of infinite extent. This has been achieved by using an absorbing region to absorb the leaking waves, thus simulating an infinite extent of the solid medium. More recently the method has been used for predicting the dispersion properties of bone [60], finned tubing [61], polygonal drill pipes [62] and helical multi-strand wires [63,64].

3. Interaction with defects

There is a vast literature on the scattering of elastic waves from different types of defect, much of it focused on the bulk longitudinal or shear waves used in conventional ultrasonic inspection. The

fundamental S0 (at low frequencies) and SH0 modes have essentially uniform displacements and major stress components through the plate thickness and so are analogous to bulk longitudinal and shear horizontal waves respectively; the higher order modes have more complex displacement and stress distributions so their interaction with defects is more complex. Mathematical modelling of scattering is challenging, and exact analytical solutions only exist for certain, geometrically simple, shapes of scatterer such as spherical and cylindrical cavities. Approximate solutions also exist that are valid in certain regimes of wavelength to scatterer size ratio. For example, if a scatterer is large and its shape is slowly varying compared to the wavelength, the Kirchhoff approximation is valid, and can be applied to scattering problems [65]; however, this approximation breaks down when the scatterer dimensions are of the order of the wavelength. Examples of analytical and semi-analytical approaches to specific guided wave scattering problems include the effect of bond quality on Lamb wave scattering in adhesive joints [66], Lamb wave scattering from both through-thickness and part-thickness circular holes in plates [67] and interaction with cracks in a waveguide [68]. Achenbach has written extensively on scattering problems, and in particular the use of reciprocity analysis [69], building on the work of Auld [70]. Shen and Giurgiutiu [71] have presented software based on simple analytical scattering models for predicting the waveforms generated by piezoelectric wafer transducers interacting with damage.

Unfortunately, there are many scatterers of practical interest with more complex geometry for which analytical and semi-analytical models are not applicable, and it is also sometimes necessary to consider irregular waveguides. In these cases numerical techniques such as the finite element (FE), finite difference [33,72] or boundary element [73,74] methods must be employed. The discretisation employed with finite element and finite difference methods is very similar but early work by Blake [75] showed that the finite element model is more stable, particularly when dealing with sloping boundaries. This led Alleyne [76,77] to follow other authors [78–80] and implement the finite element method. In the 1990s it was a challenge to model scattering from a simple notch in a plate in two dimensions, but the ever-increasing power and memory available in desktop computers has gradually made it possible to deal with scattering from complex defects in three dimensions, and to model complex waveguides. This can be done with general purpose packages such as Abaqus [81] or COMSOL [82] but can more efficiently be done with packages specifically designed for wave propagation and scattering problems such as POGO [83,84]. While the boundary element method uses Green's theorem to reduce the dimensionality of the problem by one, so requiring the meshing of the boundary of an object rather than its volume, the generality and ease of use of the finite element method means that it dominates current guided wave scattering computation.

If the reflection from a defect in a large structure remote from boundaries is to be modelled, the simplest approach is to model a region of structure around the defect that is large enough that the waves reflected from and transmitted through the defect can be separated from the reflections obtained from the boundaries of the model, and this is satisfactory in many cases in plates and pipes [77,85–87], including some 3D cases, even in composite materials [88]. However, the burden of modelling a large region around the defect in all directions becomes more severe when three dimensional models are used and it can then be beneficial to consider absorbing boundaries that suppress reflections using either ALID (absorbing layer using increasing damping) [89] or PML (perfectly matched layer) [90] schemes. Rajagopal et al. [91] compared the two approaches and concluded that the PML method generally requires a smaller total model size than ALID. However, this is balanced by ALID being easier to implement in standard finite element packages since PML requires specific programming which is only possible in a minority of packages, and also requires a different implementation for anisotropic materials. A further development of the ALID concept, involving varying the stiffness as well as the damping, has been shown to give further improvements [92].

The large models required to predict guided wave scattering remain a substantial challenge and many authors have used a hybrid formulation in which the region immediately surrounding the defect is modelled using finite elements, with the wave field at the boundaries of the region being decomposed into the modes of the uniform structure beyond the defect region. These modes can then be propagated analytically to any desired measurement position, and it is also possible to model the waves generated by a remote transducer via the reverse process. This approach was used in very early work [80] and has recently been employed by many authors (e.g. Refs. [93–95]) and has been extended to three dimensional models [96] and finite difference formulations [97]. In essentially one-dimensional waveguides such as rods or long pipes, it is usually the reflection and transmission of different modes in the forwards and backwards direction along the waveguide that is of interest, and this is relatively straightforward. There has recently been a great deal of interest in the sparse array monitoring of plate and shell structures and here transmission and scattering in all directions is of interest. Velichko and Wilcox [98] have presented a methodology for the efficient computation of the full scattering matrix from a defect in two or three dimensions using a commercial finite element code, and have applied it to both bulk and guided wave scattering.

The advances in finite element modelling capabilities have enabled, for example, 3D analyses of guided wave interaction with defects in rail [99], defects at different locations around bends in pipes [100], complex shaped corrosion patches in pipes [86] and defects between a stiffener and a composite panel [101]. They have also made it practical to model guided wave tomography configurations in which the signals from multiple send-receive transducer pairs must be predicted [102].

While finite element models are increasingly able to predict the scattering characteristics of guided waves from a variety of defects, it is often useful to combine the numerical results with the predictions of simple analytical models to gain physical insight into the relationship between the defect size and shape and the scattering amplitude and inspection frequency [103,104].

4. Mode and frequency selection

4.1. Signal requirements

The main difficulty with long range guided wave inspection is that it is very easy to obtain signals like that shown in Fig. 10a. This shows the pulse-echo signal produced on a length of uniform pipe by a group of transducers covering a quarter of the pipe circumference and connected in parallel so that they effectively act as a single transducer. Ideally the signal should contain two distinct echoes from the two ends of the pipe, rather than the very complicated trace seen in the figure. The excitation signal was a 5 cycle, 70 kHz windowed toneburst and Fig. 7 shows that over 30 modes can propagate at this frequency. The signal complication arises because many of these modes are excited and they travel at different velocities in both directions along the pipe; the velocities are in general a function of frequency (i.e. the modes are dispersive) as shown in Fig. 7, so the signals elongate rather than remaining simple windowed tonebursts.

Fig. 10b shows a signal obtained at an early stage in the development of the pipe screening system discussed later. Here, a full ring of transducers was used, rather than the quarter ring of Fig. 10a, so that the symmetric modes would be excited preferentially. Also, the transducers were designed to excite the pipe in the axial direction, so preferentially exciting the longitudinal modes rather than the torsional mode. Reflections of the L(0,2) mode (see Fig. 7) from two welds approximately 15 m apart in a long pipe can clearly be distinguished. However, there are many smaller signals between the two weld echoes which should not be present since this was a new pipe. Averaging did not improve the signal, indicating that the problem is coherent, rather than random, noise. The first weld reflection amplitude was estimated to be about

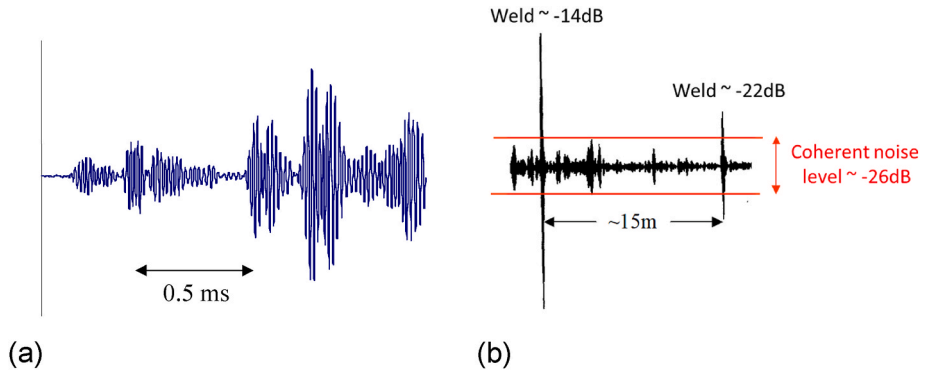

Fig. 10. (a) Signal received on length of plain pipe using transducers over quarter of circumference; (b) signal received on welded pipe in early site test [5].

−14dB, while the second weld reflection was about −22dB, the reduction being partly due to different weld geometry and partly to attenuation along the pipe. The coherent noise level was around −26dB which was the same as the target defect reflection size, so to be useable, the system needed further refinement to reduce the noise. The coherent noise has two main sources:

- the excitation and reception of unwanted modes;
- the transmission of waves in the opposite direction along the pipe and the reception of echoes from that direction.

The key to controlling coherent noise is therefore to excite and receive a single mode in one direction. It was found that the main reason for the excitation of multiple modes was that the transducers around the pipe were not sufficiently well matched, so when excited together they did not provide symmetric excitation and many unwanted flexural (F) modes (see Fig. 7) were excited; subsequent work improved the consistency of the transducers and this led to a commercial system, as discussed in section 5.4.

In addition to controlling coherent noise, it is also necessary to control dispersion. If the chosen mode is dispersive, the different frequency components in the signal travel at different velocities so the signal duration increases which compromises the spatial resolution (the ability to distinguish echoes from closely spaced reflectors). Dispersion is not very evident in Fig. 10b since it was controlled by applying narrow band excitation centred on a region where the mode of interest is non-dispersive. This strategy to overcome dispersion problems is often sufficient, though dispersion compensation [105] can also be valuable.

4.2. Mode selection

Having established that it is desirable to operate with a single mode, it is then necessary to select the mode and the operating frequency. The selection principles have been discussed by Alleyne and Cawley [38], Wilcox et al. [106] and Hay and Rose [107]. The first requirement is that the mode be sensitive to the required defect(s) type and location. For example, the Rayleigh wave is very sensitive to surface breaking cracks normal to the direction of propagation, but is insensitive to defects deeper than around one wavelength from the surface. In contrast, the S0 mode at low frequency resembles an extensional wave and is sensitive to cracks normal to the wave propagation direction at any point through the plate thickness, as is the SH0 mode. However, the S0 mode is practically insensitive to cracks parallel to the propagation direction.

The S-Parameter approach of Auld [70], based on the principle of reciprocity, is very useful for assessing the likely sensitivity of candidate modes. It implies that scattering will be strong if the stresses of the incident wave are such that they impose a large discontinuity in displacement at the defect. Thus, for example, a wave whose stresses are such that they cause a crack to open and close will tend to be scattered strongly [108]. A first assessment of the likely sensitivity of a mode to a particular type of defect is therefore to study the stress distribution through the thickness of the structure, as shown in Fig. 4f, and to assess whether there is significant stress across the plane in which defects must be detected. For example, Fig. 9 shows that in an 8-layer, cross plied carbon fibre composite plate there would be significant shear stress across a delamination between plies 1&2 and 3&4 (12.5 % and 37.5 % of the way through the thickness), but zero shear stress at 25 % and 50 % of the way through the thickness (between plies 2&3 and 4&5). By symmetry, there is also zero shear stress between plies 6&7 and a maximum between plies 5&6 and 7&8. Therefore, the mode would not be able to detect a delamination between plies 2&3, 4&5 or 6&7, as was demonstrated in Ref. [40].

The stress mode shapes give a qualitative indication of sensitivity to defects; in most cases it is then necessary to do a quantitative assessment of promising modes via finite element analysis, as discussed in section 3. In some cases, this may need to include joints or regions of changing geometry, for example pipe bends [100].

Unless the mode shape changes significantly, the sensitivity to defects in a given non-dispersive mode tends to improve with frequency as the wavelength decreases [103]. However, attenuation tends to increase with frequency, so limiting the test range. Also, more modes exist at higher frequencies which makes single mode generation more difficult and so tends to increase the coherent noise level. Since it is the signal to coherent noise ratio that usually determines the practical sensitivity to defects, it is often found that it is better to operate at lower frequencies, though this does limit the resolution (ability to distinguish two closely spaced reflectors).

A further key consideration in mode selection is ease of transduction and the extent to which the chosen mode is separated from other modes in the dispersion curves, as this determines the difficulty of exciting it while minimising the generation and reception of other modes. These issues are considered in the next section.

5. Transduction

5.1. Excitation signal

Single mode excitation requires the control of at least two parameters since at least two modes are possible at all frequencies. The most common choice of parameters is the frequency and wavelength of the excitation, the frequency being controlled by the input excitation signal, and the wavelength via the geometry of the transducer; it is also possible to use the frequency and the distribution of forces over the structure, as discussed in section 5.4. Depending on the transducer type, the wavelength at the excitation frequency is fixed via control of the phase velocity or wavenumber, as discussed below. An excitation signal with a controlled centre frequency and limited bandwidth is generally required and this is achieved by using a toneburst with a limited number of cycles

at the desired centre frequency; the amplitude of the sidelobes in the spectrum is usually limited by windowing the excitation signal via a Hanning or similar window [38,106]. The bandwidth can be reduced by increasing the number of cycles in the toneburst at a cost of increasing the length of the signal, and hence reducing the spatial resolution; excitation signals with between 3 and 10 cycles are commonly used.

5.2. Piezoelectric transducer on wedge

Guided waves are often generated by piezoelectric probes where the vibration of the piezoelectric crystal is transmitted to an angled wedge that is connected directly to the surface of the structure via a couplant. This provides strong signal amplitudes and can be implemented with basic equipment, but requires direct contact so inconsistent coupling over the footprint, particularly over curved or rough surfaces, is a problem; there are also issues with poor high temperature performance [109] and reverberations within the wedge. This method is therefore more often applied in the laboratory for development work, see e.g. Ref. [110], than in field testing, though it is sometimes used in an industrial setting [111].

Fig. 11a shows a schematic of a piezoelectric transducer mounted on an angled wedge. The wedge angle (θ) fixes the wavelength in the plate (λ) relative to the wavelength in the wedge (λ_w), as shown in the inset to Fig. 11a. Thus,

$$\lambda = \frac{\lambda_w}{sin(\theta)} \tag{3}$$

and since the phase velocity in a medium is given by $c_p = f\lambda$, the phase velocity in the plate (c) is given by

$$c = \frac{c_w}{sin(\theta)} \tag{4}$$

where c_w is the speed of longitudinal waves in the wedge. Therefore the wedge design gives a constant phase velocity, independent of frequency, and so provides excitation centred on a horizontal line in the phase velocity dispersion curves such as those of Fig. 4a, the frequency range covered being dependent on the excitation. The wavenumber within the wave guide (k) is

$$k = \frac{\omega}{c} \tag{5}$$

so from (4) and (5),

$$k = \frac{\sin(\theta)}{c_w}\omega \tag{6}$$

Therefore in wavenumber dispersion curves such as those of Fig. 4c, the excitation is centred on a straight line through the origin of slope sin $(\theta)/c_w$. As the frequency content (ω) is changed, the wavelength within the wedge (λ_w), and hence the induced wavelength on structure surface (λ), changes; this results in a change in the number of cycles in space over the excitation length (L/λ), thus affecting the wavenumber (k) bandwidth.

Fig. 12a shows the excitation of the A1 Lamb mode in a plate using a piezoelectric transducer on a wedge. The excitation zone is roughly parallel to the dispersion curve, making it relatively easy to excite a pure mode over a wide frequency bandwidth.

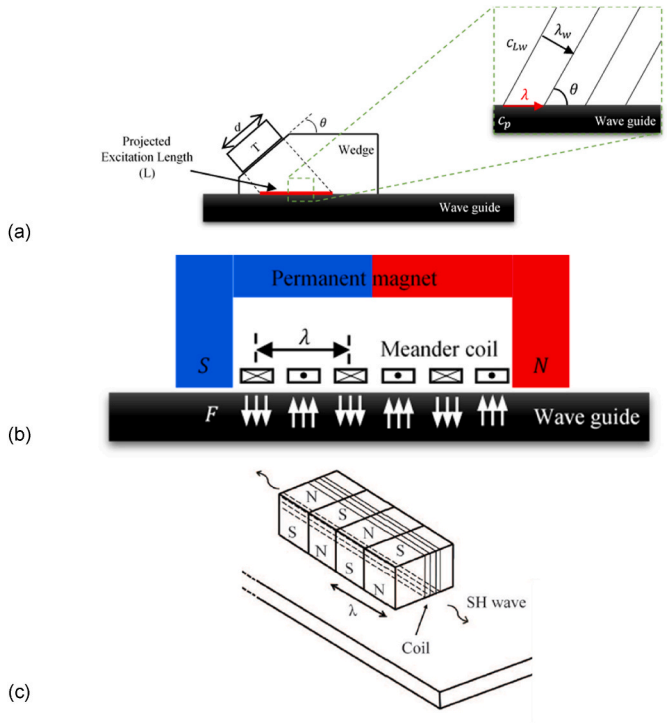

Fig. 11. Schematic diagram of excitation of guided waves by (a) piezoelectric transducer on wedge generating Lamb waves; (b) meander coil EMAT generating Lamb waves; (c) PPM EMAT generating SH waves. ((a, b) after [112]; (c) after [113]).

5.3. EMATs

Electromagnetic acoustic transducers (EMATs) rely on the interaction of the dynamic magnetic field of the eddy currents generated by a coil and the bias magnetic field of a permanent (or electro-) magnet to produce a force resulting in the generation of elastic waves in the structure [114,115]. Hence, they do not require direct contact which eliminates the need for a couplant; this allows transduction over coatings/paint and rough surfaces, and also good high temperature/long term performance [116,117]; however, they suffer from poor signal-to-noise ratio (SNR) and can require bulky and expensive electronics.

A typical configuration of a meander coil EMAT is shown in Fig. 11b [114,115]; this generates vertical forces on the testpiece and so excites Lamb waves. An alternative configuration with periodic permanent magnets (PPM) and a coil carrying current in the wave propagation direction shown in Fig. 11c generates horizontal forces normal to the propagation direction and so can be used to generate SH waves. As with the piezoelectric probe, the frequency bandwidth with EMATs is determined by the nature of the excitation signal in time; in contrast, however, the wavenumber bandwidth is dictated by the length of the transducer in wavelengths, the transducers of Fig.. 11b and c being three and two wavelengths long respectively. Therefore, due to the nature of the design, the frequency and wavenumber bandwidths are fully decoupled and can be chosen independently of each other.

For ideal modal selectivity in the case of a meander coil EMAT, for a given frequency, the wavelength of the coil (λ) is chosen according to the wavelength of the mode of interest, while the frequency and wavenumber bandwidths must be carefully calculated in order to minimize the excitation of the unwanted guided modes. For example, Fig.. 12b and c shows the excitation region for a 30 wavelength long meander coil EMAT on a steel plate; the wavelength is 1.8 mm so the axial length of the coil is 54 mm. In this case, the wavenumber bandwidth of the excitation is much narrower than that of the piezoelectric device of Fig. 12a, but if the same 5-cycle windowed excitation signal is used as with the piezoelectric probe, multiple modes are excited since the excitation region crosses the dispersion curves, rather than running parallel to them, as shown in Fig. 12b. It is possible [112] to excite a pure A1 mode by increasing the number of cycles in the excitation signal to 50, so greatly decreasing the frequency bandwidth, as shown in Fig. 12c; this is achieved at a cost of a longer signal in time and hence poorer spatial resolution.

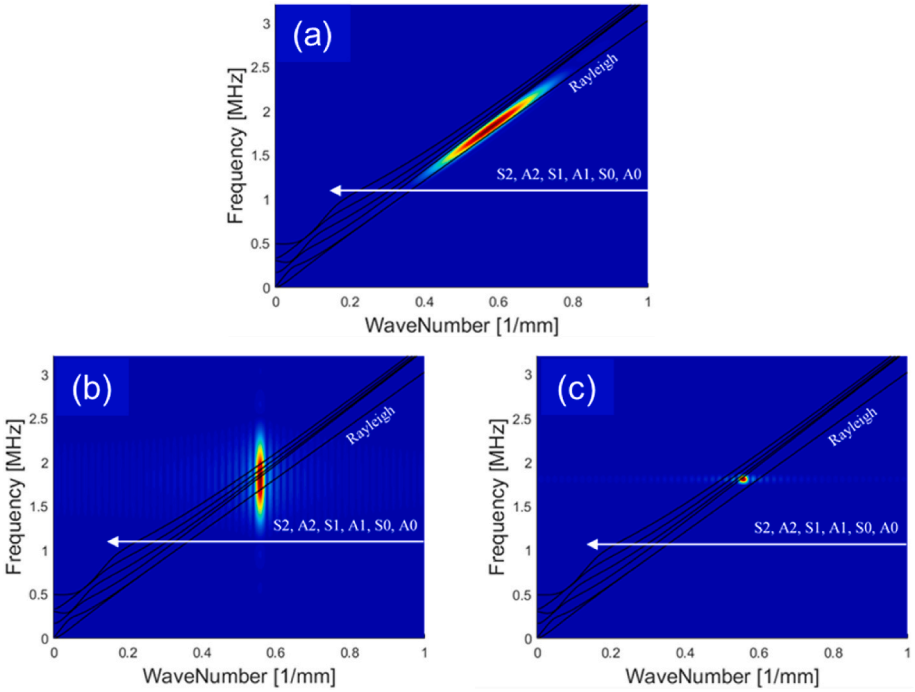

Fig. 12. Wavenumber dispersion curves for Lamb modes in 10 mm thick steel plate showing 2-D FFT colour plot [blue (low)–red (high) linear scale] of excitation zone for (a) 25.4 mm (1 inch) compression wave piezoelectric transducer mounted on a 60° PMMA wedge with a 5-cycle Hanning windowed toneburst signal in time at a centre frequency of 1.8 MHz; (b, c) 30-cycle meander coil EMAT with a 1.8 mm wavelength in space and a 1.8 MHz Hanning windowed toneburst excitation signal in time with (b) 5 cycles; (c) 50 cycles. (adapted from Ref. [112]). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

The generated forces, *F*, shown in Fig. 11b are due to the Lorentz mechanism [114], as are the forces in the PPM configuration of Fig. 11c, and the Lorentz force is the only generation mechanism in electrically conductive but non-magnetic materials. In magnetic materials, the magnetostriction mechanism is also present and, depending on the magnetostriction constants in different directions, this may reinforce the Lorentz force or may produce a dominant force in another direction [114]. On highly magnetostrictive materials such as nickel, the magnetostrictive force dominates and very strong excitation can be obtained. This has led to designs in which a thin strip of magnetostrictive material is bonded to the test structure in order to transmit the excitation; these have been used for pipe inspection using torsional waves [118,119] and Vinogradov et al. [120] have developed a magnetostrictive transducer for elevated temperature monitoring applications. Lorentz, magnetostriction in steel alone and magnetostriction via bonded nickel strip designs for SH wave generation in a steel plate have been compared by Ribichini et al. [113] and different possible transduction methods for SH and torsional wave inspection and monitoring have been reviewed by Miao and Li [121].

5.4. Array of point-like sources

In order to obtain satisfactory mode control, the transducer generally has to be around 3–5 wavelengths long. For a mode with a phase velocity of 3 mm/μs, the wavelength is 6 mm at a frequency of 500 kHz so the required transducer size is modest. However, if the frequency is reduced to 50 kHz, the wavelength increases to 60 mm and the required transducer size becomes impractical, particularly with piezoelectric devices that require coupling to the test structure. It is also difficult to use a single piezoelectric transducer to apply symmetric excitation around a pipe, unlike EMAT or magnetostrictive devices [118,119].

It has been found that an array of essentially point piezoelectric sources is very attractive in several applications. For example, the Guided Ultrasonics Ltd Wavemaker Pipe Screening System transducer array for an 8 inch pipe is shown in Fig. 13. The array comprises two rings of dry-coupled, piezoelectric transducers [122] which apply a tangential force to the pipe surface at the point of contact, so exciting the torsional mode in both directions along the pipe if all the transducers in one row are excited with the same signal. Provided the excitation frequency is below the SH1 (also T(0,2)) mode cutoff then the tangential forcing direction and the axi-symmetric excitation mean that only the T

Fig. 13. Solid transducer assembly for 8-inch pipe showing array of dry coupled piezoelectric transducers (Guided Ultrasonics Ltd [136]).

(0,1) mode out of all the modes of Fig. 7 is excited. The second ring of transducers is used for direction control, as discussed in section 5.6.

Sparse arrays of piezoelectric discs (sometimes termed piezoelectric wafer active sensors (PWAS)) bonded to a structure have been commonly proposed for SHM applications (see, for example [109, 123–129] and many others). If the disc is circular, an axially symmetric field is obtained, and this is often advantageous in SHM. A thickness-polarised disc will apply a combination of normal and radial forces to the structure, the ratio of the two depending on the disc dimensions and the stiffness of the bondline. Provided the diameter of the disc is significantly smaller than the wavelength, the excitation can often be modelled as a point source. When operated below the A1 cutoff frequency, only the A0 and S0 modes will be excited and it is possible to design the system to excite a single mode and to enhance the low frequency response [130]. Compact arrays of bonded discs giving the ability to monitor a plate-like structure from a single transducer location have also been proposed [131]. Point-like EMAT sources exciting Lamb waves [132,133] and SH waves [134] have also been developed, as have magnetostrictive sources [135].

5.5. Receiver only transducers

In applications where a single exciter is used with multiple receivers, it can be attractive to use passive receivers such as Fiber Bragg Gratings [137,138] or laser vibrometers [139,140]. Multiple gratings can be incorporated on a single fiber, hence reducing the wiring complexity for multi-point measurements. In some circumstances it may be possible to use ambient excitation and Fiber Bragg receivers, so removing the need for excitation transducers [141]. Also, a vibrometer can be mounted remote from the structure, so reducing the added mass on the structure and installation difficulties. It is also possible to use laser generation, so making the whole system remote, though this requires higher power lasers [142].

5.6. Direction control

The method used for direction control depends on the transduction system chosen and the structure being tested. The wedge system of Fig. 11a automatically controls the direction of excitation and reception, guided waves propagating to the right in this example and echoes being received from this direction. The EMAT systems of Fig.. 11b and c excite guided waves in both directions along the structure. In the case of Fig. 11b, direction control is usually achieved by introducing a second coil displaced a quarter wavelength from the first in the propagation direction. If the second coil is excited with the same signal as the first but phase shifted by 90° then it may readily be shown that the signal reinforces in one direction and cancels in the other. With the system of Fig. 11c, it is necessary to introduce a second transducer displaced axially from the first; rather than having the two transducers axially aligned so that the overall length is substantially increased, it is often more convenient to displace the second transducer from the first in the direction normal to propagation, so they are side-by-side with typically quarter wavelength axial separation [143].

In the pipe inspection system of Fig. 13, the two rings of transducers are positioned roughly a quarter wavelength apart along the pipe (the precise fraction of the wavelength depends on the test frequency used) and this enables direction control by phasing the signals to the two rings in a similar way to that used with EMATs [144,145]. If the transducers are addressed individually or in groups, the excitation signal can be applied to each transducer/group with different delays, and so can focus the excited guided waves at a particular location [146]; in reception, the different received signals can be processed to determine whether the propagating mode is symmetric [144] and to synthetically focus the received signal [147,148].

If the EMAT or wedge systems are used on plate-like structures, there will be beam spreading so the wave propagation will not be in a single direction. The extent of spread is a function of the ratio of the width of the transducer in the direction normal to propagation to the wavelength. Likewise, the angular resolution achieved by compact arrays [131] is a function of the diameter of the array to the wavelength.

6. Signal processing

Since guided wave signals often appear to be noisy, there have been extensive efforts to improve the signal to noise ratio. Simple coherent averaging and bandpass filtering deals with random noise, and this is sometimes the major noise source, particularly with EMAT transduction. However, simple averaging and filtering to the bandwidth of the excitation signal does not help with coherent noise such as that seen in the signals of Fig. 10, and this is the most common source of noise in guided wave testing. The best way of dealing with this is to solve the problem at source ie to improve the transduction system to increase mode purity and improve the cancellation of signals coming from unintended paths; the signals of Fig. 10 were improved to ~40 dB signal to noise ratio with the transduction system of Fig. 13. It is very unlikely that this improvement would have been obtained simply by signal processing, and signal processing is best used after the transduction has been optimised so it is seen as a last resort rather than a first resort. If it is necessary to operate in a dispersive regime then dispersion compensation can be applied [105]. However, it is only applicable to signals with a single, dominant mode; the processing will tend to spread the signals in other modes over a wider time interval.

If multiple modes are present, then they can be separated by the two-dimensional Fourier transform (2D FFT) method [149] that was used to produce the excitation maps of Fig. 12 from theoretical data. However, this method requires an array of reception points, so it is much less attractive to apply in practical testing. When modes have significantly different group velocities, they can sometimes be separated by simple time-gating and, if this is not possible, then improvement can sometimes be obtained via a wavelet transform [150] or the reassigned spectrogram [151]. A very useful review of the different methods used to improve the quality of guided wave signals and their pros and cons is given by Diogo et al. [152]. There has been some interest in machine learning for NDT applications, see e.g. Ref. [153], a review of the field being given by Cantero-Chinchilla et al. [154]. Machine learning is particularly attractive in SHM and there is a rapidly increasing amount of work in this area, as discussed in Section 8.2.2.

7. Applications in nondestructive testing

7.1. Pre 1990 work on simple structures

In early industrial work, the quick inspection of a variety of strips and plates was targeted by Lehfeldt and Höller in Germany [155], Ball and Shewring in UK [156] and Mansfield [157] and Alers [158] in USA; the German work on plate-like structures continued later with applications to, for example, laser welded sheets [159]. Also, Thompson et al. [160] in USA and Silk and Bainton in UK [161] developed guided wave systems to locate defects in boiler and heat exchanger piping.

Thompson, Alers and a team at Rockwell in USA working for the American Gas Association [158,162,163] and a team at the Harwell Laboratory in UK working for British Gas [33] developed systems that would be mounted on a PIG [33,164] for the detection of longitudinal cracks in gas pipelines. These employed waves propagating round the circumference of the pipe excited by EMATs in the US system and piezoelectric transducers coupled via rubber wheel probes in the UK design. The US technology used the A0 Lamb mode, while the UK system chose multiple bouncing bulk shear modes in preference to Lamb modes. However, the Harwell researchers noted that the SH guided modes would be an attractive alternative to their system but could not be generated with their preferred transduction. Later work at IZFP in Germany led to a successful system using SH modes generated by EMATs

[159]. These are now in routine commercial use by, for example, PII Pipeline Solutions, a history of their developments being provided in Ref. [165]; similar technology is available from Rosen [166] and other companies.

Early work also led to the development of higher frequency systems using Rayleigh waves for the inspection of artillery shells [163] and railway wheels [167]. In the 1980s, Salzburger et al. also produced a Rayleigh wave system for the inspection of the tread of railway wheels, and this was further developed in subsequent years [168].

There was therefore significant development and application of guided wave technology in simple structures before the surge in academic interest indicated in Fig. 1. At that stage, the instrumentation was relatively rudimentary, the display often being an oscilloscope screen, but the physics of the test and the transduction was well understood.

7.2. Post 1990 work on simple structures

7.2.1. Pipes

There are over 2 million km of long distance oil and gas pipelines [169] and they are usually inspected via magnetic and/or ultrasonic 'PIGS' [33,164]; only about 0.5 % of these lines are deemed un-piggable [170]. However, the shorter pipes both in and between plants are generally un-piggable and many are insulated or inaccessible for conventional point-by-point inspection. In the early 1990s there was a surge of interest in the inspection of un-piggable lines, initially driven by the issue of corrosion under insulation (CUI). It was seen that insulated pipes in generally good condition sometimes failed due to isolated, severe corrosion caused by, for example, a leak from a trace heating line running along the pipe. These isolated patches occurred at unpredictable locations, were very difficult to find before failure, and it was very costly to strip the insulation off the whole line in order to do a visual inspection. Guided wave inspection is potentially very attractive here as the insulation need only be removed at the location of the transduction system and tens of metres of pipe in each direction can be inspected from a single transducer position. Contemporaneous research on guided wave pipe inspection was done at Imperial College London in a consortium with TWI and a number of oil companies, and in USA at SWRI. Both developments used symmetric excitation to generate either the T(0,1) or L(0,2) mode, see Fig. 7. Both modes are essentially uniform through the pipe thickness and so provide full volumetric coverage. Imperial used dry-coupled piezoelectric transduction [122] while SWRI used a magnetostrictive system [171].

The Imperial work was commercialised in the late 1990s by two companies, Plant Integrity from TWI (now acquired by Eddyfi) [172,173] and Guided Ultrasonics Ltd (GUL) from Imperial [136,144,174]; the SWRI system is also commercially available [175,176]. Generation of a pure symmetric mode requires symmetric transduction so in the piezoelectric case it is essential that the transducers in the array such as that shown in Fig. 13 are matched in amplitude and phase, and that they are loaded uniformly onto the pipe; in the magnetostrictive case, the magnetostrictive strip must be uniform and be bonded uniformly to the pipe. This can be achieved satisfactorily, a 40 dB or better signal to coherent noise ratio typically being obtained with the GUL system.

The initial target of the Imperial College work was to detect corrosion defects removing around 5 % of the cross section of the pipe giving around a −26dB reflection coefficient, and preliminary site trials [177] showed that defects of the target size could reliably be identified. However, echoes were also seen from butt welds since the weld caps are not generally removed, so the weld presents a change in cross sectional area, and hence in effective acoustic impedance. This makes it difficult to identify defects at welds, and also introduces the possibility of a weld being incorrectly identified as a defect. This problem can be overcome by measuring the extent of mode conversion produced by a reflector.

If an axially symmetric mode is incident on an axially symmetric feature in the pipe such as a flange, square end or uniform weld (with weld cap), then only axially symmetric modes are reflected. However, if the feature is non-axially symmetric, such as a corrosion patch, some non-axially symmetric waves will be generated. These propagate back to the transducer rings and can be detected. If the T(0,1) mode is incident, the most important mode conversion is to the F(1,2) and F(2,2) modes which have similar group velocities to the T(0,1) mode in the operating frequency range (see Fig. 7). The amount of mode conversion obtained depends on the degree of asymmetry, and hence on the circumferential extent of the defect. At low circumferential extent (which is the region of interest for the detection of critical corrosion in practical situations) the mode converted F(1,2) reflection is almost as large as the direct reflection so if these two reflections are of similar size, it can be concluded that the feature is localised to a small region of the circumference; this allows the severity of the corrosion to be estimated [178]. The amplitudes of the axisymmetric and flexural modes can be obtained by analysing the distribution of motion around the pipe circumference; this is possible if the transducers of Fig. 13 are addressed individually or in segments around the pipe. For example, Fig. 14 shows the result of a test at a sleeved road crossing. The symmetric butt welds give large black (symmetric) reflections while corrosion at the entrance to the road crossing gives an extended reflection in which the symmetric (black) and non-symmetric (red) signals are of similar amplitude; smaller reflections are seen from the supports.

The sensitivity and range obtained are a function of the general condition of the pipe, how it is coated and how many features (welds, bends, tees etc) are present. The range typically varies between 5 m and 100 m in each direction along the pipe, the range being lowest on pipes with heavy general corrosion or thick bitumastic coatings. The attenuation in bitumen coated pipes is discussed in Refs. [179,180] and propagation in buried pipes was investigated in Refs. [181–183]; a coating to maximise the range on buried pipe was discussed by Leinov et al. [184]. The sensitivity is generally between 1 % cross section loss on an uncoated or fusion bonded epoxy coated pipe in good condition and 5 % loss on pipes in generally poor condition. The sensitivity is reduced at features where an echo is present in the absence of a defect, but this reduction in performance is largely removed when the system is used in a permanently installed monitoring configuration as discussed later.

The A-scan display of Fig. 14 gives no direct information on the circumferential location of the defect. (It is possible to obtain some information by identifying the principal axis of the flexural mode from the distribution of motion around the pipe.) Rose [185,186] showed that by phasing the excitation signal applied to different transducers around the pipe circumference it is possible to focus the guided waves to particular locations along and around the pipe, and so to improve the signal to noise ratio, and that the same concept can be used in reception to further enhance the focusing. This was implemented [187], but focusing by phasing the excitation signal in hardware involves a new test for each

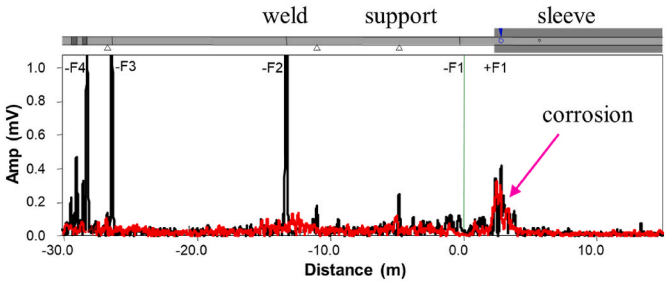

Fig. 14. Signal from test at entrance to sleeved road crossing. Cartoon at top shows series of welds (vertical line symbols) and simple supports (triangle symbols). Sleeve starts at +2 m. Weld labelled -F1 is very close to the transduction ring at 0 m so is in dead zone and no significant reflection is seen. Black indicates symmetric (T(0,1)) reflection; red indicates flexural (F(1,2)) reflection. (From Guided Ultrasonics Ltd). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

location of interest and this is very time consuming if multiple potential damage sites are to be explored. Davies and Cawley [147] showed that if each transducer or segment of a group of transducers around the pipe is excited in turn, the received signals on all transducers/segments in response to each excitation being collected, then it is possible to use synthetic focusing to give an image focused at all points along and around the pipe and so to produce a C-scan image of the unwrapped pipe; related work was also done by Mu et al. [148] and Velichko and Wilcox [188]. The procedure is analogous to the delay and sum imaging used in bulk wave phased array ultrasound [189] and is implemented in the commercial pipe testing systems.

The long range pipe testing technology is essentially a screening tool that detects and locates defects and gives a rough idea of their severity; a follow-up inspection at the identified location is then required to provide more precise sizing, and in particular a measurement of the maximum wall loss. This is usually done via ultrasonic thickness gauging that requires access to the defect site and so presents difficulties when the defect is in a buried section or is at a pipe support. It would therefore be very valuable to obtain a more accurate indication of the maximum defect depth from the guided wave results. In the sub-50 kHz frequency range used in most long range guided wave pipe testing, the reflection of the incident axisymmetric mode from a corrosion defect is roughly proportional to its circumferential extent [85,108,178,190] and has a power law relationship with the defect depth, the precise relationship being frequency dependent. However, to a first approximation, the reflection is proportional to the cross section loss. There is also dependence on the axial extent of the defect, particularly with sharp sided flaws, as discussed in the next section. This means that if the shape of the defect can be assumed then it is possible to estimate its maximum depth, but corrosion tends to be too irregular for this to be reliable [86].

Long range guided wave pipe screening is now in worldwide use, particularly in difficult-to-access applications such as road crossings, buried sections, offshore risers and locations that require extensive scaffolding to access for conventional inspection, as well as the insulated lines for which the method was originally developed; a review of the field is given by Ghavamian et al. [191]. The method is the subject of several international standards [192–194]; the equipment market is ≫$10 M per year and the value of service provision is >$100 M per year [195].

Unfortunately, while the transmission loss at the butt welds typically used on steel pipes in the oil and gas and other industries is modest, there is almost total reflection of waves propagating in the wall of cast iron pipes at the bell and spigot joints used in the water industry so long range inspection using these waves is not feasible and guided waves travelling in the water are not very sensitive to wall defects [196]; likewise the attenuation in medium density polyethylene (MDPE) [197] is too high for long range inspection to be feasible in plastic water and gas pipes.

Salzburger [159] has applied similar ideas to the inspection of lamp posts for corrosion just below the ground surface. There have also been developments of plate scanners in which a compact array of transducers analogous to the rings used in pipe testing is used to image the area of plate surrounding the transducer [132]. The original idea was to apply the method to the walls and floor of oil storage tanks, but the main requirement was to inspect the floor which is a complex construction of lap-welded plates so the application was not straightforward and required the tank to be opened and cleaned, at which point the established magnetic flux leakage (MFL) floor scanners can be applied. There was therefore insufficient demand to warrant commercial development. However, guided wave inspection of the annular ring of the tank floor from outside the wall has been successful, as discussed below.

7.2.2. Corrosion at inaccessible locations

Guided waves are also attractive for inspection at particular inaccessible locations, a common example being the detection of corrosion under pipe supports (CUPS) as illustrated in Fig. 15a. Here, the presence of corrosion is often easy to see visually, and may also be flagged via the long range guided wave screening discussed above that will screen multiple support locations in a single test. However, quantifying its severity by conventional bulk wave ultrasound involves lifting the pipe, and this can be dangerous if the corrosion is very severe. It would therefore be desirable to use guided waves not only to screen for the presence of defects, but also to obtain a quantitative indication of the minimum remaining wall thickness.

Guided wave tomography is in principle the most accurate method for mapping the topology of defects that does not require access to the defect site. It is usually configured with a ring of transducers around the area of interest, measurements being taken using each transducer as a transmitter in turn, the received waves on all the transducers being captured; it may also be possible to use two rings of transducers around a pipe on either side of the defect site. The method uses a dispersive mode so that the transit time or phase between each transmitter and receiver is dependent on the thickness over the wave path, the relevant dispersion curve providing a mapping between transit time/phase and thickness; the measurements over the different wave paths are then converted to a thickness map over the region between the transducers using a tomography algorithm. The method has been proposed by many researchers [200–208] using a variety of wave modes and tomography algorithms, and site trials have been conducted [209,210]. Huthwaite and Simonetti reviewed and explained the attractions and limitations of the different algorithms used in tomography [211] and Huthwaite [212] investigated the accuracy of the determination of the maximum defect depth for different defect geometries. He showed that the best of the algorithms gave accurate maximum depths for defects 1.6 wavelengths or more in diameter and up to 30 % of the plate thickness deep, a >10 % depth error being seen when the defect depth was 60 % of the wall. When the defect diameter was reduced to 1.1 wavelengths, there was >10 % error even at 10 % depth, and >20 % error at 60 % depth. If the A0 mode at around its maximum group velocity point is used [213] which gives a stable arrival time at which the phase can be measured (the phase velocity is dispersive at the maximum group velocity, as seen in Fig. 4a and b) then in a nominally 10 mm thick wall, the wavelength is about 18 mm, so accurate results would be expected with defect diameters of 30 mm and above. These results were on data generated from finite element predictions, so experimental results are likely to be less accurate. Experimental demonstrations of the technology have often been on defects many wavelengths in plan dimension and with no regions of sharply varying depth.

Mazeika et al. [214] have proposed using tomography for the inspection of the floors of large storage tanks using an array of transducers on the outer edge of the floor, beyond the wall. They estimated that, based on the attenuation of the S0 mode in a plate with typical floor thickness having oil on one side, it could potentially be used on tanks up to 20–30 m in diameter; the experimental validation was on an 8 m diameter model tank.

An issue with tomography is the determination of the wave arrival times/phase for an undamaged structure and this has sometimes been overcome by using the system in a monitoring mode, the transduction system being permanently installed so changes from the initial state are used as inputs [208,215]. Tomography typically requires a large number of individually addressable transducers and associated electronics as the sampling theorem dictates that full mapping of the region enclosed by the transducers requires a maximum spacing between transducers of half a wavelength, i.e. less than 10 mm in the example above. This complexity, and the concern that accurate results will not be obtained on small, sharp defects such as those shown in Fig. 15b, means that there has been a reluctance to invest in the development and deployment of tomography systems, and alternative methods giving a more attractive compromise between accuracy, cost and ease of use have been sought. One option is simply to use the transit time or phase to estimate the average thickness over the line between a pair of transducers, without attempting to produce a detailed thickness map, and this is used

Fig. 15. Photographs of (a) corrosion at a pipe support [198]; (b) deep, highly localised pits [199].

commercially [213,216]. This works very well when the corrosion is uniform, but is more problematic if the wall loss is localised [217]. Khalili and Cawley [218] reviewed the attractiveness of other alternative methods against key concerns:

- liquid loading – pipes and vessels are often full of liquid so it is important that the method is not severely affected by this due to e.g. energy leaking into the liquid, so increasing attenuation;
- surface coatings – coatings such as bitumen can cause very high attenuation;
- rough surfaces – scattering from generally corroded surfaces can cause significant attenuation, particularly at high frequencies;
- T-joints and welded patches – it is sometimes required to inspect past a T-joint e.g. transmitting under the vessel wall to access the interior of the annular ring in a storage tank, or inspecting under a welded pipe support;
- Sensitivity to different defect morphologies
  - o Sharp and shallow – small area defects with sharp edges such as those shown in Fig. 15b with maximum depth 30 % of the nominal wall thickness or less;
  - o Sharp and severe (deep) – as Fig. 15b but depth over 50 % of the nominal wall thickness;
  - o Gradual thinning over a wide area – some corrosion mechanisms tend to produce relatively large diameter (5 times the nominal wall thickness in diameter or greater) with gradually increasing depth towards the centre of the defect area.

Table 1 shows the performance of a variety of methods against these concerns:

- Long range S0/SH0 (L(0,2)/T(0,1) in a pipe) – this is the low frequency (typically <50 kHz) screening method of section 7.2.1;
- S0 in the 1.5 MHz-mm range (150 kHz in a 10 mm wall) [219] – this uses dispersion of the S0 mode to give the average thickness over the propagation path in a similar way to the A0 mode employed by Nagy et al. [213], the advantage of S0 in this range being that it is less attenuative in the presence of liquid loading, as shown in Fig. 5b;
- SH0 and SH1 in the 2.5–3 MHz-mm range (250–300 kHz in a 10 mm wall) [220,221] – this uses the non-dispersive SH0 mode to give a reference transit time and the dispersive SH1 mode to give the thickness information. Reflections of either mode can also be used to indicate the presence of sharp defects;
- HOMC/A1 - Balasubramaniam et al. [222,223] have introduced a technique that they call Higher Order Mode Cluster (HOMC) that typically uses a conventional piezoelectric transducer mounted on an angled wedge in order to generate higher order Lamb modes at around 20 MHz-mm which are claimed, due to their similar group velocities, to form a nondispersive cluster. However, Khalili and Cawley have shown that the features of HOMC are essentially those of the A1 mode in this frequency-thickness regime [224].

Table 1
Performance of different guided wave inspection methods. **Bold text** indicates desirable performance; *italic text* indicates poor performance. *When mode cut-off is possible (adapted from Ref. [218]).

Inspection method	Long range S0/SH0	S0	SH0	SH1	HOMC/A1
Centre frequency x Nominal thickness (MHz-mm)	<0.5	1.5	2.5	3	18
Effect of liquid loading	**Very low**	Average	**Very low**	**Very low**	**Low**
Effect of surface coatings	**Very low**	**Low**	*Very high*	*Very high*	**Very low**
Effect of rough surfaces	Average	**Very low**	**Very low**	Average	**Low**
Effect of T-joint/welded patch	*Poor*	Average	**Low**	Average	**Very low**
Sensitivity to sharp shallow defects	*Very poor*	Average	Average	**Good**	*Very poor*
Sensitivity to sharp severe defects	*Poor*	**Good**	**Good**	Average	**Very good**
Sensitivity to wide-area gradual thinning in reflection	**Very good**	*Poor*	*Poor*	**Good***	*Poor*
Sensitivity to wide-area gradual thinning in transmission	*Poor*	Average	*Poor*	**Very good**	**Good**

Table 1 shows that wide area gradual thinning is difficult to detect in reflection, the low frequency waves discussed in section 7.2.1 being the most suitable method. If a defect has sharp edges then there is significant reflection at all frequencies, other than at nulls caused by interference phenomena produced by reverberating reflections in the defect length [85]; this is illustrated in Fig. 16a that shows the reflection coefficient from axisymmetric defects in a pipe with different depths as a function of the defect length to wavelength ratio. The reflection coefficient reaches maxima at defect length to wavelength ratios of 0.25, 0.75, 1.25 etc, successive maxima being the same amplitude so the defect can be detected at both low and high frequencies. However, when the defect is gradual, the first peak is still evident, but successive peaks decrease rapidly in amplitude, as shown in Fig. 16b, so the defect is only detectable at low frequencies. The torsional wave speed is ~3.2 mm/μs, so for a defect 50 mm long, the maximum reflection would be expected at a wavelength of ~200 mm ($L/\lambda = 0.25$) giving a frequency of ~16 kHz; the peak in Fig. 16b is at a somewhat higher frequency than this because the defect width is not uniform over its length. Significant reflections can also be seen if a higher order mode is incident and the defect depth is sufficiently large that the product of the frequency and the remaining ligament thickness is below the cut-off frequency of the mode, so it will not propagate across the defect.

A combination of the SH0 and SH1 modes gives generally good

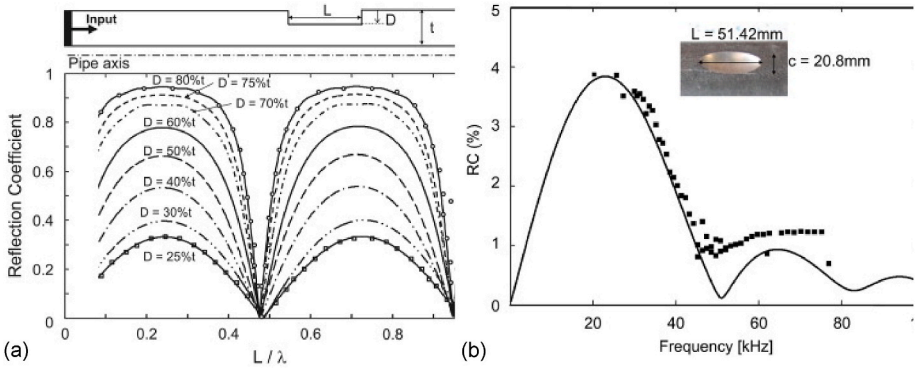

Fig. 16. (a) reflection coefficient of torsional mode from axisymmetric defects with different depths as a function of the defect axial length to wavelength ratio; (b) reflection coefficient of torsional waves from elliptical defect in 4 inch schedule 40 pipe with maximum depth 4.5 mm (75 % of wall thickness). FE predictions - solid line; experiments - square points; random noise floor ~1 % reflection (From Ref. [86]).

performance, particularly if both reflection and transmission measurements are used; this is easy in the case of corrosion at a pipe support using propagation in the circumferential direction around the pipe, as the same receiver can detect both reflections from a defect and transmission around the pipe circumference across the defect [225]. However, the size of the reflection and transmission changes produced by a defect are a function of the defect length normal to the ultrasound beam (in the axial direction if circumferential propagation is used) so a single measurement is not sufficient for depth estimation. This problem can be overcome by scanning along the pipe and testing over a wide frequency range [225].

Deep pits with small diameter are very difficult to detect and size if frequencies in the 250 kHz range or lower are used because, as discussed with tomography, the sensitivity is poor when the defect diameter drops below a wavelength. The SH0 wavelength at 250 kHz is ~13 mm so defects with diameter less than ~15 mm are unlikely to be detected and sized reliably. Here, the HOMC/A1 method is very useful as it operates at ~18 MHz-mm (1.8 MHz in a 10 mm wall thickness, so the wavelength is under 2 mm). The surface motion in HOMC/A1 is very low, so it passes easily under walls and is minimally affected by surface roughness or coatings. It is therefore ideal for the detection of severe defects in the annular ring of storage tanks, the transducer being placed outside the wall. However, the low surface motion means it is not sensitive to shallow defects and only a rough indication of severity is obtained.

The SH0/SH1 method with EMAT transduction is in common commercial use [220,221,225], sometimes to follow up indications from the low frequency screening of section 7.2.1, and there is increasing application of HOMC/A1; there is also some use of the S0 mode.

7.2.3. Other one dimensional structures

There has been extensive international interest in rail inspection using guided waves [226–232] and the topic has been reviewed by Loveday [233] and Ge et al. [234]. A commercial instrument has been produced [235] that has been used particularly at level crossings to detect foot corrosion that is not visible without lifting the rail; systems to be mounted on an inspection car have also been investigated at Pennsylvania State University and UCSD [233]. Propagation distances of well over 100 m can be achieved, particularly in modes whose energy is concentrated in the head or web of the rail and so are not significantly attenuated by contact with the sleepers (ties) and the rail clips. This is attractive for permanently installed monitoring to detect rail breaks, as discussed in section 8.

Post-tensioning is a common construction method used to increase the load bearing capacity of concrete structures. The technique is widely used in bridge construction, where steel tendons are run through ducts in the concrete after the concrete has hardened; these tendons are subsequently tensioned. Guided waves can easily be used to inspect both single and multi-wire tendons if they are simply in an unfilled duct [236,237] but if they are grouted, the guided wave energy leaks into the grout so the propagation distance is greatly reduced [238] and it is typically only possible to inspect less than 1 m from the excitation point [239], though in some cases the attenuation in the central wire of a multi-wire tendon may be lower. The attenuation in an embedded rod at a given point on the dispersion curve reduces with the rod diameter, so longer inspection ranges can be obtained in structures such as the rock bolts used in mining, the full 3 m length of bolts being inspectable [239]. This case is shown in Fig. 8 where a relatively low attenuation point is seen at a surprisingly high frequency, the explanation being that the surface motion, and hence leakage into the surrounding rock, is low at this point. A simple compression wave transducer of the type commonly used in bulk wave ultrasound can be used to excite the mode from the end of the bolt; the attenuation of most of the modes excited is very high, so the system acts as a mechanical filter and it is very easy to obtain clean signals [239]; this probably means that many practitioners think that the wave propagating is a simple compression wave. The guided wave inspection of rock bolts has also been investigated in South Africa [240].

7.3. Complex structures

7.3.1. Bonded joints

Adhesive and diffusion bonded joints are widely used in airframes and other structures, and it would in principle be attractive to test them by simply propagating a guided wave either across the joint or along it, rather than scanning point-by-point using conventional bulk wave ultrasound. Most inspection of adhesive joints involves looking for complete disbonds, but poor cohesive properties of the adhesive itself, or a weak adhesive-adherend interface caused by surface contamination or environmental attack, are also concerns; kissing bonds in which the surfaces are in intimate contact but with minimal joint strength have also generated a great deal of research. Complete disbonds are fairly straightforward to detect via changes in the reflected or transmitted guided wave signals, but the other defects are much more challenging, a review of the extensive early effort in the field being given by Lowe and Cawley [241]. None of the guided wave methods has performed as well as high frequency bulk wave ultrasound [242] for the detection of interfacial problems, and in practice the complexity of guided wave inspection due to the need to maintain mode control, together with the relatively complex received signals, means that guided waves cannot compete with conventional bulk wave scanning for disbond detection.

7.3.2. Feature guided waves

Fan and Lowe [243] showed that guided waves can propagate along a butt weld joining two plates, the energy of the wave being partially trapped by the geometry change produced by the weld cap. This means that, while energy does leak into the two plates, the attenuation is lower

than would be expected from beam spreading in a plain plate. It was shown [244] that a shear guided mode was able to detect cracking both parallel to and normal to the weld, as well as holes simulating corrosion in the heat affected zone adjacent to the weld. An inspection range of the order of 1 m from a single transducer position would be feasible, so the method is in principle attractive. However, it is less familiar than conventional bulk wave ultrasound, its sensitivity is lower and the transduction is more complex. This means that it has not been used in practice though it would potentially be very attractive in a case where the weld goes beyond an obstruction that makes access to the whole area of interest difficult. The concept has also been extended to welded T joints [245] and adhesive bonded joints [246].

7.3.3. Airframes

Given the large area to be inspected, the guided wave inspection of airframes would in principle be very attractive, removing the need for two dimensional scanning being particularly beneficial for in-service testing. The difficulty is that bonded or riveted ribs and stiffeners are typically positioned in a criss-cross pattern at a pitch of 300 mm or less and these generate significant reflections. Dalton et al. [247] showed that while propagation across one stiffener was feasible at a sufficiently small wavelength to give satisfactory spatial resolution, propagation across multiple stiffeners was not realistic so long range inspection and monitoring is not possible in these structures. A0 and S0 mode signals at frequencies below 100 kHz will propagate across multiple stiffeners, and these are the modes received in typical acoustic emission testing [248]. However, the wavelength at these frequencies is too large to obtain sufficient spatial resolution of the stiffener reflections for damage to be detected in the A-scan signals obtained in a single test. Therefore, guided wave testing of airframes is only feasible in a structural health monitoring mode, current signals being compared with a baseline; this is discussed in section 8. Even in monitoring mode, the area covered is likely to be modest since there is unlikely to be sufficient sensitivity to damage at the low frequencies at which long range propagation is possible [247]. While the work of [247] considered metallic skins, similar conclusions apply to composite material constructions, and monitoring of composite structures is also reviewed in section 8.

The acousto-ultrasonic technique involves propagation over a modest distance (typically less than 200 mm) between a transmitter and receiver, the amplitude of the received signal being affected by defects in the wave path. This has been promoted for composite laminates and adhesive joints as a more rapid technique than conventional bulk wave scanning since the scan pitch in one direction is increased to the distance between the transducers, rather than the diameter of a single transducer. The method excites Lamb waves and uses the signal processing originating in acoustic emission testing to give an easily interpretable result [249]. However, the mode control is poor and the results are not reliable for the detection of anything other than severe defects. It is therefore not widely used and its attractiveness is reduced with the advent of array scanners that substantially increase the speed of bulk wave scanning [250].

8. Applications in structural health monitoring

8.1. Initial motivation

In some cases, the original motivation for permanently installing transducers, and so enabling structural health monitoring, was simply to avoid repeated access costs. For example, in the pipe testing of section 7.2.1, if the pipe is buried or if scaffolding is required, accessing the pipe to attach the sensor can cost substantially more than purchasing the transduction system. Therefore, if regular inspection is required, it becomes very attractive to permanently attach a sensor to the pipe; the sensor can then be remotely operated from a convenient location and there is no need to access the pipe after sensor installation unless damage develops. Hence, the original applications of permanently installed guided wave pipe monitoring systems were simply designed to reduce access costs for repeat inspections, and the results of each inspection were analysed independently of previous results. However, it quickly became clear that the results of successive measurements were very consistent, so introducing the possibility of improving detection sensitivity by subtracting the current result from one taken previously [251]. This moves the system from non-destructive testing (NDT) mode to Structural Health Monitoring (SHM) mode. The lowest sensitivity in one-off guided wave inspection is achieved at locations such as a weld or tee where an echo is received even in an undamaged structure, since a defect at a feature location will cause a much smaller percentage change in the signal than at a location where the only received signal in an undamaged structure is due to noise. If subtraction is successful, the sensitivity at features may approach that in regions of simple structure.

Likewise in rail, the original motivation for permanent installation was the need for essentially continuous monitoring for rail breaks. In this application, each received signal can be processed independently as the change in transmission with a complete break is very large. However, once the system is installed, more sophisticated processing introduces the possibility of detecting defects before they have propagated to a complete break.

In feature-rich structures such as airframes, the signals obtained in a one-off test are too complicated to be interpretable and baseline subtraction in SHM mode is the only credible approach to reliable damage detection and this has led to the surge in guided wave SHM research from around 2005 seen in Fig. 1.

8.2. Compensation for environmental effects

8.2.1. Closed form algorithm approaches

Unfortunately it is not possible simply to subtract the current signal from the baseline as the signals need first to be compensated for the effects of environmental and operational conditions, the predominant factor typically being temperature [126,252–255]. This is because the wave speed is a function of temperature, so the phase of the received signal will change with temperature, leading to residual signals after subtraction that can be much larger than typical defect reflections if no compensation is applied. Two widely used methods for temperature compensation are baseline signal stretch (BSS) and optimal baseline selection (OBS). BSS uses a single baseline signal as a reference and the current signal is either compressed or stretched in time until the residual signal obtained after it is subtracted from the baseline signal is minimised [256,257]. In OBS, a set of baseline signals is stored and the one deemed most similar to the specific current measurement is used for amplitude subtraction, often after also applying BSS [254,256]. In recent years, a number of other temperature compensation techniques have also been developed [255].

The BSS method applies a single stretch factor to the whole signal, but a major cause of the coherent noise seen in guided wave signals is interference between the different unwanted modes that are generated, as seen in Fig. 10; these modes have slightly different velocities, so the interference changes with position, and it is not removed by a single, global stretch factor. Fendzi et al. [258] proposed to divide a signal into time windows containing the arrival of a single wave mode and to apply temperature dependent amplitude and phase shift compensation factors within each window. In practice this is complicated since the time windows would often contain multiple overlapping mode components. Zoubi and Mathews [259] used a mode decomposition algorithm [260] to decompose a set of baseline signals into possibly overlapped mode components and sought to determine mapping functions that describe the variations of each mode across the operating temperature range; current measurements would then be compared to reconstructed baseline signals obtained from the predicted mode contributions at the measured temperatures. Douglass and Harley [261] presented a method based on dynamic time warping that maps each signal sample of the baseline measurement to a (possibly) different sample of the current

signal, hence performing local stretches, although this can overfit the data, potentially masking defect reflections. Dao and Staszewski [262] avoided direct temperature compensation by using a cointegration technique that already takes into account environmental changes and looks for departures from a stationary residual signal to detect the occurrence of damage.

Mariani et al. [263] have developed a novel temperature compensation method denoted Location Specific Temperature Compensation (LSTC) [264] that has been successfully applied to torsional guided wave signals collected by permanently installed monitoring systems installed on pipes. The LSTC method comprises a calibration phase and a monitoring operation phase. In the former, a set of baseline signals is acquired across the temperature range of interest, which is used to construct a set of calibration curves, one for each signal sample, i.e. each point on the captured waveform, which can be associated with a particular location on the structure from knowledge of the wave speed and time of arrival. Each curve shows how the expected signal amplitude at each location in the absence of damage varies with temperature. In the monitoring operation phase, when a new measurement is acquired, at each point the expected value at the relevant temperature obtained from the calibration curves is subtracted from the measurement itself. Thus, in the absence of damage, the expected value of the residual signal is zero. More details on the procedure can be found in Mariani et al. [263] where it was also indicated that the sequence of residuals obtained at any signal sample follows a normal distribution with mean close to zero and standard deviation related to the incoherent noise level affecting the measurement.

If guided wave measurements can be processed and reduced to sequences of normally distributed samples, one can shift the paradigm of damage detection using guided waves into a task of change detection as typically performed in the statistical process control (SPC) field, the generalized likelihood ratio (GLR) change-point approach being particularly attractive. It was shown [265] that reliable detection of small defects giving reflections 1.5 times the standard deviation of incoherent noise is possible with virtually zero false calls. However, this requires the residual signals after subtraction when the structure is undamaged to be normally distributed, and this must be checked before applying the method.

The detection limit of 1.5 times the incoherent noise level shows that it is important to minimize this noise in the system design. However, this sensitivity assumes that, in the absence of damage, the only changes in the received signal after the LSTC-based temperature compensation are due to incoherent noise, and so can be removed by the averaging implicit in the GLR process. In practice there may be other effects such as loading changes [255,266]; in the pipe context, different fluid loads in the pipe could change the small reflections [267] produced at simple supports. Likewise, hysteresis effects of temperature on the sensor may be seen, meaning that its output is not a simple function of the current temperature; also, ageing of the transduction system or the adhesive attaching it to the pipe may cause the signal to drift with time. These effects are likely to be seen over longer timescales than the few weeks over which the tests of [263,265] were conducted.

SHM schemes that do not require a baseline are potentially very attractive and have been the subject of significant research. In simple structures such as pipe or rail where the signals are sufficiently simple for anomalies due to damage to be detectable in a one-off test, it is possible to operate without baselines, though the sensitivity can be improved if multiple readings are considered, and this does require compensation. Likewise, tomography is inherently self-calibrating, though some adaptation for SHM is desirable [268]. There has been much interest in time reversal schemes for damage detection and some authors have claimed that they enable damage detection without compensation for changes in environmental and operational conditions. However, the work of [269] shows that raw time reversal does not achieve this, but goes on to claim that better results can be obtained with a refinement of the method.

Self-calibration of the signals before baseline subtraction using a scheme such as that proposed by Vogt et al. [270] for pipes can be used to compensate for changes in the transducer output, though the scheme relies on the multiple reflection between two reflectors on either side of the transduction system to be above the noise floor, and this is not always the case.

There has been interest in schemes involving pairs of transducers to provide calibration, a fully self-calibrated system being proposed by Achenbach et al. [271]. Huan and Li [272] provide a useful review of various baseline free possibilities and propose a method involving the use of pairs of transducers to calibrate the outgoing signal, while Alem et al. [273] use cross correlation between the signals received by pairs of transducers on similar paths to assess whether changes have occurred. Detecting the mode conversion produced by damage by processing the signals received at transducers on the two sides of a plate to distinguish between symmetric and anti-symmetric modes [274], or using a wavelet transform to separate modes [275] has also been investigated. All these methods are much more difficult to implement on complex structures where mode conversion happens even in the absence of damage, and similar paths are difficult to define.

8.2.2. *Machine learning*

Machine learning offers great potential for identifying damage in SHM signals and distinguishing it from signal changes due to environmental and other operational effects [276–283] and it can in principle be applied alone, or after compensation via the methods discussed above; reviews of some of the work in the field are presented by Yang et al. [284], Hinders and Miller [285] and Sattarifar and Nestorovic [286]. Recent work has suggested that machine learning may perform better than the closed form algorithmic compensation methods [287, 288] and this is a very promising area for future research.

Machine learning algorithms need to be trained with large volumes of data, but much of this can be generated from tests on undamaged structures under different environmental conditions. Realistic experimental or simulated damage responses are then needed for testing the system, as with any SHM scheme; this is discussed further in Section 8.5. There has therefore been a great deal of interest in supplementing experimental data with finite element simulations [289] and this is becoming more feasible as full 3D models become tractable, as discussed in Section 3. However, in SHM applications, it is essential to combine high sensitivity to defects with a very low false call rate. For example, suppose that the probability of false call (PFA) in an inspection is 10 % and the probability of detection (POD) of a real defect is 90 %, and further suppose that the a-priori probability of a callable defect being present is 5 %. Then in 1000 inspections we expect 95 false calls, 45 correct defect calls and 5 missed defects; therefore of 140 defect calls, ~70 % are false. If the a-priori probability of a callable defect being present is reduced to 1 %, then more than 90 % of defect calls will be false [290]. Unfortunately it is very difficult to simulate reliably all the effects seen in real, long-term tests so training algorithms solely on simulated data is liable to lead to the signal changes seen in real tests on undamaged structures being interpreted as damage, leading to a high false call rate.

It is tempting to think that machine learning will extract defect signatures from arbitrary amounts of random or coherent noise, so making it less important to improve signals of the type shown in Fig. 10 by controlling the modes generated and received. However, as with any form of signal processing, the signal to noise gain from machine learning will be limited and be affected by the consistency of the signals generated and received on undamaged structures. Given the need for low false call rates, mode control is likely to remain an important factor in guided wave system design.

8.3. Simple structures

8.3.1. Pipes

Permanently installed guided wave pipe monitoring systems have been available for some years and are in worldwide use. For example, Fig. 17a shows a guided wave permanently installed monitoring system (gPIMS®) produced by Guided Ultrasonics Ltd [136]. As in the detachable system of Fig. 13, it comprises two rows of piezoelectric elements, but in this case the design is modified and the assembly is over-moulded in polyurethane to provide environmental protection. The unit is bonded to the test pipe and connected to instrumentation that can either be permanently installed to carry out the test and send the resulting data wirelessly to the relevant engineer, or can be operated via plug-in instrumentation if only occasional testing is required. Further details are given in Ref. [251] and the system performance in a blind trial is described in Ref. [291].

Fig. 17b shows an installation on a 14 inch pipe at a gas processing plant before the insulation has been reinstated, a typical A-scan in the direction of the 45° bend shown at the bottom right of Fig. 17b being presented in Fig. 17c. The bend section is welded on each side and there is also a clamp securing trace heating elements to the pipe close by; there is a similar clamp about 3.3 m from the gPIMS® transducer. Fig. 17c shows that the bend, its two associated welds and the clamp give a complex, composite reflection, while the isolated clamp is not readily separated from the background signal level. Fig. 17d shows the time history of the LSTC-compensated residual signal from the bend at the location of the red dot in Fig. 17c, together with its moving average. The baseline period over which the variation of the A-scan signal with temperature at each location was curve-fitted to enable LSTC compensation is also shown in Fig. 17c. The residual is fairly constant over the first ~8 months but then decreases significantly; this was subsequently associated with deposit building up in the pipe between the gPIMS® transducer and the bend. A cleaning PIG was then passed down the pipe and the residual signal increased markedly, as shown at the end of the monitoring period in Fig. 17d. The large increase in signal after cleaning suggests that there was already significant deposit in the pipe before the further increase detected just before cleaning. It would be necessary to re-calibrate the monitoring system by considering a fresh 'baseline' period in order to be sensitive to small changes after the cleaning operation. The detectability of defects in the presence of gradual changes due to deposit formation remains very good at locations away from pipe features where the corrosion will introduce a new echo; at a feature such as a weld or clamp where the reflection amplitude is affected by deposit growth in the pipe between the transducers and the feature, the detectability of the corrosion will be a function of the rate of corrosion growth relative to the deposit growth.

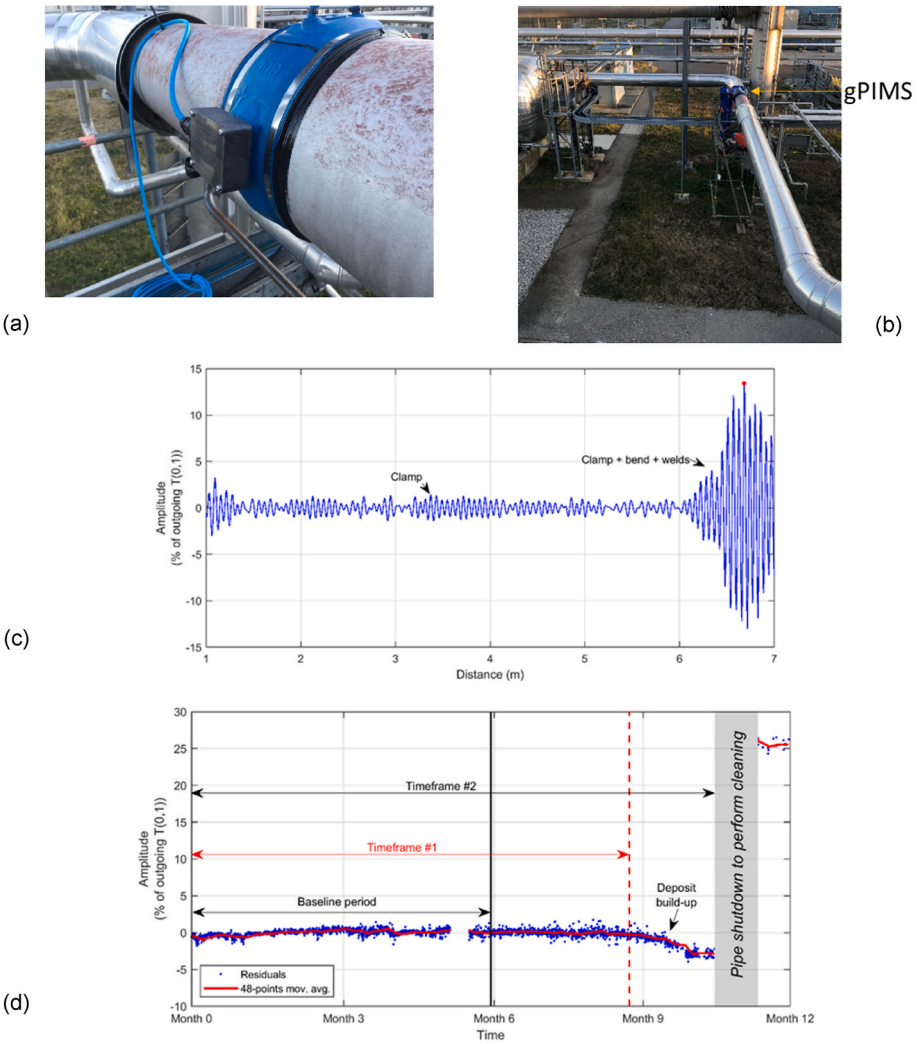

Fig. 17. Installation at refinery (a) Completed installation before insulation reinstated; (b) view of pipe during installation; (c) A-scan signal from installation of (a), (b); (d) Variation of LSTC residual with time and its moving average at position of red dot in (c) at the bend. (Moving average restarted after cleaning.) The baseline period was used in the LSTC compensation process [263] (From Ref. [251]). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Fig. 18a shows the maximum of the moving average of the residual signal at each location along the pipe over Timeframe #1 before the deposit build-up shown in Fig. 17d. The maximum variation before the bend region is about 0.75 %, suggesting that a detection threshold of 1.5 % is appropriate; the maximum variation at the bend is 1.3 %, indicating that the detection threshold here should be 2.6 %. In a one-off inspection based on the A-scan of Fig. 17c, setting the detection threshold at 6 dB above the maximum coherent noise level seen before the bend of 3 % at ~1.1 m gives a call level of 6 %. There is therefore a 12 dB improvement in sensitivity before the bend from using the monitoring approach; reliable defect detection in the bend region would be very difficult in a one-off test, but employing monitoring gives a reliable detection threshold of 2.6 % which is only ~5 dB poorer than that obtained away from features.

Fig. 18b shows the corresponding data over Timeframe #2 of Fig. 17d, including the period of rapid deposit build-up. The variation from the baseline is increased and comfortably exceeds the detection threshold at the bend location where the increased attenuation caused by the deposit causes the maximum signal reduction. The signal change is largest here both because it is the location of the largest reflection, and also because the effect of the deposit increases with signal propagation distance over the region where the deposit is formed. The residual signal also increases in the region before the bend and crosses the detection threshold at some locations. Therefore the gPIMS® monitoring system can also be used to identify deposit build-up and the need for cleaning.

A permanently installed guided wave system based on the Teletest guided wave inspection system has also been used to monitor a problem in boiler spines in a UK nuclear power plant [292]. Herdovics and Cegla have developed an EMAT-based pipe monitoring system [293], together with a temperature compensation system [294] and have investigated its long term stability [295]. A review of guided wave SHM of pipelines, and in particular the reduction of false alarms, is given by El Mountassir et al. [296].

8.3.2. Rail

As discussed above, guided waves will propagate long distances in rail so the break in transmission between a transmitter and receiver 100s of metres apart on the rail can be used to detect a complete break; the change in signal produced by the break is very large and so can be detected in a one-off test without comparing the current signal with a baseline. Loveday has worked extensively on this problem [233] and has produced a system that is installed on a heavy haul freight line in South Africa [297]. The system operates in pitch-catch mode with alternate transmit and receive transducers spaced approximately 1 km apart over a total distance of 840 km. If the acoustic signal is not received at the receive station an alarm is triggered to indicate a break in the rail between the transmit station and the receive station. The system is permanently installed, powered by solar panels and issues broken rail alarms using the GSM (Global System for Mobile Telecommunications) network where available, and digital radio technology in other areas; the entire length of rail is interrogated every 15 min. The system is now being extended to the detection of smaller defects and this does require comparison with previous signals, this being done via principal component analysis [298]. Liu et al. have worked on a system for monitoring switch rails [299] that are particularly prone to damage and Masmoudi et al. [300] have recently reviewed the use of guided waves in rail health monitoring.

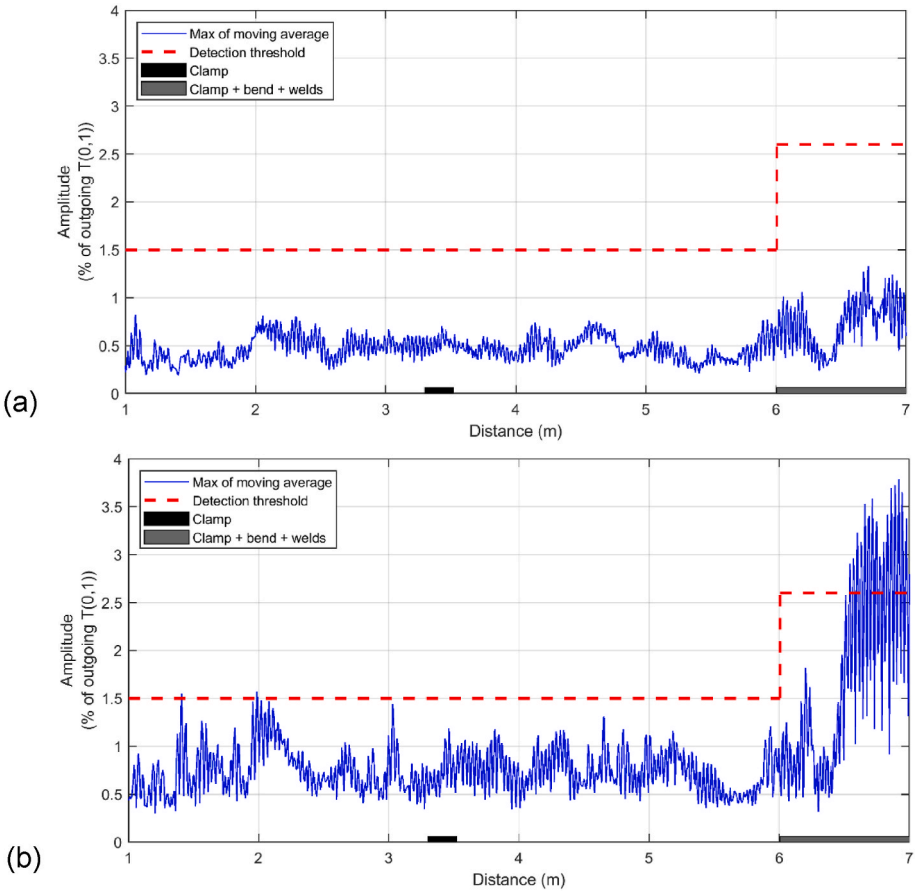

Fig. 18. Maximum of moving average residual signal at each location for case of Fig. 17. (a) in "Timeframe #1" shown in Fig. 17d; (b) in "Timeframe #2" shown in Fig. 17d. Red dashed line indicates proposed detection threshold (From Ref. [251]). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

8.3.3. Plate-like structures

There is a huge literature on the use of guided wave arrays to monitor plate-like structures. Most interest has been in sparse arrays, see e.g. Refs. [123,254,301], but concentrated arrays have also been considered [131,302,303]. Sparse arrays can give relatively uniform coverage, whereas the sensitivity of a concentrated array reduces with distance due to beam spreading; this is particularly true with the compact, omni-directional array of [131] and is less serious with a wide, linear array such as that proposed in Ref. [302]. However, full coverage is more difficult with a linear array. Wiring is a significant issue in distributed arrays since each individual transducer must be individually addressable and synchronised to a common time base. It is unlikely that trigger timing can be done sufficiently accurately with individual, self-contained transducers receiving instructions and transmitting data via a wireless link, so physical wiring connections are required; this is straightforward in laboratory conditions, but the wiring is likely to be vulnerable in field applications. One option is to mount the transducers on a flexible printed circuit board (PCB) that is then either bonded to the structure or embedded in it; this is the route taken by Chang at Stanford and an associated spinout company [128] who term the PCB and associated transducers a 'smart layer' that can either be embedded in composite materials or surface mounted on any structure [304,305]. This is connected to instrumentation that can transmit the data either wirelessly or via wired connections to an analysis centre where bespoke software assesses whether any damage is present and its severity. It was originally developed for aerospace applications, but multiple potential applications have been proposed [128]. However, these are mainly at the trial rather than routine industrial deployment stage.

Since damage will be detectable by multiple pairs of transducers, but will not be uniquely located by a single pair, it is generally necessary to combine the information from the different pairs. This is most commonly done by producing an image, the simplest imaging algorithm being 'delay and sum'. By trigonometry, the time a signal takes to travel from the location of the transmitter, (x_i, y_i), to any point of the panel, (x, y), and to the receiver, (x_j, y_j), is [306,307].

$$t_{ij}(\mathrm{x},\mathrm{y}) = \frac{\sqrt{(x_i - x)^2 + (y_i - y)^2} + \sqrt{(x_j - x)^2 + (y_j - y)^2}}{V_g} \tag{7}$$

where V_g is the group velocity of the mode of interest at the centre frequency of the input signal. If S_{ij} is the envelope of the residual signal for transducer pair ij, obtained after processing of a current signal with the chosen temperature compensation strategy and subtracting it from the baseline, then $S_{ij}(t_{ij}(x, y))$ is the amplitude of the residual signal corresponding to a given point (x, y) of the plate. By calculating this for every point of the plate, a spatial map of the amplitude of the residual signal is obtained. If a reflection is present in the residual signal, an ellipse with its foci at the locations of the two transducers will be formed; the ellipse is therefore a representation of the most probable locations of the reflection in relation to transducer pair ij. If this procedure is repeated for a minimum of three transducer pair combinations, triangulation is possible, and the point at which the three ellipses intersect gives the location of the defect. If repeated for all the transducer pair combinations in the array, and if the pulse-echo signals, S_{kk}, are not captured, the final amplitude map will be given by [308].

$$I(x,y) = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} S_{ij}\left(t_{ij}(x,y)\right) \tag{8}$$

where N is the number of sensors in the array. Although the image generated is superficially similar to that obtained in tomography, the two should not be confused; in sparse array imaging, the system is usually highly under-sampled and there is an implicit assumption that there is a limited number of locations where change has occurred, whereas tomography uses a more fully sampled array and produces a genuine property map. Qian et al. [309] have proposed using machine learning to improve the resolution with sparse arrays, showing encouraging results in comparison with results from a fully sampled array, but any analysis of under-sampled data must involve implicit assumptions about the number and distribution of defects. The sparse array imaging method has been applied to many examples of simple metallic and composite plates, see e.g. Refs. [256,310–313] and more complex structures [314,315]. For example, Fig. 19 shows an image of a ribbed container panel with a 10 mm hole drilled at one location to simulate damage; the highest image amplitude is seen at the defect location, but there are other areas with relatively high amplitudes, introducing the possibility of false calls. There has been extensive work on imaging algorithms to give the better defect discrimination than simple 'delay and sum' [316–319] but most of the work has used data captured over a fairly short period either side of real or simulated damage being introduced.

The size of defect that can be reliably detected by an SHM system is dependent on the variation in signals seen on an undamaged structure over time and on the statistical distribution of the signals [265]. It is therefore very important to design a stable transduction system and to test it over extended periods, as was done with the pipe monitoring system of section 8.3.1 [251]; only long term tests in realistic environments can determine whether the sensitivity obtained in short term tests is reproducible in industrial practice. In essentially one dimensional structures such as pipe or rail there is a single residual amplitude corresponding to each axial location, and this is tracked with time as shown in Figs. 17 and 18, whereas with a sparse array on a plate-like structure, either the image amplitude at each point can be tracked, or the contribution from each transducer pair could be monitored. Real damage will usually tend to grow more gradually than in lab tests so any trend removal operations run the risk of removing the damage signature. The issue of long-term durability of sensors and their coupling to the structure has been reviewed by Brunner [320] while Attarian et al. [109] have investigated the long-term stability of a plate monitoring system using bonded piezoelectric transducers and illustrated some of the issues that must be addressed in system design. Yang et al. [321] have investigated the impact of data compression and de-noising strategies for damage detection in long term tests on a plate and Ding et al. [322] have looked at the issue of varying structural load with time on an aircraft structure.

A long term test has been carried out on a steel tank at Bristol University, data being collected from nine transducers over a period of 8 years. Results using the optimal baseline selection (OBS) temperature compensation method discussed in section 8.2 showed that the results

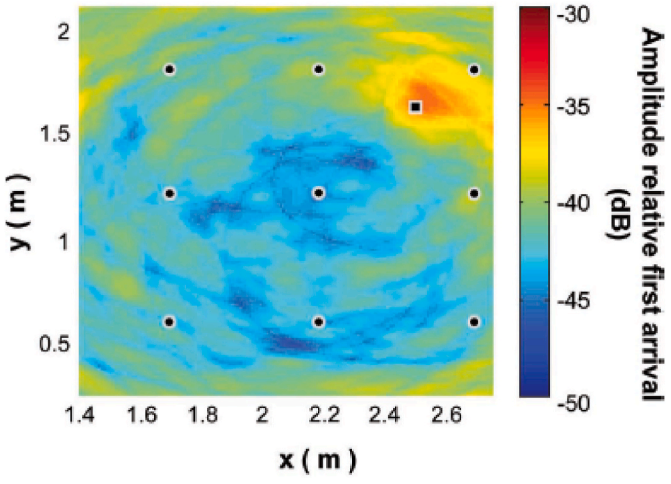

Fig. 19. Image of container panel obtained by processing signals of a 9 element S0 mode array with the algorithm of equation (8) when a 10 mm diameter hole was introduced at the position of the dark square. Dark circles indicate the transducer positions (From Ref. [314]).

after 3 years could not satisfactorily be compared with the original measurements [323], but recent work has shown much better stability from the same data when using a machine learning algorithm [287].

An alternative approach that greatly reduces the mass and complexity of the wiring required is to use a small number of permanently attached transducers only as transmitters, and to measure the wavefield produced via a remote scanning laser vibrometer [139,140]. This has the advantage that the wavefield is measured over the whole area of interest, rather than at well separated points, potentially making it more sensitive to defects. However, it depends on each location of interest being in the line of sight of the vibrometer, so excluding locations under supports or in the interior of a structure. It has potential applications in metallic or composite panels in aircraft [324,325] and wind turbines [326] and it may be possible to increase the field of view by putting the vibrometer on a drone [327]. Esfandabadi et al. [328] have investigated a deep learning approach to facilitate a reduction in the measurement point density required to achieve satisfactory spatial resolution, so speeding up the scanning. The use of wavefield measurements has been reviewed by Chia et al. [142].

8.4. Complex structures and localised monitoring

While global monitoring of a whole structure is very attractive and is achievable in some instances such as pipe or rail systems, in other cases the compromises required to make the system practical may reduce the sensitivity to damage too much to be acceptable. For example, while propagation of low frequency modes across multiple stiffeners is possible [248,314], the sensitivity to damage may not be sufficient if the sensor spacing is significantly larger than the stiffener spacing. If the sensor spacing is reduced, the weight of the system and the complexity of the cabling is increased and may become impractical if a large area is to be covered. Much of the work on stiffened panels has considered an array of sensors on either side of a stiffener, see eg Refs. [279,329]. There are many cases where monitoring is only justified in a localised area that is known to be prone to damage or is particularly critical to the integrity of the structure, an example being a repair [330]. In these cases a system such as the Smart Layer produced by Accelent [304,305] may be attractive, and it may be possible to take advantage of the reduced area coverage required by operating at higher frequencies that increases the sensitivity of the test at a cost of inspection range [331–333]. Ramalho et al. [334] have reviewed possible approaches to monitoring of adhesive joints.

8.5. Validation of performance

Receiver operating characteristic (ROC) curves [335] are routinely used in the evaluation of diagnostic tests in medicine and engineering; they plot the sensitivity (probability of detection (POD) in NDT/SHM terminology) for the test against its specificity (probability of false alarm (PFA) in SHM) for a given change in condition (defect size and type in SHM), both axes being on a (0,1) scale. The ideal operating point is at the top left-hand corner of the plot, corresponding to unity probability of detection and zero probability of false alarm, but of course this is not practically attainable.

Unfortunately, it is typically even more costly to define ROC curves experimentally in SHM applications than it is to conduct POD trials in NDT. This is because the SHM system must be installed on many structures of the required design that are cycled through the full range of operational conditions, damage being introduced in some of the structures at a representative range of locations. In most cases this is totally impractical and its cost would preclude the application of the SHM system. In NDT there is an increasing move towards model-assisted POD (MAPOD) evaluation [336] and Calmon et al. [337] have proposed a similar method for evaluating POD in guided wave SHM systems; Liu et al. [338] have proposed a related approach for ROC curve generation in SHM.

Modern computational resources mean that the signature produced by damage in bulk wave ultrasound, guided wave ultrasound and other techniques can be reliably predicted, even when it has complex shape [86]. In contrast, the reliable prediction of signal changes due to environmental and other variability is not possible since it is difficult to capture all the possible effects; it is therefore not feasible to predict the signal changes seen in the absence of damage. However, it is straightforward to obtain experimental data with environmental variation from an SHM system installed on a typical undamaged structure. Measured data can therefore be obtained over multiple environmental cycles on an undamaged structure and predicted damage signatures can be added at different locations with different growth patterns. The effect of different degrees of environmental variation on the ROC curves is then straightforward to simulate by appropriate selection of signal sets, and the effect of varying damage severity, multiple damage sites, frequency of readings etc is easy to assess. The approach is similar to the digital twin concept [339,340], and has been validated on data obtained in the blind trial of a guided wave pipe monitoring system [291] by Heinlein et al. [341]. The work of [338,341] used simple superposition of the predicted scattering from defects onto the measured signals from an undamaged structure; this approach neglects any multiple scattering of the defect signal and so becomes less accurate downstream of the defect; it is possible to model the whole system to overcome this, but at a considerable extra computational burden.

Crucially, this approach can also be used to validate the performance of an installed SHM system. Suppose that the regulator requires assurance that the SHM system installed on a nuclear power plant is correct in indicating that no defect larger than a specified size has grown at a particular location. Signals from defect growth at the specified location can be predicted and added to the sequence of raw signals measured from the installed SHM system; the resulting synthetic signal sequence can then be processed by the algorithms used to determine the health of the system, so verifying whether the required defect would be detected if present. The statistics of the measured signals can also be checked to ensure that any assumptions about, for example, normality of residuals, as discussed in section 8.2.1, are valid.

9. Conclusions

The use of guided waves in long range nondestructive testing is relatively mature, and largely limited to simple structures in applications that are difficult or impossible to address by conventional point inspection methods. Point methods are impractical in most large area SHM applications where it is necessary to detect isolated defects since it is not possible to cover the whole area of interest with transducers. Guided waves therefore have great potential in Structural Health Monitoring as they enable a large area of structure to be monitored from a limited number of transduction locations, so reducing the added weight, cost and complexity of the attachment to the structure. It is also possible to improve the sensitivity of the test relative to that obtained in one-off inspection by monitoring the change in signals with time using baseline subtraction. Current deployments are largely in the same type of structures as those where guided wave NDT is used, but there is substantial potential for increasing the range of applications.

9.1. Nondestructive testing

The main industrial applications of guided waves in NDT are in simple structures that are difficult to test using conventional inspection because they are not easily accessible, such as un-piggable pipes that are buried or insulated, and the foot of railway lines at level crossings. Guided wave NDT is also used for the detection of corrosion under pipe supports (CUPS) and in applications where it offers a great reduction in scanning requirement such as the detection of longitudinal cracking in gas pipelines using transducers mounted on a PIG, and the inspection of the whole rim of a railway wheel in a single test.

Bulk ultrasonic wave scanning technology has advanced greatly since the early guided wave work:

- The instrumentation and transduction required for simultaneous scanning on parallel channels in an array is now relatively cheap, as is the required data storage. Thus, the guided wave inspection of heat exchanger tubes has found it difficult to compete with systems such as IRIS [342]. Likewise, array scanners [343] have reduced the attraction of guided waves for large area plate scanning.
- Advances in robotics have made it easier and more reliable to access difficult areas.
- Rail inspection trains using multiple bulk wave transducers are now very advanced and reduce the attraction of guided wave inspection (as opposed to health monitoring) except for foot corrosion that is a particular issue at level crossings.
- Magnetic and ultrasonic floor scanners make it difficult for guided waves to compete in storage tank inspection unless they enable all the testing to be done from outside without opening the tank, and this usually limits the inspection to the annular ring adjacent to the wall.

Therefore, many of the early applications developed pre-1990 have not advanced and there is a limited number of cases where the reduced sensitivity compared with local testing using bulk wave ultrasound or other methods is acceptable. A further difficulty is that the design of test procedures is more complicated with guided waves than bulk waves, and there is much less off-the-shelf transduction and instrumentation available. The sensitivity and reliability of long range guided wave inspection on complex structures is unlikely to be sufficient to make the method competitive.

9.2. Structural health monitoring

The main current commercial applications of guided wave monitoring are in pipes and rail that are relatively simple structures where in many cases the critical defect size that must be detected is relatively large, for example a corrosion patch removing over 30 % of the wall thickness in a pipe. Progress has been much slower in other structures such as airframes that have attracted a great deal of research interest, but with limited pull through into applications. A key issue here is that the sensor spacing required to obtain adequate sensitivity on a complex structure can be ~100 mm, so a large number of sensors and associated wiring is needed to monitor a large area, resulting in a significant weight penalty; it can therefore be more cost effective to strengthen the structure to make it more damage tolerant. Hence, it is essential to design SHM systems with a clear business case [344], and if this is done then SHM applications have the potential to expand significantly as a result of future research and development effort.

9.3. Research needs

The transduction, wave propagation prediction and damage interaction modelling tools required to develop a guided wave NDT test procedure are readily available, making it possible to design an inspection for a new application as the need arises, as was done, for example, for an aero-engine component by Corcoran et al. [345]. However, signal interpretation remains a significant issue and increasing the test reliability by easing the demands on the test technician would significantly increase the attraction of guided wave inspection. Machine learning has great promise here, though it is a bigger challenge in one-off inspection, where the signal has to be interpreted with limited prior knowledge, than in SHM where the current signal can be compared with a baseline signal on the same structure.

There are more extensive research needs in guided wave SHM where the key is to develop systems for real applications, not unrealistic model structures such as plain plates. Key issues to improve are:

- The long-term stability of signals in realistic environments – the sensitivity of monitoring is dependent on the signal changes with time in the absence of damage being smaller than those due to the damage that must be detected. This will be affected by the stability of the transducers and their attachment to the structure, as well as that of the transmission and receiver electronics.
- Compensation for changes in environmental and operational conditions; great progress has been made with temperature compensation, but load changes and benign structural changes such as deposit formation or ageing of coatings are much more difficult to address.
- Dealing with the signals obtained in complex structures such as airframes where individual reflections are unlikely to be resolved, making baseline subtraction very demanding.

Machine learning is showing great promise in interpreting signal changes in long term tests, and this is potentially a very fruitful research area. It is essential to combine high sensitivity to defects with a very low false call rate and this is very challenging if the high POD is to be maintained [290]. Addressing this problem means it is essential for the research to be on realistic structures and for data to be collected over an extended period in realistic environments.

If the sensitivity of the test at a given sensor spacing can be increased then there is the possibility of reducing the sensor density, and hence the cost and complexity, while maintaining detection performance; alternatively, the improved sensitivity can be used to expand the range of applications to cases where detection performance close to that of local bulk wave inspections is required.

CRediT authorship contribution statement

Peter Cawley: Conceptualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Peter Cawley reports a relationship with Guided Ultrasonics Ltd that includes: board membership and equity or stocks.

Data availability

No data was used for the research described in the article.

Acknowledgements

The author is grateful to Professors Mike Lowe and Bruce Drinkwater for helpful comments on the first draft of the manuscript and also to the referees for making some helpful suggestions.

References

[1] Worlton DC. Ultrasonic testing with Lamb wave. Nondestruct Test 1957;15:218–22.
[2] Worlton DC. Experimental confirmation of Lamb waves at megacycle frequencies. J Appl Phys 1961;32:967–71.
[3] Rose JL. Ultrasonic waves in solid media. New York: Cambridge University Press; 1999.
[4] Rose JL. Ultrasonic guided waves in solid media. New York: Cambridge University Press; 2014.
[5] Cawley P. Practical long range guided wave inspection - managing complexity. In: Review of progress in quantitative NDE; 2003.
[6] Chimenti DE, Martin RW. Nondestructive evaluation of composite laminates by leaky Lamb waves. Ultrasonics 1991;29:13–21.
[7] Briggs GA. Acoustic microscopy. Oxford: Clarendon Press; 1992.
[8] Clark AV, Thompson RB, Li Y, Reno RC, Blessing GV, Mitrakovic DV. Ultrasonic measurement of sheet steel texture and formability: comparison with neutron diffraction and mechanical measurements. Res Nondestr Eval 1990;2:239–57.