Efficient, Selective Piezoeletric Wave Transduction Using Interdigitated Electrodes: COM Versus COMSOL Multiphysics®

Harry T.D. Grigg[*1],Thomas H. Hanley[*1], and Barry J. Gallacher[1]
[1]Newcastle University,
*School of Mechanical and Systems Engineering, Stephenson Building, Claremont Road, Newcastle, Tyne and Wear, UK NE1 7RU

Abstract: The controlled transduction of elastic waves is a prerequisite for the development of elastic resonators. One of the most common and powerful designs in this field is the interdigitated transducer, which is both versatile and easy to fabricate using lithographic techniques and piezoelectric materials. The resulting devices will, in general, produce a whole gamut of wave types. In order to maximize energy transfer into the desired wave modes while suppressing spurious components as required by a particular design, a common approximation is to treat the waves as unperturbed by the transducer grating using a Coupling of Modes (COM) approach. Here, we use the Structural Mechanics and Piezoelectric Devices interfaces to study the frequency reponse as calculated by COMSOL, and compare the results to the more traditional COM methods.

Keywords: SAW, transducers, waves, MEMS, biosensors.

1. Introduction

Interdigitated Transducers (IDTs) have been used since the 1960s to generate surface and bulk acoustic waves for a variety of applications[1]. Their popularity is due in part to versatility, ease and low cost of manufacture, and suitability for mass production. IDTs find application to a wide variety of device classes where a requirement exists for the generation of elastic waves in solids. Different transducer geometries have been used, both alone and in combination with coupling gratings, to produce waves including bulk transverse horizontal, shear vertical, and longitudinal waves; surface Rayleigh waves, Love waves, Bleustein-Gulyaev waves, Sezawa waves, Stoneley waves, and others. The present work will focus on their use in Rayleigh wavedevices, but the principles enunciated generalize easily.

1.1 Surface Acoustic Wave types

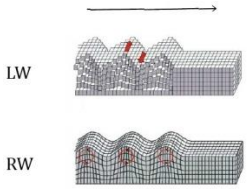

Figure 1. Graphic showing some of the various elastic surface wave types: Love waves and Rayleigh waves, respectively.

Surface Acoustic Waves (SAW) is a term referring here to elastic waves propagating in a (possibly inhomogeneous, anisotropic or layered) solid, distinguished from bulk waves by an evanescent behavior in the direction normal to the surface.

The simplest surface waves are Rayleigh waves (RW), which can exist in an isotropic, homogeneous solid with a free surface, provided that the wavelengths are much smaller than the dimensions transverse to the surface. In this case, plane waves are symmetric under translation in the direction normal to the sagittal plane and exhibit no dispersion. However, in layered media, the participation of the various media changes as the wavelength varies, introducing dispersion. For the same reason, periodic surface conditions such as those introduced by the deposition of IDTs also introduce nontrivial dispersion and attendant phase effects to the problem.

1.2 Biosensors and Surface Acoustic Waves

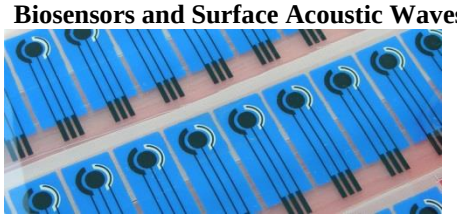

Figure 2. Image of an electrochemical biosensor array, patterned using lithographic techniques. The ability to mass-produce high performance portable transducers capable of rapid sample analysis at low cost is set to change the fundamentals of healthcare in the coming decade.

Biosensors have received an increasing level of attention in the literature during the past 20 years[2], due to their attractive applications in point-of-care medicine and patient self-testing. Although many physical principles are exploited, including electrochemistry, plasmon resonance, etc., elastic wave biosensors are particularly promising for specific subfields. The principle of operation is based on some coupling of an elastic wave to the concentration of a biological species, typically introduced to the vicinity of the propagating medium in the form of a solution such as a blood or plasma sample. The modification of the wave properties is then combined with an understanding of the device dynamics to infer the concentration of the analyte species.

Surface waves have particular advantages here[3]. Due to the localization of the mechanical energy near to a solid interface, the change in the observable wave properties pursuant to a modification of the surface by the analyte coupling is larger than for bulk wave devices. The higher the frequency of the surface waves, the more localized the energy becomes and the greater the sensitivity can be made. The ability to detect very low analyte concentrations at high precision in a cheap, mass-produced device is extremely attractive.

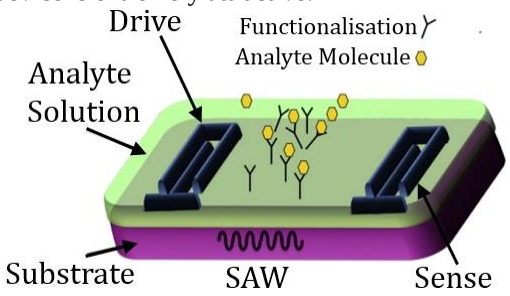

Figure 3. Operating principle of a SAW biosensor. The SAW could make a single timed journey from the Drive IDT to the Sense IDT, or be reflected at the boundaries and the resulting frequency response characterized, as possible sensing mechanisms.

The coupling is typically achieved by selective immobilization of a chemical monolayer with specific affinity for the analyte molecule to a gold layer in the SAW path defined between a drive IDT, which excites the SAW, and a sense IDT, which picks off the quantities of interest. If the waves are reflected by a carefully designed periodic array, then the sensitivity of the devices can be substantially augmented by the resonant sensing principle. However, to control the modal content of the resonance adequately, it is necessary to have an accurate prediction of the phase of the generated waves from the drive IDTs, i.e. the complex phase of the input impedance. Examining fine corrections to this quantity is the thrust of this paper.

1.3 Function and Structure of IDT

An IDT structure consists of interdigitated electrodes patterned in metal, usually gold, via photolithography or other means on a piezoelectrically active substrate or layer. It is therefore a quasi-two dimensional geometry, although the out-of-plane field components cannot be neglected in general. Two bus tracks are deposited parallel to the desired direction of wave propagation. Electrodes are then arranged transverse to the buses in the gap between them, such that, in the simplest case, alternate electrodes are shorted to alternate buses, with a short gap isolating them from the other bus. The pattern thus formed resembles the interlocked digits of clasped hands, leading to the transverse electrodes being termed digits. This arrangement creates two topologically disconnected regions which are addressed separately and excited with an AC voltage. An complex pattern of AC electric fields is thus induced, which interacts with the substrate via the piezoelectric effect to excite wave propagation.

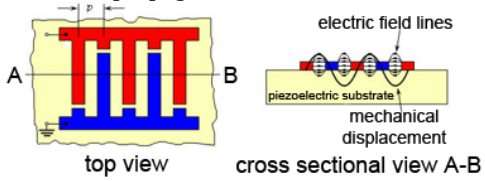

Figure 4. Operating principle of an IDT. The disconnected regions, illustrated in blue and red, are excited with AC voltages of equal amplitude and frequency but opposing phase. This gives rise to an AC electric field pattern of the form illustrated on the right.

In Figure 4, it can be seen that each digit does not reach close proximity of the opposing bus, but instead is separated from a short "dummy" digit extending from it by a small gap. This configuration, which is common design practice, has at least two advantages. One is the ability to vary the width of the excitation region, which extends between the lines defined by the gaps,

without requiring curved bus geometries. This is useful for shaping the frequency response of the sensor as the application dictates, particularly in SAW filters for signal processing applications. The second is to prevent anomalous dispersion distorting plane waves propagating along the transducer, which would reduce sensitivity and introduce unwanted distortion to the designed phase response.

The origins of the anomalous dispersion are in the boundary conditions presented to the substrate by the metallised regions. These can be modelled as an electrical component, arising from the shorting of the surface where gold is deposited, and a mechanical loading due to the presence of deposited mass and associated stiffness. The prevalence of the dummy digit in practical design, even when apodisation is not required, is a reminder of the importance of carefully controlling the dispersion of the surface waves of interest in the vicinity of the transducers for optimal design control.

2.Physics and governing equations

The governing equations for the electrostatics and time-harmonic forced mechanical dynamics of piezoelectric media can be written in the respective forms:

$$_{\mathrm{ij,j}} - \rho_{\mathrm{v}} = 0 \ (\mathbf{1})$$

And

$$\rho\omega^2 u_i + \sigma_{ij,j} + \mathrm{F}_i e^{i\varphi} = 0 \ (\mathbf{2})$$

Where **D** is the electric displacement field, ρ_{v} is the charge density, ρ is the material density, **u** is the mechanical displacement field, ω and φ are the excitation frequency and phase respectively, $\boldsymbol{\sigma}$ is the mechanical stress tensor, and $\mathbf{F}_v$ is the external force density. Due to the inherent anisotropy in the piezoelectric media in the problem and the zero initial values in the reference configuration, we introduce the linearised constitutive relations of piezoelectricity for the stress and displacement respectively as:

$$s_{ik} - C_{ijkl}\epsilon_{jl} + e^T_{kij}E_k = 0 \ \ (\mathbf{3})$$

And

$$_i - e_{ijk}\epsilon_{jk} + \varepsilon_{ij}E_j = 0 \ \ (\mathbf{4})$$

Where s_{ik} is the stress tensor, C_{ijkl} is the fourth order elasticity tensor, ϵ_{jl} is the strain tensor, e_{ijk} is the third order piezoelectric coupling tensor, E_k is the electric field, D_i is the displacement field, and ε_{ij} is the anisotropic permittivity. It can be inferred that, from an initial rest condition, instantaneous introduction of voltage boundary conditions will lead to a nonzero electric field on an electromagnetic timescale, leading to nonzero stresses via the coupling tensor, which in turn act as a mechanical forcing on the mechanical timescale.

3. Elastic Radiation/COM model

In order to reduce the complexity of the full problem, and given that the electromagnetic field problem is essentially electrostatic due to the great disparity in mechanical and electromagnetic velocities as reflected in the chosen dynamical equations, the full field solution and resulting stress distribution is often replaced by a representative mechanical surface force.

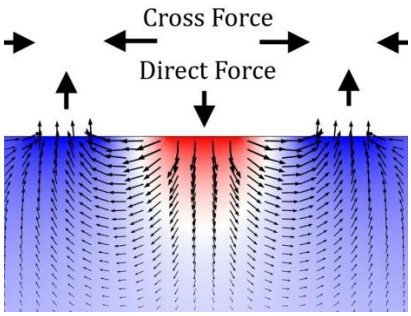

Figure 5. COMSOL model of electric field distribution and effective lumped "cross" and "direct" force terms for a Y-polarised PZT sample used in the COM model. Red and blue denote positive and negative voltages, respectively.

In Figure 5, the electric field distribution arising from a static IDT voltage is complicated and not available in analytical closed form. A corresponding harmonic forcing arises from the interaction between the anisotropic substrate constitutive properties and the field, according to (**1-4**). For design purposes, often the full field is reduced to its dominant components. For out-of-plane polarization in the common materials, these can be considered as an out of plane "direct" component centred on the digits and an in plane "cross" traction, as illustrated.

Here, we consider the effect of the cross terms arising from two digits in order to draw more general insight into the nature of the interactions before looking at the complete problem. The governing equations are reduced to those of

planar two dimensional isotropic elasticity, namely

$$\rho\omega^2 u_j - C_{11}(u_{j,jj}) - C_{44}(u_{j,ij} + u_{i,ji}) = 0 \quad (5)$$

And the time-harmonic boundary conditions are illustrated in Figure 6:

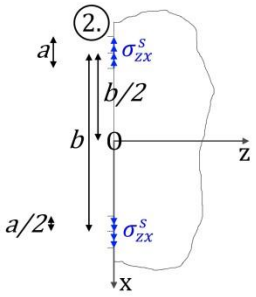

Figure 6. Schematic of 2D isotropic elasticity model of the crossed-field problem.

An integral expression for these quantities was derived in [4] and solved numerically in MATLAB, and the corresponding problem solved in COMSOL using the Structural Mechanics module. A PML was used to prevent spurious effects associated with boundary reflection. The mesh and geometry are illustrated in Figure 7.

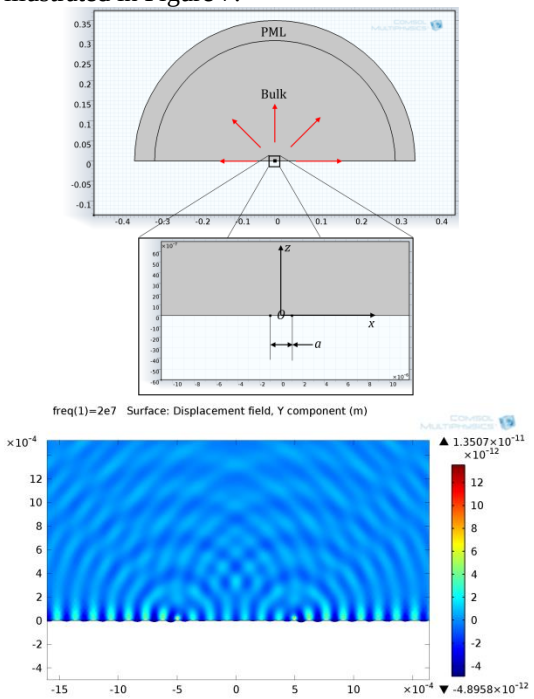

Figure 7. COMSOL geometry used to solve the elastic wave radiation problem for the double mechanical stress source, approximating the crossed-field IDT(top), and a typical resulting displacement pattern(bottom).

It can be observed in Figure 6 that a complex diffraction pattern emerges. The integral formulation in [4], which is the first solution of the elastic double slit problem in the literature, shows that P and S waves are generated, as well as a surface Rayleigh wave.

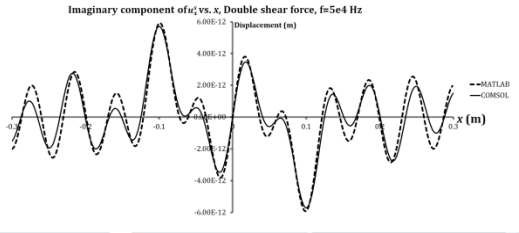

Figure 8. Comparison of COMSOL elastic model and integral form expression for displacements in crossed field model solved in MATLAB.

By applying a known force in both the integral formulation and COMSOL with matched geometry and material properties, a validating comparison can be drawn by plotting out the resulting displacements at constant phase along the free surface; the result of doing so is shown in Figure 8. Since Rayleigh, P, and S waves with differing spatial components are present, the excellent agreement shown gives confidence that the underlying physics is being replicated well. Next, we turn to the numerical components of displacement averaged over the digits, as given by the validated integral form.

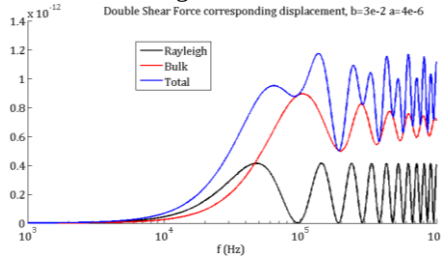

Figure 9. Lumped mechanical acceptances corresponding to the crossed field model of a two-finger IDT, as calculated in MATLAB.

Figure 9 shows the reciprocal of the lumped mechanical port impedances predicted by the model, over a wide frequency range. It can be seen that the Rayleigh wave component oscillates between zero and a positive maximum. The zeroes correspond to infinite impedance; no wave is transduced; the maxima, to a minimum

impedance. This result can be interpreted physically in terms of constructive and destructive interference of the Rayleigh waves with the sources, as the fixed digit separation represents a varying number of wavelengths for the propagating wave at differing frequencies. This insight underlies the physical principle governing the frequency response structure of IDTs in general.

4. Interdigitated Transducer Model

The mechanical model of the previous sector makes no account of several important factors in practical IDT design, including material anisotropy, the mass loading associated with the electrode metallization, and the electrical boundary condition changes and associated variation in wave impedance. Accessing these results analytically would be very involved. To assess the magnitude of these issues, a COMSOL model was developed using the Piezoelectric Devices (PZD) physics interface and a Frequency Domain structure, in conjunction with a double-IDT geometric configuration and Cartesian PMLs surrounding a propagation channel.

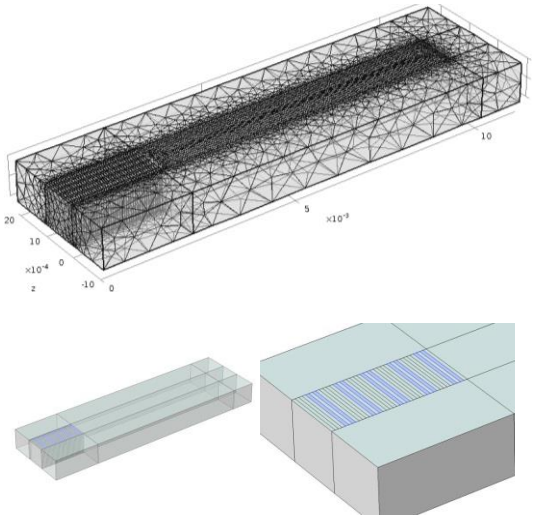

Figure 10. COMSOL mesh(top) and geometry used to solve the full piezoelectric IDT problem: overview(bottom left) showing transducer region, propagation channel and Cartesian PML domains, and closeup of transducer zone(bottom right), highlighting double digit IDT structure.

The mesh and geometry are exhibited in Figure 10. To capture the fast variation near the surface associated with the RW within computational constraints, the propagating surface was first meshed with a Free Triangular structure and a custom element size. Next, the transducer region was meshed with a Free Tetrahedral structure, tuned to provide good resolution of the electromechanical fields in the vicinity of the IDTs. Finally, the remaining domains were automatically meshed, allowing for element growth away from the sensitive regions. The resulting element number was approximately 7×10^4 , allowing the problem to be solved for 300 parameter combinations in 6 hours on a 2 core Xeon machine with 12GB ram.

Piezoelectric constitutive relations were used throughout, with the outer regions forming an absorbing layer matched via the inbuilt PML functionality. All bottom surfaces were fixed to approximate the surface wave form, while all top and side surfaces were free, with the exception of the top transducer surface. Here, a double IDT pattern was defined by voltages on every other rectangular region. It can be seen in Figure 10 that the voltage alternates every second transition, rather than every first, as would be the case for single IDTs.

In addition, an Added Mass boundary condition and a Thin Elastic Layer boundary condition were defined for the metallised regions. The strength of each was modulated by a parameter $rhoadd$, corresponding to the added mass and hence thickness of the metallization. In this manner, by running a parametric sweep over three values of $rhoadd$ corresponding to 200 nm and 2 um of gold on a PZT-5A substrate, the mass and stiffness effect of varying the deposition thickness on the resulting displacement fields could be explored.

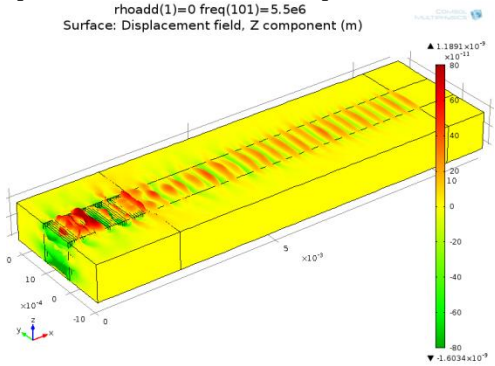

Figure 11. COMSOL geometry used to solve the elastic wave radiation problem for the double mechanical stress source, approximating the crossed-

field IDT(top), and a typical resulting distribution pattern.

A consistent and stable displacement pattern was obtained from the model, as shown in Figure 11. It agreed quite well with the form given by a corresponding modification of the elastic model when the variation in wavelength engendered by the piezoelectricity-induced change in wave speed was properly accounted for, when the condition $rhoadd = 0$ was enforced. However, significant differences could be observed when the added mass was simulated.

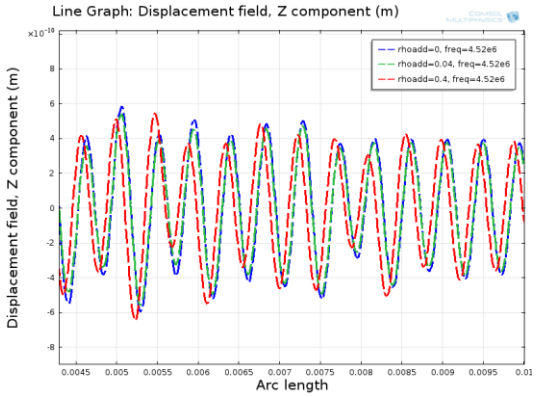

Figure 12. Transverse displacement at the free surface in the propagation channel, for three values of the parameter $rhoadd$. Note that the relative amplitudes of the peaks are significantly changed, and that the peaks are progressively retarded, as the parameter increases from 0.

Figure 12 is a representative result showing the changes in the surface displacements engendered by changing the added mass parameter. One notable conclusion from the observed variation in the relative peak amplitudes is that the relative contributions of P, S, and Rayleigh waves is influenced by the addition of mass. This has important implications when the goal is to selectively excite one wave type only; the results of using the COM model to eliminate bulk modes in design would not capture this information and hence lead to a suboptimal design and the generation of deleterious wave components. Furthermore, as mass is added, progressive phase retardation occurs. This is particularly significant when designing phased reflector arrays in SAW resonator design, where such errors would lead to the introduction of undesirable mode components and contamination of the desired output signal. e

5. Conclusions

The standard approach to IDT design based on Coupling of Modes is validated as an acceptable approximation for many purposes. However, the results of this study make clear that, when output phase is of importance, the electrical and mechanical loadings imposed by the IDT structure cannot be neglected. It is further shown that for high phase precision applications, COMSOL Multiphysics provides a rapid, efficient, and accurate design tool capable of yielding information about perturbations from anisotropy, field distribution, and surface loading not captured by the standard approach.

6. References

1. A.A. Oliner et al., *Acoustic Surface Waves*. Springer-Verlag, Berlin (1978)

2. A.P. Turner, Biosensors: sense and sensibility, *Chem Soc Reviews*, **42(8)**, 3184-3196 (2013)

3. K. Lange et al., Surface acoustic wave biosensors: A review, *Anal Bioanal Chem*, **391(5),** 1509-1519 (2008)

4. H.T.D Grigg, *Priciples and Practice of the Xylophone Bar Magnetometer*, PhD Thesis, University of Newcastle, 2014. Available online: 2015 (pending publications)

7. Acknowledgements

Many thanks to EPSRC for supporting this work, to my mentor Dr. Barry Gallacher, and to Professor James Burdess at Newcastle for their time and insight into the problems discussed here.