IntechOpen

Acoustic Waves

Edited by Don Dissanayake

WEB OF SCIENCE™

Acoustic Waves

edited by
Don W. Dissanayake

Acoustic Waves
http://dx.doi.org/10.5772/277
Edited by Don Dissanayake

First published in Croatia, 2010 by INTECH d.o.o.
eBook (PDF) Published by IN TECH d.o.o.
Place and year of publication of eBook (PDF): Rijeka, 2019. IntechOpen is the global imprint of IN TECH d.o.o.
Printed in Croatia

Legal deposit, Croatia: National and University Library in Zagreb

Additional hard and PDF copies can be obtained from orders@intechopen.com

Acoustic Waves
Edited by Don Dissanayake

p. cm.

ISBN 978-953-307-111-4

eBook (PDF) ISBN 978-953-51-4543-1

Contents

Preface

Surface Acoustic Wave (SAW) devices are widely used in multitude of device concepts mainly in Micro Electro Mechanical Systems (MEMS) and communication electronics. As such, SAW based micro-sensors, actuators and communication electronic devices are well known applications of SAW technology. Due to their solid state design and fabrication compatible with other modern technologies such as Microwave Integrated Circuits (MIC), MEMS, (Charge Coupled Devices) CCD and integrated optic devices, SAW based sensors are considered to be extremely reliable. For example, SAW based passive micro sensors are capable of measuring physical properties such as temperature, pressure, variation in chemical properties, and SAW based communication devices perform a range of signal processing functions, such as delay lines, filters, resonators, pulse compressors, and convolvers. In recent decades, SAW based low-powered actuators and microfluidic devices have significantly contributed towards their popularity.

SAW devices are based on propagation of acoustic waves in elastic solids and the coupling of these waves to electric charge signals via an input and an output Inter Digital Transducers (IDT) that are deposited on the piezoelectric substrate. Since the introduction of the first SAW devices in the 1960s, this flexibility has facilitated a great level of creativity in the design of different types of devices, which has resulted in low cost mass production alongside with modern electronic, biomedical and similar systems.

In recent times, SAW devices have become an indispensable part of the modern electronic communication industry due to their usefulness as IF, RF, and GPS filters for various applications. Over the years, SAW devices are known to offer superior performance in communication due to a range of factors such as high stability, excellent aging properties, low insertion attenuation, high stopband rejection and processing gain, and narrow transition width from passband to the stopband. Therefore, it is evident that SAW based wireless communication is a well-established field in RF–MEMS and Bio–MEMS devices, and has a great potential to incorporate with modern biosensors, micro actuators and biological implants. Furthermore, passive SAW sensors can be RF controlled wirelessly through a transceiver unit over distances of several meters, without the need of a battery. Hence, such devices are well suited for use in a wide range of sensor and identification systems.

This book consists of 20 exciting chapters composed by researchers and engineers active in the field of SAW technology, biomedical and other related engineering disciplines. The topics range from basic SAW theory, materials and phenomena to advanced applications such as sensors actuators, and communication systems. As such, first part of this book is dedicated to several chapters that present the theoretical analysis and numerical modelling such as Finite Element Modelling (FEM) and Finite Difference Methods (FDM) of SAW devices. Then, some exciting research contributions in SAW based actuators and micro motors are presented in

part two of this book. In part three, a fruitful collection of research outcomes that are based on SAW sensing applications are presented.

We are excited that this collection of up-to-date information and research outcomes on SAW technology will be of great interest, not only to all those working in SAW based technology, but also to many more who stand to benefit from an insight into the rich opportunities that this technology has to offer, especially to develop advanced, low-powered biomedical implants and passive communication devices.

Editor

Don W. Dissanayake,
The School of Electrical and Electronic Engineering,
The University of Adelaide,
Australia

1

The Eigen Theory of Waves in Piezoelectric Solids

Shaohua Guo
Zhejiang University of Science and Technology
P.R.China

1. Introduction

It is difficult to answer how many elastic waves exist in anisotropic solids. The traditional viewpoint believes that there are only two bulk elastic waves in solids, one is the dilation wave discovered by Poisson in 1892, and the other the shear wave discovered by Stokes in 1899. The existence of the P-wave and the S-wave was also verified by the classical elastic theory. However, with the discovery of some new phenomena of elastic waves in anisotropic solids, it is found that the limitations of classical elastic theory have become obvious. Furthermore, the current concepts and theories of elastic waves can not answer several basic questions of elastic wave propagation in anisotropic solids. For example, how many elastic waves are there? How many wave types are there? What is the space pattern of elastic waves? As we know, the Christoffel's equation, which is often used to describe anisotropic elastic waves in the classical elastic theory, can not indicate the space pattern and the complete picture of elastic wave propagation in anisotropic solids, but only show the difference of propagation in the different directions along an axis or a section (Vavrycuk, 2005). The reason for this is that the classical elastic wave equations, expressed by displacements can not distinguish the different elastic sub-waves (except for isotropic solids), because the elasticity and anisotropy of solids are synthesized in an elastic matrix. Similarly, for the electromagnetic fields, except for the Helmholtz's equation of electromagnetic waves in isotropic media, the laws of propagation of electromagnetic waves in anisotropic media are also not clear to us. From the Maxwell's equation, the explicit equations of electromagnetic waves in anisotropic media could not be obtained because the dielectric permittivity matrix and magnetic permeability matrix were all included in these equations, so that only local behaviour of electromagnetic waves, for example, in a certain plane or along a certain direction, can be studied (Yakhno et al., 2006).

The theory of linear piezoelectricity is based on a quasi-static approximation (Tiersten et al., 1962). In this theory, although the mechanical equations are dynamic, the electromagnetic equations are static and the electric field and the magnetic field are not coupled. Therefore it does not describe the wave behaviour of electromagnetic fields. Electromagnetic waves generated by mechanical fields (Mindlin, 1972) need to be studied in the calculation of radiated electromagnetic power from a vibrating piezoelectric device (Lee et al., 1990), and are also relevant in acoustic delay lines (Palfreeman, 1965) and wireless acoustic wave sensors (Sedov et al., 1986), where acoustic waves produce electromagnetic waves or vise versa. When electromagnetic waves are involved, the complete set of Maxwell equation needs to be used,

coupled to the mechanical equations of motion. Such a fully dynamic theory is called piezoelectromagnetism by some researchers (Lee, 1991). Piezoelectromagnetic SH waves were studied by Li (Li, 1996) using scalar and vector potentials, which results in a relatively complicated mathematical model of four equations. Two of these equations are coupled, and the other two are one-way coupled. In addition, a gauge condition needs to be imposed. A different formulation was given by Yang and Guo (Yang et al., 2006), which leads to two uncoupled equations. Piezoelectromagnetic SH waves over the surface of a circular cylinder of polarized ceramics were analyzed. Although many works have been done for the piezoelectromagnetic waves in piezoelectric solids, the explicit uncoupled equations of piezoelectromagnetic waves in the anisotropic media could not be obtained because of the limitations of classical theory. In this chapter, the idea of eigen theory presented by author (Guo, 1999; 2000; 2001; 2002; 2005; 2007; 2009; 2009; 2010; 2010; 2010) is used to deal with both the Maxwell's electromagnetic equation and the Newton's motion equation. By this method, the classical Maxwell's equation and Newton's equation under the geometric presentation can be transformed into the eigen Maxwell's equation and Newton's equation under the physical presentation. The former is in the form of vector and the latter is in the form of scalar. As a result, a set of uncoupled modal equations of electromagnetic waves and elastic waves are obtained, each of which shows the existence of electromagnetic and elastic sub-waves, meanwhile the propagation velocity, propagation direction, polarization direction and space pattern of these sub-waves can be completely determined by the modal equations.

In section 2, the elastic waves in anisotropic solids were studied under six dimensional eigen spaces. It was found that the equations of elastic waves can be uncoupled into the modal equations, which represent the various types of elastic sub-waves respectively. In section 3, the Maxwell's equations are studied based on the eigen spaces of the physical presentation, and the modal electromagnetic wave equations in anisotropic media are deduced. In section 4, the quasi-static theory of waves in piezoelectric solids (mechanical equations of motion, coupled to the equations of static electric field, or Maxwell's equations, coupled to the mechanical equations of equilibrium) are studied based on the eigen spaces of the physical presentation. The complete sets of uncoupled elastic or electromagnetic dynamic equations for piezoelectric solids are deduced. In section 5, the Maxwell's equations, coupled to the mechanical equations of motion, are studied based on the eigen spaces of the physical presentation. The complete sets of uncoupled fully dynamic equations for piezoelectromagnetic waves in anisotropic media are deduced, in which the equations of electromagnetic waves and elastic ones are both of order 4. The discussions are given in section 6.

2. Elastic waves in anisotropic solids

2.1 Concepts of eigen spaces

The eigen value problem of elastic mechanics can be written as

$$\boldsymbol{C}\boldsymbol{\varphi}_i = \lambda_i \boldsymbol{\varphi}_i \qquad i = 1,2,\cdots,6 \tag{1}$$

where $\boldsymbol{C}$ is a standarded matrix of elastic coefficients, λ_i is eigen elasticity, and is invariables of coordinates, $\boldsymbol{\varphi}_i$ is the corresponding eigen vector, and satisfies the orthogonality condition of basic vectors.

$$\boldsymbol{C} = \boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}^T \tag{2}$$

where $\boldsymbol{\Lambda} = diag\left[\lambda_1, \lambda_2, \cdots, \lambda_6\right]$, $\boldsymbol{\Phi} = \left\{\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \cdots, \boldsymbol{\varphi}_6\right\}$ is the modal matrix of elastic solids, it is orthogonal and positive definite one, satisfies $\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi} = I$.
The eigen spaces of anisotropic elastic solids consist of independent eigen vectors, it has the structure as follows

$$W = W_1[\boldsymbol{\varphi}_1^*] \oplus \cdots \oplus W_m[\boldsymbol{\varphi}_m^*] \tag{3}$$

where the possible overlapping roots are considered, and $m(\leq 6)$ is used to represent the number of independent eigen spaces. Projecting the stress vector $\boldsymbol{\sigma}$ and strain vector $\boldsymbol{\varepsilon}$ on the eigen spaces, we get

$$\boldsymbol{\sigma} = \sigma_1^* \boldsymbol{\varphi}_1^* + \cdots + \sigma_m^* \boldsymbol{\varphi}_m^* \tag{4}$$

$$\boldsymbol{\varepsilon} = \varepsilon_1^* \boldsymbol{\varphi}_1^* + \cdots + \varepsilon_m^* \boldsymbol{\varphi}_m^* \tag{5}$$

where σ_i^* and ε_i^* are modal stress and modal strain, which are stress and strain under the eigen spaces respectively, and are different from the traditional ones in the physical meaning. Eqs.(4) and (5) are also regarded as a result of the sum of finite number of normal modes.
The modal stress and modal strain satisfy the normal Hook's law

$$\sigma_i^* = \lambda_i \varepsilon_i^* \qquad i = 1, 2, \cdots, m \tag{6}$$

2.2 Modal elastic wave equations

When neglecting body force, the dynamics equation and displacement equation of elastic solids are the following respectively

$$\sigma_{ik'k} = \rho \ddot{u}_i \tag{7}$$

$$\varepsilon_{ij} = \frac{1}{2}(u_{i'j} + u_{j'i}) \tag{8}$$

From Eqs.(7) and (8), we can get the following equation

$$\sigma_{ik'kj} + \sigma_{jk'ki} = 2\rho \ddot{\varepsilon}_{ij} \tag{9}$$

Because of the symmetry on (i, j) in Eq.(9) we can rewrite it in the form of matrix.

$$\Delta\boldsymbol{\sigma} = \rho \Delta_{tt} \boldsymbol{\varepsilon} \tag{10}$$

where

$$\boldsymbol{\Delta} = \begin{bmatrix} \partial_{11} & 0 & 0 & 0 & \partial_{31} & \partial_{21} \\ 0 & \partial_{22} & 0 & \partial_{32} & 0 & \partial_{21} \\ 0 & 0 & \partial_{33} & \partial_{32} & \partial_{31} & 0 \\ 0 & \partial_{23} & \partial_{23} & (\partial_{22} + \partial_{33}) & \partial_{21} & \partial_{31} \\ \partial_{13} & 0 & \partial_{13} & \partial_{12} & (\partial_{11} + \partial_{33}) & \partial_{32} \\ \partial_{12} & \partial_{12} & 0 & \partial_{13} & \partial_{23} & (\partial_{22} + \partial_{11}) \end{bmatrix} \tag{11}$$

It is seen that $\boldsymbol{\Delta}$ is a symmetrical differential operator matrix, and $\partial_{ij} = \partial_{ji} = \partial^2 / \partial x_i \partial x_j$, $\Delta_{tt} = \partial^2 / \partial_t \partial_t$.

It is proved by author that the elastic dynamics equation (10) under the geometrical spaces of three dimension can be converted into the modal equations under the eigen spaces of six dimension

$$\Delta_i^* \lambda_i \varepsilon_i^* = \rho \Delta_{tt} \varepsilon_i^* \qquad i = 1,2,\cdots,m \tag{12}$$

and

$$\Delta_i^* = \boldsymbol{\varphi}_i^{*\mathrm{T}} \Delta \boldsymbol{\varphi}_i^* \qquad i = 1,2,\cdots,m \tag{13}$$

where Δ_i^* is called as the stress operator. From Eq.(12), we have

$$\Delta_i^* \varepsilon_i^* = \frac{1}{v_i^2} \nabla_{tt} \varepsilon_i^*, \quad i = 1,2,\cdots,m \tag{14}$$

The calculation shows that the stress operators are the same as Laplace's operator (either two dimensions or three dimensions) for isotropic solids, and for most of anisotropic solids. In the modal equations of elastic waves, the speeds of propagation of elastic waves are the following

$$v_i = \sqrt{\frac{\lambda_i}{\rho}}, \quad i = 1,2,\cdots,m \tag{15}$$

Eqs.(14) and (15) show that the number of elastic waves in anisotropic solids is equal to that of eigen spaces of anisotropic solids, and the speeds of propagation of elastic waves are related to the eigen elasticity of anisotropic solids.

2.3 Elastic waves in isotropic solids

There are two independent eigen spaces in isotropic solids

$$W = W_1^{(1)}[\boldsymbol{\varphi}_1] \oplus W_2^{(5)}[\boldsymbol{\varphi}_2,\cdots,\boldsymbol{\varphi}_6] \tag{16}$$

where

$$\left.\begin{aligned} \boldsymbol{\varphi}_1 &= \frac{\sqrt{3}}{3}[1,1,1,0,0,0]^T, \boldsymbol{\varphi}_2 = \frac{\sqrt{2}}{2}[0,1,-1,0,0,0]^T \\ \boldsymbol{\varphi}_3 &= \frac{\sqrt{6}}{6}[2,-1,-1,0,0,0]^T, \boldsymbol{\varphi}_i = \xi_i \quad (i = 4,5,6) \end{aligned}\right\} \tag{17}$$

where $\boldsymbol{\xi}_i$ is a vector of order 6, in which i th element is 1 and others are 0.

The eigen elasticity and eigen operator of isotropic solids are the following

$$\left.\begin{aligned} \lambda_1 &= 3(\lambda + 2\mu), \quad \lambda_2 = 2\mu, \\ \Delta_1^* &= \frac{1}{3}\nabla_{III}^2, \quad \Delta_2^* = \frac{1}{2}\nabla_{III}^2, \end{aligned}\right\} \tag{18}$$

where λ and μ are Lame constants, ∇_{III}^2 is Laplace's operator of three dimention. Thus, there exist two independent elastic waves in isotropic solids, which can be described by the following equations

$$(\lambda+2\mu)\nabla_{\mathrm{III}}^2\varepsilon_1^*(x,t)=\rho\ddot{\varepsilon}_1^*(x,t) \tag{19}$$

$$\mu\nabla_{\mathrm{III}}^2\varepsilon_2^*(x,t)=\rho\ddot{\varepsilon}_2^*(x,t) \tag{20}$$

It will be seen as follows that Eqs.(19) and (20) represent the dilation wave and shear wave respectively.
Using Eq. (5), the modal strain of order 1 of isotropic solids is

$$\varepsilon_1^*=\boldsymbol{\varphi}_1^{*\mathrm{T}}\cdot\boldsymbol{\varepsilon}=\frac{\sqrt{3}}{3}(\varepsilon_{11}+\varepsilon_{22}+\varepsilon_{33}) \tag{21}$$

Eq. (21) represents the relative change of the volume of elastic solids. So Eq. (19) shows the motion of pure longitudinal wave.
Also from Eq. (5), the modal strain of order 2 of isotropic solids is

$$\varepsilon_2^*\boldsymbol{\varphi}_2^*=\boldsymbol{\varepsilon}-\varepsilon_1^*\boldsymbol{\varphi}_1^* \tag{22}$$

By the orthogonality condition of eigenvectors, we have

$$\begin{aligned}\left|\varepsilon_2^*\right|&=[(\boldsymbol{\varepsilon}-\varepsilon_1^*\boldsymbol{\varphi}_1^*)^{\mathrm{T}}\cdot(\boldsymbol{\varepsilon}-\varepsilon_1^*\boldsymbol{\varphi}_1^*)]^{1/2}\\&=\{\frac{1}{3}[(\varepsilon_1-\varepsilon_2)^2+(\varepsilon_2-\varepsilon_3)^2+(\varepsilon_3-\varepsilon_1)^2]\}^{1/2}\end{aligned} \tag{23}$$

Eq. (23) represents the pure shear strain on elastic solids. So Eq. (20) shows the motion of pure transverse wave.

2.4 Elastic waves in anisotropic solids

2.4.1 Cubic solids

There are three independent eigen spaces in a cubic solids

$$W=W_1^{(1)}[\boldsymbol{\varphi}_1]\oplus W_2^{(2)}[\boldsymbol{\varphi}_2,\boldsymbol{\varphi}_3]\oplus W_3^{(3)}[\boldsymbol{\varphi}_4,\boldsymbol{\varphi}_5,\boldsymbol{\varphi}_6] \tag{24}$$

where $\boldsymbol{\varphi}_1,\boldsymbol{\varphi}_2,\cdots,\boldsymbol{\varphi}_6$ are the same as in Eq. (17).
The eigenelasticity and eigenoperator of cubic solids are

$$\left.\begin{aligned}&\lambda_1=c_{11}+2c_{12},\quad \lambda_2=c_{11}-c_{12},\quad \lambda_3=c_{44},\\&\Delta_1^*=\frac{1}{3}\nabla_{III}^2,\quad \Delta_2^*=\frac{1}{2}\nabla_{III}^2,\quad \Delta_3^*=\nabla_{III}^*,\end{aligned}\right\} \tag{25}$$

Thus, there exist three independent elastic waves in cubic solids, which can be described by following equations

$$\lambda_i\Delta_i^*\varepsilon_i^*(x,t)=\rho\ddot{\varepsilon}_i^*(x,t)\qquad i=1,2,3 \tag{26}$$

where

$$\varepsilon_1^{\ *} = \frac{\sqrt{3}}{3}(\varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33}) \tag{27}$$

$$\varepsilon_2^{\ *} = [\frac{1}{2}(\varepsilon_{22} - \varepsilon_{33})^2 + \frac{1}{6}(2\varepsilon_{11} - \varepsilon_{22} - \varepsilon_{33})^2]^{1/2} \tag{28}$$

$$\varepsilon_3^{\ *} = \sqrt{\frac{1}{2}(\varepsilon_{32}^{\ 2} - \varepsilon_{31}^{\ 2} - \varepsilon_{12}^{\ 2})} \tag{29}$$

It is seen that there exist three elastic waves in cubic solids, one of which is the quasi-dilation wave, and two others are the quasi-shear waves.

2.4.2 Hexagonal solids

There are four independent eigen spaces in a hexagonal (transversely isotropic) solids

$$W = W_1^{(1)}[\boldsymbol{\varphi}_1] \oplus W_2^{(1)}[\boldsymbol{\varphi}_2] \oplus W_3^{(2)}[\boldsymbol{\varphi}_3, \boldsymbol{\varphi}_6] \oplus W_4^{(2)}[\boldsymbol{\varphi}_4, \boldsymbol{\varphi}_5] \tag{30}$$

where

$$\left.\begin{aligned} \boldsymbol{\varphi}_{1,2} &= \frac{c_{13}}{\sqrt{(\lambda_{1,2} - c_{11} - c_{12})^2 + 2c_{13}^{\ 2}}} \times [1, 1, \frac{\lambda_{1,2} - c_{11} - c_{12}}{c_3}, 0, 0, 0]^{\mathrm{T}} \\ \boldsymbol{\varphi}_3 &= \frac{\sqrt{2}}{2}[1, -1, 0, 0, 0, 0]^{\mathrm{T}}, \qquad \boldsymbol{\varphi}_i = \xi_i, \ i = 4, 5, 6 \end{aligned}\right\} \tag{31}$$

The eigenelasticity and eigenoperator of hexagonal solids are

$$\left.\begin{aligned} \lambda_{1,2} &= \frac{c_{11} + c_{12} + c_{33}}{2} \pm \sqrt{\left(\frac{c_{11} + c_{12} + c_{33}}{2}\right)^2 + 2c_{13}^2}, \\ \lambda_3 &= c_{11} - c_{12}, \quad \lambda_4 = c_{44}, \\ \Delta_{1,2}^* &= \frac{c_{13}^2}{\left(\lambda_{1,2} - c_{11} - c_{12}\right)^2 + 2c_{13}^2} \nabla_{III}^*, \\ \Delta_3^* &= \frac{2}{3}\nabla_{II}^2, \quad \Delta_4^* = \frac{1}{2}\left(\nabla_{III}^2 + 2\partial_{12}\right), \end{aligned}\right\} \tag{32}$$

Thus there exist four independent elastic waves in hexagonal solids, which can be described by following equations

$$\lambda_i \Delta_i^* \varepsilon_i^{\ *}(x,t) = \rho \ddot{\varepsilon}_i^{\ *}(x,t) \qquad i = 1,2,3,4 \tag{33}$$

where

$$\varepsilon_{1,2}^* = \frac{c_{13}}{\sqrt{(\lambda_{1,2} - c_{11} - c_{12})^2 + 2c_{13}^{\ 2}}} \times [\varepsilon_{11} + \varepsilon_{22} + (\frac{\lambda_{1,2} - c_{11} - c_{12}}{c_{13}})\varepsilon_{33}] \tag{34}$$

$$\varepsilon_3^* = \sqrt{\frac{1}{2}\left(\varepsilon_{11} - \varepsilon_{22}\right)^2 + \varepsilon_{12}^2} \tag{35}$$

$$\varepsilon_4^{\ *} = \sqrt{\frac{1}{2}\left(\varepsilon_{32}^{\ 2} + \varepsilon_{31}^{\ 2}\right)} \tag{36}$$

It is seen that there exist four elastic waves in hexagonal solids, two of which are the quasi-dilation wave, and two others are the quasi-shear waves.

3. Electromagnetic waves in anisotropic solids

3.1 Eigen spaces of electromagnetic solids

In anisotropic electromagnetic solids, the dielectric permittivity and magnetic permeability are tensors instead of scalars. The constitutive relations are expressed as follows

$$\boldsymbol{D} = \boldsymbol{e} \cdot \boldsymbol{E} \tag{37}$$

$$\boldsymbol{B} = \boldsymbol{\mu} \cdot \boldsymbol{H} \tag{38}$$

where the dielectric permittivity matrix $\boldsymbol{e}$ and the magnetic permeability matrix $\boldsymbol{\mu}$ are usually symmetric ones, and the elements of the matrixes have a close relationship with the selection of reference coordinate. Suppose that if the reference coordinates is selected along principal axis of electrically or magnetically anisotropic solids, the elements at non-diagonal of these matrixes turn to be zero. Therefore, Eqs. (37),(38) are called the constitutive equations of electromagnetic solids under the geometric presentation. Now we intend to get rid of effects of geometric coordinate on the constitutive equations, and establish a set of coordinate-independent constitutive equations of electromagnetic media under physical presentation. For this purpose, we solve the following problems of eigen-value of matrixes

$$\left(\boldsymbol{e} - \eta \boldsymbol{I}\right)\boldsymbol{\phi} = 0 \tag{39}$$

$$\left(\boldsymbol{\mu} - \gamma \boldsymbol{I}\right)\boldsymbol{\vartheta} = \boldsymbol{0} \tag{40}$$

where $\eta_I \left(I = 1,2,3\right)$ and $\gamma_I \left(I = 1,2,3\right)$ are respectively eigen dielectric permittivity and eigen magnetic permeability, which are constants of coordinate-independent. $\boldsymbol{\phi}_I \left(I = 1,2,3\right)$ and $\boldsymbol{\vartheta}_I \left(I = 1,2,3\right)$ are respectively eigen electric vector and eigen magnetic vector, which show the electrically principal direction and magnetically principal direction of anisotropic solids, and are all coordinate-dependent. We call these vectors as eigen spaces. Thus, the matrix of dielectric permittivity and magnetic permeability can be spectrally decomposed as follows

$$\boldsymbol{e} = \boldsymbol{\Psi} \boldsymbol{\Gamma} \boldsymbol{\Psi}^T \tag{41}$$

$$\boldsymbol{\mu} = \boldsymbol{\Theta} \boldsymbol{\Pi} \boldsymbol{\Theta}^{\mathrm{T}} \tag{42}$$

where $\boldsymbol{\Gamma} = diag\left[\eta_1, \eta_2, \eta_3\right]$ and $\boldsymbol{\Pi} = diag\left[\gamma_1, \gamma_2, \gamma_3\right]$ are the matrix of eigen dielectric permittivity and eigen magnetic permeability, respectively. $\boldsymbol{\Psi} = \left\{\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \boldsymbol{\phi}_3\right\}$ and

$\boldsymbol{\Theta} = \{\boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2, \boldsymbol{\vartheta}_3\}$ are respectively the modal matrix of electric media and magnetic media, which are both orthogonal and positive definite matrixes, and satisfy, $\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Psi} = \boldsymbol{I}$, $\boldsymbol{\Theta}^{\mathrm{T}}\boldsymbol{\Theta} = \boldsymbol{I}$.
Projecting the electromagnetic physical qualities of geometric presentation, such as the electric field intensity vector $\boldsymbol{E}$, magnetic field intensity vector $\boldsymbol{H}$, magnetic flux density vector $\boldsymbol{B}$ and electric displacement vector $\boldsymbol{D}$ into eigen spaces of physical presentation, we get

$$\boldsymbol{D}^* = \boldsymbol{\Psi}^{\mathrm{T}} \cdot \boldsymbol{D} \qquad \boldsymbol{E}^* = \boldsymbol{\Psi}^{\mathrm{T}} \cdot \boldsymbol{E} \tag{43}$$

$$\boldsymbol{B}^* = \boldsymbol{\Theta}^{\mathrm{T}} \cdot \boldsymbol{B} \qquad \boldsymbol{H}^* = \boldsymbol{\Theta}^{\mathrm{T}} \cdot \boldsymbol{H} \tag{44}$$

Rewriting Eqs.(43) and (44) in the form of scalar, we have

$$D_I^* = \boldsymbol{\phi}_I^{\mathrm{T}} \cdot \boldsymbol{D} \quad I = 1 \text{-} n \qquad E_I^* = \boldsymbol{\phi}_I^{\mathrm{T}} \cdot \boldsymbol{E} \quad I = 1 \text{-} n \tag{45}$$

$$B_I^* = \boldsymbol{\vartheta}_I^{\mathrm{T}} \cdot \boldsymbol{B} \quad I = 1 \text{-} n \qquad H_I^* = \boldsymbol{\vartheta}_I^{\mathrm{T}} \cdot \boldsymbol{H} \quad I = 1 \text{-} n \tag{46}$$

where $n(\leq 3)$ are number of electromagnetic independent subspaces. These are the electromagnetic physical qualities under the physical presentation.
Substituting Eqs. (43) and (44) into Eqs. (37) and (38) respectively, and using Eqs.(45) and (46) yield

$$D_I^* = \eta_I E_I^* \quad I = 1 \text{-} n \tag{47}$$

$$B_I^* = \gamma_I H_I^* \quad I = 1 \text{-} n \tag{48}$$

The above equations are just the modal constitutive equations of electromagnetic media in the form of scalar.

3.2 Matrix expression of Maxwell's equation

The classical Maxwell's equations in passive region can be written as

$$e_{IJK} H_{K,J} = \dot{D}_I \tag{49}$$

$$e_{IJK} E_{K,J} = -\dot{B}_I \tag{50}$$

Now we rewrite above equations in the form of matrix as follows

$$\begin{bmatrix} 0 & -\partial_z & \partial_y \\ \partial_z & 0 & -\partial_x \\ -\partial_y & \partial_x & 0 \end{bmatrix} \begin{Bmatrix} H_1 \\ H_2 \\ H_3 \end{Bmatrix} = \nabla_t \begin{Bmatrix} D_1 \\ D_2 \\ D_3 \end{Bmatrix} \tag{51}$$

or

$$[\Delta_c]\{H\} = \nabla_t \{D\} \tag{52}$$

$$\begin{bmatrix} 0 & -\partial_z & \partial_y \\ \partial_z & 0 & -\partial_x \\ -\partial_y & \partial_x & 0 \end{bmatrix} \begin{Bmatrix} E_1 \\ E_2 \\ E_3 \end{Bmatrix} = -\nabla_t \begin{Bmatrix} B_1 \\ B_2 \\ B_3 \end{Bmatrix} \tag{53}$$

Or

$$[\Delta_d]\{E\} = -\nabla_t \{B\} \tag{54}$$

where $[\Delta_d]$ and $[\Delta_c]$ are respectively defined as the matrix of electric operators and magnetic operators, and $[\Delta_d] = [\Delta_c]$.
Substituting Eqs.(37) and (38) into Eqs. (52) and (54), we have

$$[\Delta_c]\{H\} = \nabla_t [e]\{E\} \tag{55}$$

$$[\Delta_d]\{E\} = -\nabla_t [\mu]\{H\} \tag{56}$$

Substituting Eq. (55) into (56) or Eq. (56) into (55) yield

$$[\square]\{H\} = -\nabla_t^2 [\mu][e]\{H\} \tag{57}$$

$$[\square]\{E\} = -\nabla_t^2 [\mu][e]\{E\} \tag{58}$$

where $[\square] = [\Delta_d][\Delta_c]$ is defined as the matrix of electromagnetic operators as follows

$$[\square] = \begin{bmatrix} -\left(\partial_z^2 + \partial_y^2\right) & \partial_{xy} & \partial_{xz} \\ \partial_{yx} & -\left(\partial_x^2 + \partial_z^2\right) & \partial_{yz} \\ \partial_{zx} & \partial_{zy} & -\left(\partial_x^2 + \partial_y^2\right) \end{bmatrix} \tag{59}$$

3.2.1 Electrically anisotropic solids

In anisotropic dielectrics, the dielectric permittivity is a tensor, while the magnetic permeability is a scalar. So Eqs. (57) and (58) can be written as follows

$$[\square]\{H\} = -\nabla_t^2 \mu_0 [e]\{H\} \tag{60}$$

$$[\square]\{E\} = -\nabla_t^2 \mu_0 [e]\{E\} \tag{61}$$

Substituting Eqs. (43), (44) and (47) into Eqs. (60) and (61), we have

$$\left[\square^*\right]\left\{H^*\right\} = -\nabla_t^2 \mu_0 [\Gamma]\left\{H^*\right\} \tag{62}$$

$$\left[\square^*\right]\left\{E^*\right\} = -\nabla_t^2 \mu_0 [\Gamma]\left\{E^*\right\} \tag{63}$$

where $\left[\square^*\right] = [\Psi]^{\mathrm{T}}[\square][\Psi]$ is defined as the eigen matrix of electromagnetic operators under the eigen spaces, is a diagonal matrix. Thus Eqs. (62) and (63) can be uncoupled in the form of scalar.

$$\square_I^* H_I^* + \mu_0 \eta_I \nabla_t^2 H_I^* = 0 \qquad I = 1 - n \tag{64}$$

$$\square_I^* E_I^* + \mu_0 \eta_I \nabla_t^2 E_I^* = 0 \qquad I = 1 - n \tag{65}$$

Eqs.(64) and (65) are the modal equations of electromagnetic waves in anisotropic dielectrics.

3.2.2 Magnetically anisotropic solids

In anisotropic magnetics, the magnetic permeability is a tensor, while the dielectric permittivity is a scalar. So Eqs. (57) and (58) can be written as follows

$$[\square]\{H\} = -\nabla_t^2 e_0 [\mu]\{H\} \tag{66}$$

$$[\square]\{E\} = -\nabla_t^2 e_0 [\mu]\{E\} \tag{67}$$

Substituting Eqs. (43), (44) and (48) into Eqs. (66) and (67), we have

$$\left[\square^*\right]\left\{H^*\right\} = -\nabla_t^2 e_0 [\Gamma]\left\{H^*\right\} \tag{68}$$

$$\left[\square^*\right]\left\{E^*\right\} = -\nabla_t^2 e_0 [\Gamma]\left\{E^*\right\} \tag{69}$$

where $\left[\square^*\right] = [\Theta]^{\mathrm{T}}[\square][\Theta]$ is defined as the eigen matrix of electromagnetic operators under the eigen spaces, is a diagonal matrix. Thus Eqs. (68) and (69)can be uncoupled in the form of scalar.

$$\square_I^* H_I^* + e_0 \gamma_I \nabla_t^2 H_I^* = 0 \qquad I = 1 - n \tag{70}$$

$$\square_I^* E_I^* + e_0 \gamma_I \nabla_t^2 E_I^* = 0 \qquad I = 1 - n \tag{71}$$

Eqs.(70) and (71) are the modal equations of electromagnetic waves in anisotropic magnetics.

3.3 Electromagnetic waves in anisotropic solids

In this section, we discuss the propagation behaviour of electromagnetic waves only in anisotropic dielectrics.

3.3.1 Isotropic crystal

The matrix of dielectric permittivity of isotropic dielectrics is following

$$\boldsymbol{e} = \begin{bmatrix} e_{11} & 0 & 0 \\ 0 & e_{11} & 0 \\ 0 & 0 & e_{11} \end{bmatrix} \tag{72}$$

The eigen-values and eigen-vectors are respectively shown as below

$$\boldsymbol{\Gamma} = diag[e_{11}, e_{11}, e_{11}] \tag{73}$$

$$\boldsymbol{\Psi} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{74}$$

We can see from the above equations that there is only one eigen-space in isotropic crystal, which is a triple-degenerate one, and the space structure is following

$$\boldsymbol{W} = W_1^{(3)}\left[\phi_1, \phi_2, \phi_3\right] \tag{75}$$

The basic vector of one dimension in a triple-degenerate subspace is

$$\boldsymbol{\phi}_1^* = \frac{\sqrt{3}}{3}\{1,1,1\}^T \tag{76}$$

The eigen-qualities and eigen-operators of isotropic crystal are respectively shown as below

$$E_1^* = \frac{1}{3}\left(E_1 + E_2 + E_3\right) \tag{77}$$

$$\square_1^* = \frac{1}{3}\left[-\left(\partial_x^2 + \partial_y^2 + \partial_z^2\right)\right] \tag{78}$$

Thus the equations of electromagnetic waves in isotropic crystal can be written as

$$\left(\partial_x^2 + \partial_y^2 + \partial_z^2\right)\left(E_1 + E_2 + E_3\right) = \mu_0 e_{11}\partial_t^2\left(E_1 + E_2 + E_3\right) \tag{79}$$

or

$$\left(\partial_x^2 + \partial_y^2 + \partial_z^2\right)E_1 = \mu_0 e_{11}\partial_t^2 E_1 \tag{80}$$

$$\left(\partial_x^2 + \partial_y^2 + \partial_z^2\right)E_2 = \mu_0 e_{11}\partial_t^2 E_2 \tag{81}$$

$$\left(\partial_x^2 + \partial_y^2 + \partial_z^2\right)E_3 = \mu_0 e_{11}\partial_t^2 E_3 \tag{82}$$

The velocity of the electromagnetic wave is

$$c = \frac{1}{\sqrt{\mu_0 e_{11}}} \tag{83}$$

Eq. (79) is just the Helmholtz's equation of electromagnetic wave.

3.3.2 Uniaxial crystal

The matrix of dielectric permittivity of uniaxial dielectrics is following

$$\boldsymbol{e} = \begin{bmatrix} e_{11} & 0 & 0 \\ 0 & e_{11} & 0 \\ 0 & 0 & e_{33} \end{bmatrix} \tag{84}$$

The eigen-values and eigen-vectors are respectively shown as below

$$\boldsymbol{\Lambda} = diag\left[e_{11}, e_{11}, e_{33}\right] \tag{85}$$

$$\boldsymbol{\Psi} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{86}$$

We can see from the above equations that there are two eigen-spaces in uniaxial crystal, one of which is a double-degenerate space, and the space structure is following

$$\boldsymbol{W} = W_1^{(2)}\left[\phi_1, \phi_2\right] \oplus W_2^1\left[\phi_3\right] \tag{87}$$

The basic vectors in two subspaces are following

$$\boldsymbol{\phi}_1^* = \frac{\sqrt{2}}{2}\left\{1,1,0\right\}^T \tag{88}$$

$$\boldsymbol{\phi}_2^* = \left\{0,0,1\right\}^T \tag{89}$$

The eigen electric strength qualities of uniaxial crystal are respectively shown as below

$$E_2^* = \boldsymbol{\phi}_2^{\mathrm{T}} \cdot \boldsymbol{E} = E_3 \tag{90}$$

$$\boldsymbol{\phi}_1^{\mathrm{T}} E_1^* = \boldsymbol{E} - \boldsymbol{\phi}_2^{\mathrm{T}} E_2^* \tag{91}$$

Multiplying Eq.(91) with $\boldsymbol{\phi}_2$, using $\boldsymbol{\phi}_2^{\mathrm{T}} \cdot \boldsymbol{\phi}_1 = 0$ and $\boldsymbol{\phi}_i^{\mathrm{T}} \cdot \boldsymbol{\phi}_i = 1\left(i = 1,2\right)$, we get

$$\left|E_1^*\right| = \sqrt{\left(\boldsymbol{E} - \boldsymbol{\phi}_2^{\mathrm{T}} E_2^*\right)^{\mathrm{T}} \left(\boldsymbol{E} - \boldsymbol{\phi}_2^{\mathrm{T}} E_2^*\right)} = \sqrt{E_1^2 + E_2^2} \tag{92}$$

The eigen-operators of uniaxial crystal are respectively shown as below

$$\square_1^* = -\left(\partial_x^2 + \partial_y^2 + 2\partial_z^2 - 2\partial_{xy}^2\right) \tag{93}$$

$$\square_2^* = -\left(\partial_x^2 + \partial_y^2\right) \tag{94}$$

Therefore, the equations of electromagnetic waves in uniaxial crystal can be written as below

$$\left(\partial_x^2 + \partial_y^2 + 2\partial_z^2 - 2\partial_{xy}^2\right)\sqrt{E_1^2 + E_2^2} = \mu_0 e_{11} \partial_t^2 \sqrt{E_1^2 + E_2^2} \tag{95}$$

$$\left(\partial_x^2 + \partial_y^2\right) E_3 = \mu_0 e_{33} \partial_t^2 E_3 \tag{96}$$

The velocities of electromagnetic waves are respectively as follows

$$c^{(1)} = \frac{1}{\sqrt{\mu_0 e_{11}}} \tag{97}$$

$$c^{(2)} = \frac{1}{\sqrt{\mu_0 e_{33}}} \tag{98}$$

It is seen that there are two kinds of electromagnetic waves in uniaxial crystal.

3.3.3 Biaxial crystal

The matrix of dielectric permittivity of biaxial dielectrics is following

$$\boldsymbol{e} = \begin{bmatrix} e_{11} & 0 & 0 \\ 0 & e_{22} & 0 \\ 0 & 0 & e_{33} \end{bmatrix} \tag{99}$$

The eigen-values and eigen-vectors are respectively shown as below

$$\boldsymbol{\Lambda} = diag\left[e_{11}, e_{22}, e_{33}\right] \tag{100}$$

$$\boldsymbol{\Psi} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{101}$$

We can see from the above equations that there are three eigen-spaces in biaxial crystal, and the space structure is following

$$\boldsymbol{W} = W_1^{(1)}\left[\phi_1\right] \oplus W_2^{(1)}\left[\phi_2\right] \oplus W_3^{(1)}\left[\phi_3\right] \tag{102}$$

The eigen-qualities and eigen-operators of biaxial crystal are respectively shown as below

$$E_1^* = \boldsymbol{\phi}_1^{\mathrm{T}} \cdot \boldsymbol{E} = E_1 \tag{103}$$

$$E_2^* = \boldsymbol{\phi}_2^{\mathrm{T}} \cdot \boldsymbol{E} = E_2 \tag{104}$$

$$E_3^* = \boldsymbol{\phi}_3^{\mathrm{T}} \cdot \boldsymbol{E} = E_3 \tag{105}$$

$$\square_1^* = -\left(\partial_z^2 + \partial_y^2\right) \tag{106}$$

$$\square_2^* = -\left(\partial_x^2 + \partial_z^2\right) \tag{107}$$

$$\square_3^* = -\left(\partial_x^2 + \partial_y^2\right) \tag{108}$$

Therefore, the equations of electromagnetic waves in biaxial crystal can be written as below

$$\left(\partial_z^2 + \partial_y^2\right) E_1 = \mu_0 e_{11} \partial_t^2 E_1 \tag{109}$$

$$\left(\partial_x^2+\partial_z^2\right)E_2=\mu_0 e_{22}\partial_t^2 E_2 \tag{110}$$

$$\left(\partial_x^2+\partial_y^2\right)E_3=\mu_0 e_{33}\partial_t^2 E_3 \tag{111}$$

The velocities of electromagnetic waves are respectively as follows

$$c^{(1)}=\frac{1}{\sqrt{\mu_0 e_{11}}} \tag{112}$$

$$c^{(2)}=\frac{1}{\sqrt{\mu_0 e_{22}}} \tag{113}$$

$$c^{(3)}=\frac{1}{\sqrt{\mu_0 e_{33}}} \tag{114}$$

It is seen that there are three kinds of electromagnetic waves in biaxial crystal.

3.3.4 Monoclinic crystal

The matrix of dielectric permittivity of monoclinic dielectrics is following

$$\boldsymbol{e}=\begin{bmatrix} e_{11} & e_{12} & 0 \\ e_{12} & e_{22} & 0 \\ 0 & 0 & e_{33} \end{bmatrix} \tag{115}$$

The eigen-values and eigen-vectors are respectively shown as below

$$\boldsymbol{\Gamma}=diag\left[\eta_1,\eta_2,e_{33}\right] \tag{116}$$

$$\begin{cases} \boldsymbol{\phi}_1=\dfrac{e_{12}}{\sqrt{\left(\eta_1-e_{11}\right)^2+e_{12}^2}}\left[\dfrac{\eta_1-e_{11}}{e_{12}},1,0\right]^{\mathrm{T}} \\ \boldsymbol{\phi}_2=\dfrac{e_{12}}{\sqrt{\left(\eta_2-e_{11}\right)^2+e_{12}^2}}\left[1,\dfrac{\eta_2-e_{11}}{e_{12}},0\right]^{\mathrm{T}} \\ \boldsymbol{\phi}_3=\left[0,0,1\right]^{\mathrm{T}} \end{cases} \tag{117}$$

where

$$\eta_{1,2}=\frac{\left(e_{11}+e_{22}\right)}{2}\pm\sqrt{\left[\frac{1}{2}\left(e_{11}-e_{22}\right)\right]^2+e_{12}^2}\ ,\ \eta_3=e_{33} \tag{118}$$

We can see from the above equations that there are also three eigen-spaces in monoclinic crystal, and the space structure is following

$$\boldsymbol{W} = W_1^{(1)}[\phi_1] \oplus W_2^{(1)}[\phi_2] \oplus W_3^{(1)}[\phi_3] \tag{119}$$

The eigen-qualities and eigen-operators of monoclinic crystal are respectively shown as below

$$E_1^* = \boldsymbol{\phi}_1^{\mathrm{T}} \cdot \boldsymbol{E} = \frac{1}{\sqrt{(\eta_1 - e_{11})^2 + e_{12}^2}}\left[(\eta_1 - e_{11})E_1 + e_{12}E_2\right] \tag{120}$$

$$E_2^* = \boldsymbol{\phi}_2^{\mathrm{T}} \cdot \boldsymbol{E} = \frac{1}{\sqrt{(\eta_2 - e_{11})^2 + e_{12}^2}}\left[e_{12}E_1 + (\eta_2 - e_{11})E_2\right] \tag{121}$$

$$E_3^* = \boldsymbol{\phi}_3^{\mathrm{T}} \cdot \boldsymbol{E} = E_3 \tag{122}$$

$$\square_1^* = -\frac{e_{12}^2}{(\eta_1 - e_{11})^2 + e_{12}^2}\left[\left(\frac{\eta_1 - e_{11}}{e_{12}}\right)^2\left(\partial_z^2 + \partial_y^2\right) + \left(\partial_x^2 + \partial_z^2\right) - 2\left(\frac{\eta_1 - e_{11}}{e_{12}}\right)\partial_{xy}^2\right] \tag{123}$$

$$\square_2^* = -\frac{e_{12}^2}{(\eta_2 - e_{11})^2 + e_{12}^2}\left[\left(\frac{\eta_2 - e_{11}}{e_{12}}\right)^2\left(\partial_x^2 + \partial_z^2\right) + \left(\partial_z^2 + \partial_y^2\right) - 2\left(\frac{\eta_2 - e_{11}}{e_{12}}\right)\partial_{xy}^2\right] \tag{124}$$

$$\square_3^* = -\left(\partial_x^2 + \partial_y^2\right) \tag{125}$$

Therefore, the equations of electromagnetic waves in monoclinic crystal can be written as below

$$\square_1^*\left[(\eta_1 - e_{11})E_1 + e_{12}E_2\right] = \mu_0\eta_1\partial_t^2\left[(\eta_1 - e_{11})E_1 + e_{12}E_2\right] \tag{126}$$

$$\square_2^*\left[e_{12}E_1 + (\eta_2 - e_{11})E_2\right] = \mu_0\eta_2\partial_t^2\left[e_{12}E_1 + (\eta_2 - e_{11})E_2\right] \tag{127}$$

$$\left(\partial_x^2 + \partial_y^2\right)E_3 = \mu_0 e_{33}\partial_t^2 E_3 \tag{128}$$

The velocities of electromagnetic waves are respectively as follows

$$c^{(1)} = \frac{1}{\sqrt{\mu_0\left\{\frac{(e_{11} + e_{22})}{2} + \sqrt{\left[\frac{1}{2}(e_{11} - e_{22})\right]^2 + e_{12}^2}\right\}}} \tag{129}$$

$$c^{(2)} = \frac{1}{\sqrt{\mu_0\left\{\frac{(e_{11} + e_{22})}{2} - \sqrt{\left[\frac{1}{2}(e_{11} - e_{22})\right]^2 + e_{12}^2}\right\}}} \tag{130}$$

$$c^{(3)} = \frac{1}{\sqrt{\mu_0 e_{33}}} \tag{131}$$

It is seen that there are also three kinds of electromagnetic waves in monoclinic crystal. In comparison with the waves in biaxial crystal, the electromagnetic waves in monoclinic crystal have been distorted.

4. Quasi-static waves in piezoelectric solids

4.1 Modal constitutive equation of piezoelectric solids

For a piezoelectric but nonmagnetizable dielectric body, the constitutive equations is the following

$$\boldsymbol{\sigma} = \boldsymbol{c} \cdot \boldsymbol{\varepsilon} - \boldsymbol{h}^{\mathrm{T}} \cdot \boldsymbol{E} \tag{132}$$

$$\boldsymbol{D} = \boldsymbol{h} \cdot \boldsymbol{\varepsilon} + \boldsymbol{e} \cdot \boldsymbol{E} \tag{133}$$

$$\boldsymbol{B} = \boldsymbol{\mu} \cdot \boldsymbol{H} \tag{134}$$

where $\boldsymbol{h}$ is the piezoelectric matrix.

Substituting Eqs. (5), (43) and (44) into Eqs. (132)-(134), respectively, and multiplying them with the transpose of modal matrix in the left, we have

$$\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\sigma} = \boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{c}\boldsymbol{\Phi}\boldsymbol{\varepsilon}^{*} - \boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{T}}\boldsymbol{\Psi}\boldsymbol{E}^{*} \tag{135}$$

$$\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{D} = \boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{h}\boldsymbol{\Phi}\boldsymbol{\varepsilon}^{*} + \boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{e}\boldsymbol{\Psi}\boldsymbol{E}^{*} \tag{136}$$

$$\boldsymbol{\Theta}^{\mathrm{T}}\boldsymbol{B} = \boldsymbol{\Theta}^{\mathrm{T}}\boldsymbol{\mu}\boldsymbol{\Theta}\boldsymbol{H}^{*} \tag{137}$$

Let $\boldsymbol{G} = \boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{h}\boldsymbol{\Phi}$, $\boldsymbol{G}^{\mathrm{T}} = \boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{h}^{\mathrm{T}}\boldsymbol{\Psi}$, that is a coupled piezoelectric matrix, and using Eqs.(2), (41) and (42), we get

$$\boldsymbol{\sigma}^{*} = \boldsymbol{\Lambda}\boldsymbol{\varepsilon}^{*} - \boldsymbol{G}^{\mathrm{T}}\boldsymbol{E}^{*} \tag{138}$$

$$\boldsymbol{D}^{*} = \boldsymbol{G}\boldsymbol{\varepsilon}^{*} + \boldsymbol{\Gamma}\boldsymbol{E}^{*} \tag{139}$$

$$\boldsymbol{B}^{*} = \boldsymbol{\Pi}\boldsymbol{H}^{*} \tag{140}$$

Rewriting the above equations in the form of scalar, we have

$$\sigma_i^{*} = \lambda_i \varepsilon_i^{*} - g_{iJ}^{T} E_J^{*} \qquad i = 1\text{-}m \qquad J = 1\text{-}n \tag{141}$$

$$D_I^{*} = \eta_I E_I^{*} + g_{Ij}\varepsilon_j^{*} \qquad I = 1\text{-}n \qquad j = 1\text{-}m \tag{142}$$

$$B_I^{*} = \gamma_I H_I^{*} \qquad I = 1\text{-}n \tag{143}$$

Eqs.(138)-(140) are just the modal constitutive equations for anisotropic piezoelectric body, in which $g_{Ij} = \{\phi_I\}^T [h] \{\varphi_j\}$, $g_{iJ}^T = \{\varphi_i\}^T [h]^T \{\phi_J\}$, $g_{iJ} = g_{Ji}^T$ are the coupled piezoelectric coefficients.

4.2 Modal equation of elastic waves in piezoelectric solids

When only acoustic waves are considered, we can use the of quasi-static electromagnetic approximation. In this case, the mechanical equations are dynamic, the electromagnetic equations are static and the electric field and the magnetic field are not coupled.
The static electric field equations and dynamic equations are given as follows

$$\nabla_I^* D_I^* = 0 \qquad I = 1 - n \tag{144}$$

$$\Delta_i^* \sigma_i^{\,*} = \rho \Delta_{tt} \varepsilon_i^{\,*} \qquad i = 1, 2, \cdots, m \tag{145}$$

where

$$\left[\nabla^*\right] = \left[\Psi\right]^T \left[\nabla\right]\left[\Psi\right] \tag{146}$$

$$\left[\nabla\right] = \begin{bmatrix} \partial_{11} & \partial_{21} & \partial_{31} \\ \partial_{12} & \partial_{22} & \partial_{32} \\ \partial_{13} & \partial_{23} & \partial_{33} \end{bmatrix} \tag{147}$$

Substituting Eq.(142) into Eq.(144), we have

$$\eta_I \nabla_I^* E_I^* + \nabla_I^* g_{Ij} \varepsilon_j^* = 0 \tag{148}$$

The above can also be rewritten as follows

$$\nabla_J^* E_J^* + \nabla_J^* \left(\frac{g_{Ji}}{\eta_J} \right) \varepsilon_i^* = 0 \qquad J = 1 \text{ - } n \qquad i = 1 \text{ - } m \quad \text{(sum to i)} \tag{149}$$

According to the principle of operator, Eq.(148) becomes

$$E_J^* = -\left(\frac{g_{Jk}}{\eta_J} \right) \varepsilon_k^* \qquad J = 1 \text{ - } n \qquad k = 1 \text{ - } m \tag{150}$$

In same way, Substituting Eq.(141) into Eq.(145), we have

$$\Delta_i^* \lambda_i \varepsilon_i^* - \Delta_i^* g_{ij}^T E_J^* = \rho \nabla_{tt} \varepsilon_i^* \qquad i = 1 - m \tag{151}$$

Using Eq.(150), Eq.(151) becomes

$$\Delta_i^* \lambda_i \varepsilon_i^* + \Delta_i^* g_{ij}^T \left(\frac{g_{Jk}}{\eta_J} \right) \delta_{ik} \varepsilon_i^* = \rho \nabla_{tt} \varepsilon_i^* \qquad i = 1 - m \qquad J = 1 - n \tag{152}$$

According to the principle of operator, Eq.(152) becomes

$$\left[g_{ij}^T \left(\frac{g_{Jk}}{\eta_J} \right) \delta_{ik} + \lambda_i \right] \Delta_i^* \varepsilon_i^* = \rho \nabla_{tt} \varepsilon_i^* \qquad i = 1 - m \qquad J = 1 - n \quad \text{(sum to J)} \tag{153}$$

They are the equations of elastic waves in quasi-static piezoelectricity, in which the propagation speed of elastic waves is the folloing

$$v_i = \sqrt{\frac{\lambda_i + \sum_J g_{ij}^T \left(\frac{g_{Jk}}{\eta_J} \right) \delta_{ik}}{\rho}} \qquad i = 1 - m \tag{154}$$

4.3 Modal equation of electromagnetic waves in piezoelectric solids

When electromagnetic waves are involved, the complete set of Maxwell equation needs to be used, coupled to the static mechanical equations as follows

$$\Delta_i^* \sigma_i^* = 0 \qquad i = 1, 2, \cdots, m \tag{155}$$

$$\left\{ \Delta_I^* \right\} E_I^* = -\nabla_t \left\{ \vartheta_I^* \right\} B_I^* \qquad I = 1 - n \tag{156}$$

$$\left\{ \Delta_I^* \right\} H_I^* = \nabla_t \left\{ \phi_I^* \right\} D_I^* \qquad I = 1 - n \tag{157}$$

Substituting Eq. .(141)-(143) into Eq.(155)-(157), we have

$$\left\{ \Delta_I^* \right\} E_I^* = -\nabla_t \left\{ \vartheta_I \right\} \gamma_I H_I^* \qquad I = 1 - n \tag{158}$$

$$\left\{ \Delta_I^* \right\} H_I^* = \nabla_t \left\{ \phi_I \right\} \left(\eta_I E_I^* + g_{Ij} \varepsilon_j^* \right) \qquad I = 1 - n \tag{159}$$

$$\Delta_i^* \left(\lambda_i \varepsilon_i^* - g_{iJ}^T E_J^* \right) = 0 \qquad i = 1 - m \tag{160}$$

Transposing Eq.(158), and multiplying it with $\left\{ \Delta_I^* \right\}$, and also using Eq.(159), we have

$$\left\{ \Delta_I^* \right\} \left\{ \Delta_I^* \right\}^T E_I^* = -\nabla_{tt} \left\{ \vartheta_I \right\}^T \left\{ \phi_I \right\} \gamma_I \left(\eta_I E_I^* + g_{Ij} \varepsilon_j^* \right) \tag{161}$$

Let $\square_I^* = \left\{ \Delta_I^* \right\} \cdot \left\{ \Delta_I^* \right\}^T$ and $\xi_I = \left\{ \vartheta_I \right\}^T \left\{ \phi_I \right\}$, Eq.(159) can be written as

$$\square_I^* E_I^* + \nabla_{tt} \xi_I \gamma_I \eta_I E_I^* = -\nabla_{tt} \xi_I \gamma_I g_{Ij} \varepsilon_j^* \tag{162}$$

From Eq.(160), Using the principle of operator, and changing the index, we have

$$\varepsilon_j^* = \left(\frac{g_{jK}^T}{\lambda_j} \right) E_K^* \qquad j = 1, 2, \cdots, 6 \text{ (sum to K)} \tag{163}$$

Substituting Eq.(163) into Eq.(162), we get the equations of electric fields

$$\square_I^* E_I^* + \nabla_{tt} \xi_I \gamma_I \left[\eta_I + g_{Ij} \left(\frac{g_{jK}^T}{\lambda_j} \right) \delta_{IK} \right] E_I^* = 0 \text{ (sum to j, K)} \tag{164}$$

In same way, we can get the equations of magnetic fields

$$\square_I^* H_I^* + \nabla_{tt}\xi_I\gamma_I\left[\eta_I + g_{Ij}\left(\frac{g_{jK}^T}{\lambda_j}\right)\delta_{IK}\right]H_I^* = 0 \qquad \text{(sum to j, K)} \tag{165}$$

Eqs.(164) and (165) are just the eigen equations of electromagnetic waves in piezoelectric solids, the speed of electromagnetic waves are the following

$$c_I^2 = \frac{1}{\xi_I\gamma_I\left[\eta_I + g_{Ij}\left(\frac{g_{jK}^T}{\lambda_j}\right)\delta_{IK}\right]} \qquad I = 1-n \tag{166}$$

4.4 Eigen properties of polarized ceramics

In this section, we discuss the propagation laws of piezoelectromagnetic waves in an polarrized ceramics poled in the x_3-direction. The material tensors in Eqs.(132)-(134) are represented by the following matrices under the compact notation

$$\begin{bmatrix} c_{11} & c_{12} & c_{13} & 0 & 0 & 0 \\ c_{12} & c_{11} & c_{13} & 0 & 0 & 0 \\ c_{13} & c_{13} & c_{13} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{66} \end{bmatrix}, \begin{bmatrix} 0 & 0 & h_{31} \\ 0 & 0 & h_{31} \\ 0 & 0 & h_{33} \\ 0 & h_{15} & 0 \\ h_{15} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} e_{11} & 0 & 0 \\ 0 & e_{11} & 0 \\ 0 & 0 & e_{33} \end{bmatrix}, \begin{bmatrix} \mu_{11} & 0 & 0 \\ 0 & \mu_{11} & 0 \\ 0 & 0 & \mu_{33} \end{bmatrix} \tag{167}$$

where $c_{66} = \frac{1}{2}(c_{11} - c_{12})$.

There are four independent mechanical eigenspaces as follows

$$W_{mech} = W_1^{(1)}[\boldsymbol{\varphi}_1] \oplus W_2^{(1)}[\boldsymbol{\varphi}_2] \oplus W_3^{(2)}[\boldsymbol{\varphi}_3, \boldsymbol{\varphi}_6] \oplus W_4^{(2)}[\boldsymbol{\varphi}_4, \boldsymbol{\varphi}_5] \tag{168}$$

where

$$\left.\begin{aligned} \boldsymbol{\varphi}_{1,2} &= \frac{c_{13}}{\sqrt{(\lambda_{1,2} - c_{11} - c_{12})^2 + 2c_{13}^{\ 2}}} \times [1, 1, \frac{\lambda_{1,2} - c_{11} - c_{12}}{c_{13}}, 0, 0, 0]^{\mathrm{T}} \\ \boldsymbol{\varphi}_3 &= \frac{\sqrt{2}}{2}[1, -1, 0, 0, 0, 0]^{\mathrm{T}}, \qquad \boldsymbol{\varphi}_i = \xi_i,\ i = 4,5,6 \end{aligned}\right\} \tag{169}$$

$$\left.\begin{aligned} \lambda_{1,2} &= \frac{c_{11} + c_{12} + c_{33}}{2} \pm \sqrt{\left(\frac{c_{11} + c_{12} + c_{33}}{2}\right)^2 + 2c_{13}^2} \\ \lambda_3 &= c_{11} - c_{12}, \quad \lambda_4 = c_{44} \end{aligned}\right\} \tag{170}$$

Then, we have

$$\boldsymbol{\varphi}_{1,2}^* = \boldsymbol{\varphi}_{1,2},\ \boldsymbol{\varphi}_3^* = \frac{\sqrt{3}}{3}\left[1, -1, 0, 0, 0, 1\right]^T,\ \boldsymbol{\varphi}_4^* = \frac{\sqrt{2}}{2}\left[0, 0, 0, 1, 1, 0\right]^T \tag{171}$$

There are two independent electric or magnetic eigenspaces as follows

$$W_{ele} = W_1^{(2)}[\boldsymbol{\phi}_1, \boldsymbol{\phi}_2] \oplus W_2^{(1)}[\boldsymbol{\phi}_3] \tag{172}$$

$$W_{mag} = W_1^{(2)}[\boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2] \oplus W_2^{(1)}[\boldsymbol{\vartheta}_3] \tag{173}$$

where

$$\boldsymbol{\Gamma} = diag[e_{11}, e_{11}, e_{33}],\ \boldsymbol{\Pi} = diag[\mu_{11}, \mu_{11}, \mu_{33}] \tag{174}$$

$$\boldsymbol{\Psi} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\ \boldsymbol{\Theta} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{175}$$

Then, we have

$$\boldsymbol{\phi}_1^* = \boldsymbol{\vartheta}_1^* = \frac{\sqrt{2}}{2}[1,1,0]^T,\ \boldsymbol{\phi}_2^* = \boldsymbol{\vartheta}_2^* = [0,0,1]^T \tag{176}$$

It is seen that the electric subspaces are the same as magnetic ones for polarrized ceramics. Thus, the physical quantities, the coupled coefficients and the corresponding operators of polarrized ceramics are calculated as follows

$$\varepsilon_{1,2}^* = a_{1,2}[\varepsilon_{11} + \varepsilon_{22} + b_{1,2}\varepsilon_{33}],\ \varepsilon_3^* = \frac{\sqrt{3}}{3}(\varepsilon_{11} - \varepsilon_{22} + \varepsilon_{12}),\ \varepsilon_4^* = \frac{\sqrt{2}}{2}(\varepsilon_{32} + \varepsilon_{31}) \tag{177}$$

$$\left.\begin{aligned} &\Delta_{1,2}^* = a_{1,2}^2\left(\partial_x^2 + \partial_y^2 + b_{1,2}^2\partial_z^2\right), \\ &\Delta_3^* = \frac{2}{3}\left(\partial_x^2 + \partial_y^2\right),\quad \Delta_4^* = \frac{1}{2}\left(\partial_x^2 + \partial_y^2 + 2\partial_z^2 + 2\partial_{xy}\right) \end{aligned}\right\} \tag{178}$$

$$E_1^* = \frac{\sqrt{2}}{2}(E_1 + E_2),\ H_1^* = \frac{\sqrt{2}}{2}(H_1 + H_2),\ E_2^* = E_3,\ H_2^* = H_3 \tag{179}$$

$$\square_1^* = -\frac{1}{2}\left(\partial_x^2 + \partial_y^2 + 2\partial_z^2 - 2\partial_{xy}^2\right),\ \square_2^* = -\left(\partial_x^2 + \partial_y^2\right) \tag{180}$$

and

$$\xi_I = 1 \qquad I = 1,2$$

$$\begin{aligned} &g_{11} = 0, \quad g_{21} = 0, \quad g_{31} = 0, \quad g_{41} = h_{15} \\ &g_{12} = a_1\left(2h_{31} + b_1h_{33}\right), g_{22} = a_2\left(2h_{31} + b_2h_{33}\right), g_{32} = 0, g_{42} = 0 \end{aligned}$$

$$\begin{aligned} &g_{11}^T = 0, \quad g_{12}^T = 0, \quad g_{13}^T = 0, \quad g_{14}^T = h_{15} \\ &g_{21}^T = a_1\left(2h_{31} + b_1h_{33}\right), g_{22}^T = a_2\left(2h_{31} + b_2h_{33}\right), g_{23}^T = 0, g_{24}^T = 0 \end{aligned}$$

where

$$a_{1,2} = \frac{c_{13}}{\sqrt{(\lambda_{1,2} - c_{11} - c_{12})^2 + 2c_{13}^2}},\ b_{1,2} = \frac{\lambda_{1,2} - c_{11} - c_{12}}{c_{13}}.$$

4.5 Elastic waves in polarized ceramics

There are four equations of elastic waves in polarized ceramics, in which the propagation speed of elastic waves is the folloing, respectively

$$v_1^2 = \frac{\lambda_1}{\rho} \tag{181}$$

$$v_2^2 = \frac{\lambda_2 + \dfrac{a_1^2\left(2h_{31} + b_1 h_{33}\right)^2}{e_{11}} + \dfrac{a_2^2\left(2h_{31} + b_2 h_{33}\right)^2}{e_{33}}}{\rho} \tag{182}$$

$$v_3^2 = \frac{c_{11} - c_{12}}{\rho} \tag{183}$$

$$v_4^2 = \frac{c_{44} + \dfrac{h_{15}^2}{e_{11}}}{\rho} \tag{184}$$

It is seen that two of elastic waves are the quasi-dilation wave, and two others are the quasi-shear waves, and only two waves were affected by the piezoelectric coefficients, which speeds up the propagation of second and fourth waves.

4.6 Electromagnetic waves in polarized ceramics

There are two equations of electromagnetic waves in polarized ceramics, in which the propagation speed of electromagnetic waves is the folloing, respectively

$$c_1^2 = \frac{1}{\mu_{11}\left[e_{11} + \dfrac{a_1^2\left(2h_{31} + b_1 h_{33}\right)^2}{\lambda_2} + \dfrac{h_{15}^2}{c_{44}}\right]} \tag{185}$$

$$c_2^2 = \frac{1}{\mu_{22}\left[e_{22} + \dfrac{a_2^2\left(2h_{31} + b_2 h_{33}\right)^2}{\lambda_2}\right]} \tag{186}$$

It is seen that two electromagnetic waves are all affected by the piezoelectric coefficients, which slow down the propagation of electromagnetic waves.

5. Fully dynamic waves in piezoelectric solids

When fully dynamic waves are considered, the complete set of Maxwell equation needs to be used, coupled to the mechanical equations of motion as follows

$$\Delta_i^* \sigma_i^* = \rho \nabla_{tt} \varepsilon_i^* \qquad i = 1, 2, \cdots, m \tag{187}$$

$$\left\{\Delta_I^*\right\} E_I^* = -\nabla_t \left\{\vartheta_I^*\right\} B_I^* \qquad I = 1 - n \tag{188}$$

$$\left\{\Delta_I^*\right\} H_I^* = \nabla_t \left\{\phi_I^*\right\} D_I^* \qquad I = 1-n \tag{189}$$

Substituting Eqs. (141)-(143) into Eqs. (187)- (189), respectively, we have

$$\Delta_i^* \left(\lambda_i \varepsilon_i^* - g_{iJ}^T E_J^*\right) = \rho \nabla_{tt} \varepsilon_i^* \quad \text{(sum to J)} \tag{190}$$

$$\left\{\Delta_I^*\right\} E_I^* = -\nabla_t \left\{\vartheta_I\right\} \gamma_I H_I^* \qquad I = 1-n \tag{191}$$

$$\left\{\Delta_I^*\right\} H_I^* = \nabla_t \left\{\phi_I\right\} \left(\eta_I E_I^* + g_{Ij} \varepsilon_j^*\right) \qquad I = 1-n \tag{192}$$

Transposing Eq.(191), and multiplying it with $\left\{\Delta_I^*\right\}$, and using Eq.(192), we have

$$\left(\square_I^* + \nabla_{tt} \xi_I \gamma_I \eta_I\right) E_I^* = -\nabla_{tt} \xi_I \gamma_I g_{Ij} \varepsilon_j^* \qquad I = 1-n \qquad j = 1-m \quad \text{(sum to j)} \tag{193}$$

where $\square_I^* = \left\{\Delta_I^*\right\} \left\{\Delta_I^*\right\}^T$ and $\xi_I = \left\{\vartheta_I^*\right\}^T \left\{\phi_I^*\right\}$. Eq.(190) can be written as follows

$$\left(\lambda_j \Delta_j^* - \rho \nabla_{tt}\right) \varepsilon_j^* = \Delta_j^* g_{jI}^T E_I^* \qquad I = 1-n \qquad j = 1-m \quad \text{(sum to I)} \tag{194}$$

where $\left(\square_I^* + \nabla_{tt} \xi_I \gamma_I \eta_I\right)$ and $\left(\lambda_j \Delta_j^* - \rho \nabla_{tt}\right)$ are the electromagnetic dynamic operator and mechanical dynamic operator, respectively. In order to investigate the mutual effects between mechanical subspaces and electromagnetic subspaces, multiplying Eq.(191) with the mechanical dynamic operator and Eq.(192) with the electromagnetic dynamic operator, and substituting Eq. (194) into Eq. (193) and Eq. (193) into Eq. (194), we have

$$\left(\lambda_j \Delta_j^* - \rho \nabla_{tt}\right) \left(\square_I^* + \nabla_{tt} \xi_I \gamma_I \eta_I\right) E_I^{*(j)} = -\nabla_{tt} \xi_I \gamma_I g_{Ij} \Delta_j^* g_{jK}^T E_K^{*(j)} \quad \text{(sum to K)} \tag{195}$$

$$\left(\square_I^* + \nabla_{tt} \xi_I \gamma_I \eta_I\right) \left(\lambda_j \Delta_j^* - \rho \nabla_{tt}\right) \varepsilon_j^{*(I)} = g_{jI}^T \Delta_j^* \left(-\nabla_{tt} \xi_I \gamma_I g_{Ik} \varepsilon_k^{*(I)}\right) \quad \text{(sum to k)} \tag{196}$$

where $E_I^{*(j)}$ notes the I th modal electric field induced by j th mechanical subspace, and $\varepsilon_j^{*(I)}$ the j th modal strain field induced by I th electromagnetic subspace. Reorganizing Eqs.(195) and (196), we get

$$\rho \nabla_{tttt} \xi_I \gamma_I \eta_I E_I^{*(j)} + \nabla_{tt} \left[\rho \square_I^* - \Delta_j^* \xi_I \gamma_I \left(\lambda_j \eta_I + g_{Ij} g_{jK}^T \delta_{KI}\right)\right] E_I^{*(j)} - \lambda_j \Delta_j^* \square_I^* E_I^{*(j)} = 0$$

$$I = 1-n \qquad j = 1-m \quad \text{(sum to k)} \tag{197}$$

$$\rho \nabla_{tttt} \xi_I \gamma_I \eta_I \varepsilon_j^{*(I)} + \nabla_{tt} \left[\rho \square_I^* - \Delta_j^* \xi_I \gamma_I \left(\lambda_j \eta_I + g_{Ij} g_{jK}^T \delta_{KI}\right)\right] \varepsilon_j^{*(I)} - \lambda_j \Delta_j^* \square_I^* \varepsilon_j^{*(I)} = 0$$

$$I = 1-n \qquad j = 1-m \quad \text{(sum to k)} \tag{198}$$

In same way, we can obtain the modal magnetic field equations as follows

$$\rho \nabla_{tttt} \eta_I \gamma_I \xi_I H_I^{(j)*} + \nabla_{tt} \left[\rho \square_I^* - \Delta_j^* \gamma_I \xi_I \left(\lambda_j \eta_I + g_{Ij} g_{jK}^T \delta_{KI}\right)\right] H_I^{(j)*} - \lambda_j \Delta_j^* \square_I^* H_I^{(j)*} = 0$$

$$I = 1 - n \qquad j = 1 - m \quad \text{(sum to k)} \tag{199}$$

From the above results, we can see that when considering the mutual effects between mechanical subspaces and electromagnetic subspaces, the electromagnetic wave equations become four-order partial differential ones, meanwhile elastic wave equation still keep in the form of four-order partial differential ones.

6. Conclusions

In the first place, we analyzed here the elastic waves and electromagnetic waves in anisotropic solids. The calculation shows that the propagation of elastic waves in anisotropic solids consist of the incomplete dilation type and the incomplete shear type except for the pure longitudinal or pure transverse waves in isotropic solids. Several novel results for elastic waves were obtained, for example, there are two elastic waves in isotropic solids, which are the P-wave and the S-wave. There are three elastic waves in cubic solids, one of which is a quasi-P-waves and two are quasi-S-waves. There are four elastic waves in hexagonal (transversely isotropic) solids, half of which are quasi-P-waves and half of which are quasi-S-waves. There are five elastic waves in tetragonal solids, two of which are quasi-P-waves and three of which are quasi-S-waves. There are no more than six elastic waves in orthotropic solids or in the more complicated anisotropic solids. For electromagnetic waves, the similar results were obtained: 1) the number of electromagnetic waves in anisotropic media is equal to that of eigen-spaces of anisotropic media; 2) the velocity of propagation of electromagnetic waves is dependent on the eigen-dielectric permittivity and eigen-magnetic permeability; 3) the direction of propagation of electromagnetic waves is related on the eigen electromagnetic operator in the corresponding eigen-space; 4) the direction of polarization of electromagnetic waves is relevant to the eigen-electromagnetic quantities in the corresponding eigen-space. In another word, there is only one kind of electromagnetic wave in isotropic crystal. There are two kinds of electromagnetic waves in uniaxial crystal. There are three kinds of electromagnetic waves in biaxial crystal and three kinds of distorted electromagnetic waves in monoclinic crystal. Secondly, the elastic waves and electromagnetic waves in piezoelectric solids both for static theory and for fully dynamic theory are analyzed here based on the eigen spaces of physical presentation. The results show that the number and propagation speed of elastic or electromagnetic waves in anisotropic piezoelectric solids are determined by both the subspaces of electromagnetically anisotropic media and ones of mechanically anisotropic media. For the piezoelectric material of class 6mm, it is seen that there exist four elastic waves, respectively, but only two waves were affected by the piezoelectric coefficients. There exist two electromagnetic waves, respectively, but the two waves were all affected by the piezoelectric coefficients. The fully dynamic theory of Maxwell's equations, coupled to the mechanical equations of motion, are studied here. The complete set of uncoupled dynamic equations for piezoelectromagnetic waves in anisotropic media are deduced. For the piezoelectric material of class 6mm, it will be seen that there exist eight electromagnetic waves and also eight elastic waves, respectively. Furthermore, in fact of $c_I \gg v_j$, except for the classical electromagnetic waves and elastic waves, we can obtain the new electromagnetic waves propagated in speed of elastic waves and new elastic waves propagated in speed of electromagnetic waves.

7. References

Guo, S. H. (1999). Eigen theory of elastic dynamics for anisotropic solids. *Trans. Nonferr. Met. Soc. China*, 9., 2., 63–74, 1003-6326

Guo, S. H. (2000). Eigen theory of elastic mechanics for anisotropic solids. *Trans. Nonferr. Met. Soc. China*, 10., 2., 173–184, 1003-6326

Guo, S. H. (2001). Eigen theory of viscoelastic mechanics for anisotropic solids. *Acta Mech. Solida Sinica*, 22., 4., 368–374, 0894-9166

Guo, S. H. (2002). The standard expression of the engineering mechanical vector and the invariant of elastic mechanics. Acta Mech. Solida Sinica, 15., 1., 74-80, 0254-7805

Guo, S. H. (2005). Eigen theory of viscoelastic dynamics based on the Kelvin-Voigt model. *Appl. Math. Mech.,* 25., 7., 792–798, 0253-4827

Guo, S. H. (2008). An eigen theory of rheology for complex media. *Acta Mech.*, 198., 3/4., 253-260, 0001-5970

Guo, S. H. (2009). An eigen theory of electromagnetic waves based on the standard spaces. *Int. J. Engng. Sci.*, 47., 3., 405-412, 0020-7225

Guo, S. H. (2009). An Eigen theory of Static Electromagnetic Field for Anisotropic Media. Appl. Math. Mech., 30., 5., 528-533, 0253-4827

Guo, S. H. (2010). An eigen theory of waves in piezoelectric solids. *Acta Mech. Sinica*, 26., 2., 280-288, 0567-7718

Guo, S. H. (2010). An eigen theory of electro-magnetic acoustic waves in magnetoelectroelastic media. *Acta Mech.*, 211., 4., 173-180, 0001-5970

Guo, S. H. (2010). An eigen expression for piezoelectrically stiffened elastic and dielectric constants based on the static theory. *Acta Mech.*, 210., 3., 345-350, 0001-5970

Lee, P. C. Y.; Kim, Y. G.& Prevost, J. H. (1990). Electromagnetic radiation from doubly rotated piezoelectric crystal plates vibrating at thickness frequencies. *J. Appl. Phys.*, 67., 6633-6642, 0021-8979

Lee, P. C. Y. (1991). A variational principle for the equations of piezoelectromagnetism in elastic dielectric crystals. *J. Appl. Phys.*, 69., 7470-7473, 0021-8979

Li, S. (1996). The electromagneto-acoustic surface wave in a piezoelectric medium: the Bleustein-Gulyaev mode. *J. Appl. Phys.*, 80., 5264-5269, 0021-8979

Mindlin, R. D. (1973). Electromagnetic radiation from a vibrating quartz plate. *Int. J. Solids Struct.*, 9., 6., 697-702, 0020-7683

Palfreeman, J. S.(1965). Acoustic delay lines-A survey of types and uses. *Ultrasonics*, 3., 1., 1-8, 0041-624x

Sedov, A.; Schmerr, Jr.; et al. (1986). Some exact solutions for the propagation of transient electroacoustic waves I: piezoelectric half-space. *Int. J. Engng. Sci.*, 24., 4., 557-568, 0020-7225

Tiersten, H. F.; Mindlin, R. D. (1962). Forced vibrations of piezoelectric crystal plates. *Quart. Appl. Mech.,* 20., 2., 107-119, 0033-569x

Vavrycuk, V. (2005). Acoustic axes in weak triclinic anisotropy. *Geophysical Journal International*, 163., 2., 629-638, 0956-540x

Yakhno, V. G.; Yakhno, T. M. & Kasap, M. (2006). A novel approach for modeling and simulation of electromagnetic waves in anisotropic dielectrics. *Int. J. Solids Struct.*, 43., 20., 6261-6276, 0020-7683

Yang, J.S.; Guo, S. H. (2006). Piezoelectromagnetic waves guided by the surface of a ceramic cylinder. *Acta Mech.*, 181., 3/4., 199-205, 0001-5970

2

Reverberation-Ray Matrix Analysis of Acoustic Waves in Multilayered Anisotropic Structures

Yongqiang Guo[1] and Weiqiu Chen[2]
[1]*Key Laboratory of Mechanics on Disaster and Environment in Western China, Ministry of Education, and School of Civil Engineering and Mechanics, Lanzhou University*
[2]*Department of Engineering Mechanics, Zhejiang University*
P.R.China

1. Introduction

Natural multilayers can be frequently observed, like the layered soils and rocks for example (Kausel & Roesset, 1981; Kennett, 1983). They are also increasingly used as artificial materials and structures in engineering practices for their high performances (Nayfeh, 1995). For instances, cross-ply and fibrous laminated composites have been applied in naval vessels, aeronautical and astronautical vehicles, and so on for the sake of high strength and light weight (Nayfeh, 1995); piezoelectric thin film systems have been used in various surface acoustic wave (SAW) and bulk acoustic wave (BAW) devices in electronics and information technology in order to accomplish smaller size, lower energy consumption, higher operating frequency and sensitivity, greater bandwidth, and enhanced reception characteristics (Auld, 1990; Adler, 2000). Consequently, as a widespread category of inhomogeneous materials and structures, multilayered structures deserve special concern about their mechanical and acoustical behavior, especially the dynamic behavior since it is what these structures differ most markedly from the homogeneous materials and structures.
Investigation of acoustic wave propagation in multilayered structures plays an essential role in understanding their dynamic behavior, which is the main concern in design, optimization, characterization and nondestructive evaluation of multilayered composites (Lowe, 1995; Chimenti, 1997; Rose, 1999) and acoustic wave devices (Auld, 1990; Rose, 1999; Adler, 2000). Nevertheless, the top and bottom surfaces and the interfaces in a multilayered structure cause reflection and/or transmission of elastic waves, giving rise to coupling of various fundamental wave modes in adjacent layers. In multilayered structures consisting of anisotropic media, even the fundamental wave modes themselves are mutually coupled in each layer (Achenbach, 1973). As a result, the analysis of acoustic waves in multilayered structures always remains an extraordinary complex problem, and it is very difficult to obtain a simple and yet numerically well-performed, closed-form analytical solution for a general multilayered structure.
For the above reasons, various matrix formulations have been developed for the analysis of acoustic wave propagation in multilayered media from diverse domains (Ewing et al., 1957; Brekhovskikh, 1980; Kennett, 1983; Lowe, 1995; Nayfeh, 1995; Rose, 1999), since the beginning of this research subject in the midst of last century. Most of these matrix methods

were presented initially for multilayered structures consisting of isotropic (transversely isotropic) materials, and then extended to those structures made of anisotropic elastic and piezoelectric layers. These matrix methods fall into two groups. One group is those numerical methods based on discrete models, such as the boundary element method (BEM) (Makkonen, 2005), the finite difference method (FDM) (Igel et al., 1995; Makkonen, 2005), the finite element method (FEM) (Datta et al., 1988; Makkonen, 2005) and the hybrid method of BEM and FEM (BEM/FEM) (Makkonen, 2005). This group of methods is powerful for modeling acoustic waves in multilayered structures with various geometries and boundaries. However, they have the disadvantage that the results are approximate, and particularly certain high frequency components must be thrown off in any discrete model. The accuracy of the computational results and the stability of the numerical algorithms depend greatly on the discretization in the temporal and spatial domains. Calculation efficiency will be dramatically decreased if higher accuracy is pursued. The other group is those analytical methods based on continuous (distributed-parameter) model, among which the transfer matrix method (TMM) (Lowe, 1995), also referred to as the propagator matrix method (PMM) (Alshits & Maugin, 2008), is the typical one. TMM (Thomson, 1950; Haskell, 1953; Lowe, 1995; Nayfeh, 1995; Adler, 1990, 2000) leads to a system equation with dimension keeping small and unchanged as the number of layers increases, since in the formulation the basic unknowns of the intermediate layers are eliminated by matrix products. Thus, TMM has the advantage of high accuracy and high efficiency in most cases, but it suffers from numerical instability in the case of high frequency-thickness products (Nayfeh, 1995; Adler, 1990, 2000; Lowe, 1995; Tan, 2007). Aiming at circumventing this kind of numerical difficulty, different variant forms of TMM as well as analytical matrix methods have been proposed, including the stiffness matrix method (Kausel & Roesset, 1981; Shen et al., 1998; Wang & Rokhlin, 2001, 2002a; Rokhlin & Wang, 2002a; Tan, 2005), the spectral element method (Rizzi & Doyle, 1992; Chakraborty & Gopalakrishnan, 2006), the surface impedance matrix method (Honein et al., 1991; Degettekin et al., 1996; Zhang et al., 2001; Hosten & Castaings, 2003; Collet, 2004), the hybrid compliance/stiffness matrix method (Rokhlin & Wang, 2002b; Wang & Rokhlin, 2004a; Tan, 2006), the recursive asymptotic stiffness matrix method (Wang & Rokhlin, 2002b, 2004b, 2004c), the scattering matrix method (Pastureaud et al., 2002) and the compound matrix method (Fedosov et al., 1996), for instances. Tan (Tan, 2007) compared some of these methods in mathematical algorithm, computational efficiency and numerical stability. However, most of these alternative formulations lack uniformity in a certain degree, and are computationally complicated and inefficient, especially for the high frequency analysis.

Lately, Pao and his coworkers (Pao et al, 2000; Su et al., 2002; Tian et al., 2006) developed the method of reverberation-ray matrix (MRRM) for evaluating the transient wave propagation in layered isotropic and transversely isotropic media. It is shown that MRRM has many advantages and its comparison to the TMM in various aspects was discussed by Pao et al. (Pao et al., 2007). However, the original formulation of MRRM is based on the wave potential functions, which confines this numerically stable and uniform matrix method from extending to layered anisotropic structures. In fact, it is impossible to use wave potential functions for an arbitrarily anisotropic medium, in which the fundamental wave modes are mutually coupled (Achenbach, 1973). Thus, Guo and Chen (Guo & Chen, 2008a, 2008b; Guo, 2008; Guo et al., 2009) presented a new formulation of MRRM based on state-space formalism and plane wave expansion for the analysis of free waves in anisotropic elastic and piezoelectric layered media.

The objective of this chapter is to present the general and unified formulation of the method of reverberation-ray matrix (MRRM) for the analysis of acoustic wave propagation in multilayered structures of arbitrarily anisotropic elastic and piezoelectric media based on the state-space formalism and Fourier transforms. In Section 2, the state equation for each layer made of an arbitrarily anisotropic elastic/piezoelectric material is derived from the three-dimensional linear theory of elasticity/piezoelectricity with the help of Fourier transforms, and the solution to the state equation boils down to an eigenvalue problem from which the propagation constants and characteristic mode coefficients can be obtained numerically for a specified frequency. Then the traveling wave solution to the state equation can be written in explicit form in terms of unknown amplitudes as well as known propagation constants and characteristic mode coefficients. In Section 3, we show how the multilayered anisotropic structure is described in both the global and the local dual coordinates. From the boundary conditions on the upper and lower surfaces with applied external forces and the continuity conditions at the interfaces, the scattering relation, which expresses one group of equations for the unknown wave amplitudes in dual local coordinates, is appropriately constructed such that matrix inversion is avoided. Due to the uniqueness of physical essence, the two solutions expressed in dual local coordinates should be compatible with each other, leading to the phase relation, which represents the other group of equations for the unknown wave amplitudes in dual local coordinates. Care must be taken of to properly establish the phase relation such that all exponentially growing functions are excluded. The number of simultaneous equations from the phase and scattering relations amount exactly to the number of unknown wave amplitudes in dual local coordinates, and hence the wave solution can be determined. To reduce the dimension, we substitute the phase relation into the scattering relation to obtain a system equation, from which the dispersion relation for free wave propagation is obtained by letting the determinant of coefficient matrix vanish, and the steady-state and transient wave propagation due to the external force excitations can be obtained by inverse Fourier transforms. Section 4 gives numerical examples of guided wave propagation in cross-ply elastic composite structures. Dispersion curves for different configurations, various boundary conditions and in particular at the high frequency range are illustrated to show the versatility and numerical stability of the proposed formulation of MRRM. Effects of configurations and boundaries on the dispersion spectra are clearly demonstrated through comparison. Conclusions are drawn in Section 5, with highlights of advantages of the proposed general formulation of MRRM for characterizing the acoustic waves in multilayered anisotropic structures.

2. State space formalism for anisotropic elastic and piezoelectric layers

2.1 Governing equations and state vectors

Consider a homogeneous, arbitrarily anisotropic elastic medium. From the three-dimensional linear elasticity (Synge, 1956; Stroh, 1962; Nayfeh, 1995) we have the constitutive relations

$$\sigma_{ij} = c_{ijkl}\varepsilon_{kl} \tag{1}$$

the strain-displacement relations

$$\varepsilon_{kl} = (u_{k,l} + u_{l,k})/2 \tag{2}$$

and the equations of motion (in absence of body forces)

$$\sigma_{ij,j} = \rho \ddot{u}_i \tag{3}$$

where the comma in the subscripts and superposed dot imply spatial and time derivatives, σ_{ij}, ε_{kl} and u_i are respectively the stress, strain and displacement tensors, c_{ijkl} are the elastic constants, and ρ is the material density. The dynamic governing equations can be simplified by eliminating the strain tensor to

$$\sigma_{ij} = c_{ijkl}(u_{k,l} + u_{l,k})/2 \qquad \sigma_{ij,j} = \rho \ddot{u}_i \tag{4}$$

It is commonly difficult to obtain solutions to Eq. (4) for an anisotropic medium of the most general kind as there are 21 independent elastic constants in total, and the deformations in different directions and of different kinds are coupled. However, for an arbitrarily anisotropic elastic layer, the state space formulation (Tarn, 2002a) can be established by grouping the field variables properly. Assume that the correspondence between the digital indices and coordinates follows a usual rule, i.e. $1 \to x$, $2 \to y$, and $3 \to z$. If the z axis is along the thickness direction of the laminate, we divide the stresses into two groups: the first consists of the components on the surface of z=const., and the second consists of the remaining components. The combination of the displacements $\mathbf{v}_u = [u, v, w]^{\mathrm{T}}$ and the first group of stresses $\mathbf{v}_\sigma = [\tau_{zx}, \tau_{zy}, \sigma_z]^{\mathrm{T}}$ gives the state vector $\mathbf{v} = [(\mathbf{v}_u)^{\mathrm{T}}, (\mathbf{v}_\sigma)^{\mathrm{T}}]^{\mathrm{T}}$.
For piezoelectric materials of the most general kind, in the catalogue of three-dimensional linear theory (Ding & Chen, 2001), we have the constitutive relations instead of Eq. (1)

$$\sigma_{ij} = c_{ijkl}\varepsilon_{kl} - e_{kij}E_k, \; D_i = e_{ikl}\varepsilon_{kl} + \beta_{ik}E_k \tag{5}$$

the strain-displacement relations of Eq. (2) are further supplemented by

$$E_k = -\varphi_{,k} \tag{6}$$

and the equations of motion in Eq.(3) are complemented with (in absence of free charges)

$$D_{i,i} = 0 \tag{7}$$

where D_i, E_k and φ are respectively the electric displacement, field and potential tensors, and e_{kij} and β_{ik} are the piezoelectric and permittivity constants, respectively. In view of Eqs. (2), (6) and (5), the dynamic governing equations become

$$\begin{cases} \sigma_{ij} = c_{ijkl}(u_{k,l} + u_{l,k})/2 + e_{kij}\varphi_{,k} \\ D_i = e_{ikl}(u_{k,l} + u_{l,k})/2 - \beta_{ik}\varphi_{,k} \end{cases} \qquad \begin{cases} \sigma_{ij,j} = \rho \ddot{u}_i \\ D_{i,i} = 0 \end{cases} \tag{8}$$

where the coupling between the mechanical and electrical fields is clearly seen. It is noted that the independent piezoelectric and permittivity constants of arbitrarily piezoelectric media should be 18 and 6 respectively, adding further complexity to the solution procedure. However, for an anisotropic piezoelectric layer of the most general kind, the state space formalism (Tarn, 2002b) can also be established just as for arbitrarily anisotropic elastic layer. This will be illustrated in the following section. For piezoelectric materials, the state

vector is defined by $\mathbf{v}=[(\mathbf{v}_u)^{\mathrm{T}},(\mathbf{v}_\sigma)^{\mathrm{T}}]^{\mathrm{T}}$, with $\mathbf{v}_u=[u,v,w,\varphi]^{\mathrm{T}}$ being the generalized displacements and $\mathbf{v}_\sigma=[\tau_{zx},\tau_{zy},\sigma_z,D_z]^{\mathrm{T}}$ the first group of generalized stresses.

2.2 Fourier transforms and state equations

By virtue of the triple Fourier transform pairs as follows

$$\hat{f}(k_x;k_y,z;\omega)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y,z,t)\mathrm{e}^{-\mathrm{i}k_x x}\mathrm{e}^{-\mathrm{i}k_y y}\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t \tag{9}$$

$$f(x,y,z,t)=(\frac{1}{2\pi})^3\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\hat{f}(k_x;k_y,z;\omega)\mathrm{e}^{\mathrm{i}k_x x}\mathrm{e}^{\mathrm{i}k_y y}\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{d}k_x\,\mathrm{d}k_y\,\mathrm{d}\omega \tag{10}$$

the generalized displacements and stresses as well as dynamic governing equations given in Eqs. (4) and (8) in the time-space domain can be transformed into those in the frequency-wavenumber domain, where ω is the circular frequency; k_x and k_y are the wavenumbers in the x and y directions, respectively; $\mathrm{i}=\sqrt{-1}$ is the unit imaginary; and the z-dependent variable in the frequency-wavenumber domain is indicated with an over caret.

By eliminating the second group of generalized stresses, the transformed Eqs. (4) and (8) in a right-handed coordinate system can be reduced to a system of first-order ordinary differential equations with respect to the state vector, which contains $n_v/2$ generalized displacement components and $n_v/2$ generalized stress components, as follows

$$\frac{\mathrm{d}\hat{\mathbf{v}}(z)}{\mathrm{d}z}=\mathbf{A}\hat{\mathbf{v}}(z) \tag{11}$$

which is usually referred to as the state equation. The coefficient matrix $\mathbf{A}$ of order $n_v\times n_v$, with all elements being functions of the material constants, the circular frequency ω or the wavenumbers k_x and k_y, can be written in a blocked form

$$\mathbf{A}=\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{bmatrix} \tag{12}$$

where

$$\begin{gathered}\mathbf{A}_{11}=-\mathrm{i}\mathbf{G}_{33}^{-1}\mathbf{W},\ \mathbf{A}_{12}=\mathbf{G}_{33}^{-1},\ \mathbf{A}_{22}=-\mathrm{i}\mathbf{W}^{\mathrm{T}}\mathbf{G}_{33}^{-1}\\ \mathbf{A}_{21}=-\rho\omega^2\mathbf{M}+k_x^2\mathbf{G}_{11}+k_y^2\mathbf{G}_{22}+k_xk_y\left(\mathbf{G}_{12}+\mathbf{G}_{21}\right)-\mathbf{W}^{\mathrm{T}}\mathbf{G}_{33}^{-1}\mathbf{W}\end{gathered} \tag{13}$$

with $\mathbf{W}=k_x\mathbf{G}_{31}+k_y\mathbf{G}_{32}$. For a layer of arbitrarily anisotropic elastic material, we have

$$n_v=6,\ \mathbf{G}_{kl}=\begin{bmatrix}c_{1k1l} & c_{1k2l} & c_{1k3l}\\ c_{2k1l} & c_{2k2l} & c_{2k3l}\\ c_{3k1l} & c_{3k2l} & c_{3k3l}\end{bmatrix},\ \mathbf{M}=\mathbf{I}_3 \tag{14}$$

while for a layer of arbitrarily anisotropic piezoelectric material, we have

$$n_v = 8\,,\ \mathbf{G}_{kl} = \begin{bmatrix} c_{1k1l} & c_{1k2l} & c_{1k3l} & e_{l1k} \\ c_{2k1l} & c_{2k2l} & c_{2k3l} & e_{l2k} \\ c_{3k1l} & c_{3k2l} & c_{3k3l} & e_{l3k} \\ e_{k1l} & e_{k2l} & e_{k3l} & -\beta_{kl} \end{bmatrix},\ \mathbf{M} = \begin{bmatrix} \mathbf{I}_3 & 0 \\ 0 & 0 \end{bmatrix} \tag{15}$$

where $\mathbf{I}_3$ denotes the identity matrix of order 3.

2.3 Traveling wave solutions to the state equation

According to the theory of ordinary differential equation (Coddington & Levinson, 1955), the solution to the state equation (11) can be expressed as

$$\hat{\mathbf{v}}(z) = \mathbf{\Phi}\exp(\mathbf{\Lambda}z)\mathbf{w} = \begin{Bmatrix} \hat{\mathbf{v}}_u(z) \\ \hat{\mathbf{v}}_\sigma(z) \end{Bmatrix} = \begin{Bmatrix} \mathbf{\Phi}_u \\ \mathbf{\Phi}_\sigma \end{Bmatrix}\exp(\mathbf{\Lambda}z)\mathbf{w} \tag{16}$$

where $\exp(\cdot)$ denotes the matrix exponential function; $\mathbf{\Lambda}$ and $\mathbf{\Phi}$ are the $n_v \times n_v$ diagonal eigenvalue matrix and square eigenvector matrix of the coefficient matrix $\mathbf{A}$, respectively; $\mathbf{\Phi}_u$ and $\mathbf{\Phi}_\sigma$ are the $n_v/2 \times n_v$ sub-matrices of $\mathbf{\Phi}$ corresponding to the generalized displacement and stress vectors, respectively; and $\mathbf{w}$ is the vector of undetermined coefficients with n_v components.

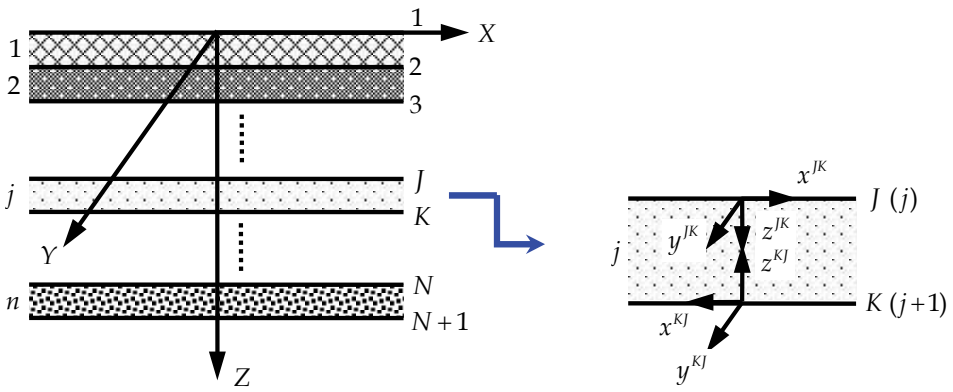
When combined with the common factors in the integrand of inverse transform in Eq. (10), the solutions in Eq. (16) are interpreted as the total response corresponding to harmonic plane traveling waves with different wavenumbers $-\lambda_i$ in the z direction ($i = 1,2,\cdots,n_v$) and common wavenumbers $-k_x$ and $-k_y$ in the x and y-directions at a common radian frequency ω, where λ_i is the i th component of the diagonal wavenumber matrix $\mathbf{\Lambda}$. The corresponding undermined coefficient w_i denotes the wave amplitude, and the components of the corresponding eigenvector $\mathbf{\Phi}_i$ give the state variable response coefficients of the ith wave. According to the nature of wavenumber $-\lambda_i$, the characteristic waves can be divided into two groups, the n_a arriving waves with their z-axis projection along the negative direction and the n_d departing waves with their z-axis projection along the positive direction (Ingebrigtsen & Tonning, 1969). Denote the respective sub eigenvalue matrices as $\mathbf{\Lambda}_-$ (of order $n_a \times n_a$) and $\mathbf{\Lambda}_+$ (of order $n_d \times n_d$), both being diagonal, the respective sub eigenvector matrices as $\mathbf{\Phi}_-$ (of order $n_v \times n_a$) and $\mathbf{\Phi}_+$ (of order $n_v \times n_d$), and the respective amplitude vectors as $\mathbf{a}$ (the arriving wave vector) and $\mathbf{d}$ (the departing wave vector). Obviously, we have $\mathbf{w} = [\mathbf{a}^{\mathrm{T}}, \mathbf{d}^{\mathrm{T}}]^{\mathrm{T}}$. Therefore, the matrices $\mathbf{\Lambda}_-$ and $\mathbf{\Phi}_-$ and the vector $\mathbf{a}$ correspond to the eigenvalues λ_i, which satisfy $\mathrm{Re}(-\lambda_i) < 0$ or $\mathrm{Re}(-\lambda_i) = 0, \mathrm{Im}(-\lambda_i) < 0$, while the matrices $\mathbf{\Lambda}_+$ and $\mathbf{\Phi}_+$ and the vector $\mathbf{d}$ are associated with the remaining eigenvalues. It is easily seen that we always have $n_a + n_d = n_v$ with $n_v = 6$ for elastic materials and $n_v = 8$ for piezoelectric materials. Consequently, the solution to the state equation given in Eq. (16) can be rewritten as

$$\begin{aligned} \hat{\mathbf{v}}(z) &= \begin{bmatrix} \mathbf{\Phi}_- & \mathbf{\Phi}_+ \end{bmatrix} \begin{bmatrix} \exp(\mathbf{\Lambda}_- z) & \mathbf{0} \\ \mathbf{0} & \exp(\mathbf{\Lambda}_+ z) \end{bmatrix} \begin{Bmatrix} \mathbf{a} \\ \mathbf{d} \end{Bmatrix} \\ &= \begin{Bmatrix} \hat{\mathbf{v}}_u(z) \\ \hat{\mathbf{v}}_\sigma(z) \end{Bmatrix} = \begin{bmatrix} \mathbf{\Phi}_{u-} & \mathbf{\Phi}_{u+} \\ \mathbf{\Phi}_{\sigma-} & \mathbf{\Phi}_{\sigma+} \end{bmatrix} \begin{bmatrix} \exp(\mathbf{\Lambda}_- z) & \mathbf{0} \\ \mathbf{0} & \exp(\mathbf{\Lambda}_+ z) \end{bmatrix} \begin{Bmatrix} \mathbf{a} \\ \mathbf{d} \end{Bmatrix} \end{aligned} \tag{17}$$

where $\mathbf{\Phi}_{u-}$ and $\mathbf{\Phi}_{\sigma-}$ are $n_v/2\times n_a$ sub eigenvector matrices of $\mathbf{\Phi}_{-}$ corresponding to the generalized displacement and stress vectors, respectively; $\mathbf{\Phi}_{u+}$ and $\mathbf{\Phi}_{\sigma+}$ are those $n_v/2\times n_d$ sub eigenvector matrices of $\mathbf{\Phi}_{+}$. It is noted from Eq. (17) that the only unknowns in the solutions are the wave amplitudes, which should be determined from the system equation formulated by simultaneously considering the dynamic state of all constituent layers of the structure and their interactions. This will be shown in the following section within the framework of reverberation-ray matrix analysis.

3. Unified formulation of MRRM

The schematic of a multilayered anisotropic structure of infinite lateral extent is depicted in Fig. 1, which consists of a perfect stacking of n homogeneous arbitrarily anisotropic elastic or piezoelectric layers. From up to down, the layers are denoted in order by numbers 1 to n, and the top surface, interfaces and bottom surface in turn are denoted by numbers 1 to $N+1$, bearing the fact that $n=N$. The top and bottom surfaces are denoted by 1 and $N+1$ respectively, and the upper and lower bounding faces of an arbitrary layer, j for instance, are respectively denoted by J and K, with $J=j$ and $K=j+1$ and the layer will be referred to as JK or KJ.



(a) Description of the laminated structure in global coordinates

(b) Description of a typical layer j in local dual coordinates

Fig. 1. The schematic of an *n*-layered anisotropic laminate

3.1 Description of the structural system

A global coordinate system (X,Y,Z) with its origin located on the top surface and the Z-axis along the thickness direction, as shown in Fig. 1a, is established for the system analysis of the whole structure. The physical variables associated with a surface/interface will be described in the global coordinates, and will be affixed with a single superscript, J or K for example, to indicate their affiliation.

As a unique feature of MRRM, a pair of dual local coordinates (x^{JK},y^{JK},z^{JK}) and (x^{KJ},y^{KJ},z^{KJ}) is used to describe each layer JK (or KJ), with z^{JK} originating from J to K and z^{KJ} from K to J, y^{JK} and y^{KJ} in the same direction as the Y-axis, and x^{JK} in the same and x^{KJ} in the opposite direction of the X-axis, as shown in Fig. 1b. The physical variables

inside the layers will be described in the local dual coordinates and double superscripts, JK or KJ for instance, will be affixed to any physical quantity to denote the corresponding coordinate system and the pertaining layer. As an example, $\hat{\mathbf{v}}^{JK}$ and $\hat{\mathbf{v}}^{KJ}$ are the state vectors for layer JK (or KJ) in the coordinates (x^{JK}, y^{JK}, z^{JK}) and (x^{KJ}, y^{KJ}, z^{KJ}), respectively. To make the sign convection more clear, physical variables are deemed to be positive as it is along the positive direction of the pertinent coordinate axis.

It is seen from Fig. 1b that the dual local coordinates are both right-handed, thus the state equations in Eq. (11) and the traveling wave solutions in Eqs. (16) and (17) all come into existence for an arbitrary layer JK (or KJ) in (x^{JK}, y^{JK}, z^{JK}) and (x^{KJ}, y^{KJ}, z^{KJ}), which are written as

$$\frac{\mathrm{d}\hat{\mathbf{v}}^{JK}(z^{JK})}{\mathrm{d}z^{JK}} = \mathbf{A}^{JK}\hat{\mathbf{v}}^{JK}(z^{JK}) \tag{18}$$

$$\frac{\mathrm{d}\hat{\mathbf{v}}^{KJ}(z^{KJ})}{\mathrm{d}z^{KJ}} = \mathbf{A}^{KJ}\hat{\mathbf{v}}^{KJ}(z^{KJ}) \tag{19}$$

$$\hat{\mathbf{v}}^{JK}(z^{JK}) = \mathbf{\Phi}^{JK}\exp(\mathbf{\Lambda}^{JK}z^{JK})\mathbf{w}^{JK} = \begin{bmatrix}\mathbf{\Phi}_{-}^{JK} & \mathbf{\Phi}_{+}^{JK}\end{bmatrix}\begin{bmatrix}\exp\left(\mathbf{\Lambda}_{-}^{JK}z^{JK}\right) & \mathbf{0}\\ \mathbf{0} & \exp\left(\mathbf{\Lambda}_{+}^{JK}z^{JK}\right)\end{bmatrix}\begin{Bmatrix}\mathbf{a}^{JK}\\ \mathbf{d}^{JK}\end{Bmatrix} \tag{20}$$

$$\hat{\mathbf{v}}^{KJ}(z^{KJ}) = \mathbf{\Phi}^{KJ}\exp(\mathbf{\Lambda}^{KJ}z^{KJ})\mathbf{w}^{KJ} = \begin{bmatrix}\mathbf{\Phi}_{-}^{KJ} & \mathbf{\Phi}_{+}^{KJ}\end{bmatrix}\begin{bmatrix}\exp\left(\mathbf{\Lambda}_{-}^{KJ}z^{KJ}\right) & \mathbf{0}\\ \mathbf{0} & \exp\left(\mathbf{\Lambda}_{+}^{KJ}z^{KJ}\right)\end{bmatrix}\begin{Bmatrix}\mathbf{a}^{KJ}\\ \mathbf{d}^{KJ}\end{Bmatrix} \tag{21}$$

From Eqs. (20) and (21) we see that there are totally $n_v \times N$ arriving wave amplitudes and $n_v \times N$ departing wave amplitudes for all layers in dual local coordinates, which should be determined by $2n_v \times N$ relations. It is deduced that the basic unknowns (wave amplitudes) in the MRRM double in number due to the particular description of dynamic state in dual local coordinates, as compared with that in other analytical methods which are usually based on single local coordinates. However, by doing so in the MRRM, the boundary conditions on surfaces and continuous conditions at interfaces take on an extremely simple form since the exponential functions in the solutions no longer appear, as will be seen from Section 3.2. Furthermore, as will be shown in Section 3.3, the arriving and departing wave amplitudes in dual local coordinates are related directly from the point of view of wave propagation through the layer. Thus it shall be possible to deduce a system equation in terms of only the departing wave vectors of all layers. In such a case, the dimension of the system equation will be the same as the one of other analytical methods based on single local coordinates, such as the stiffness matrix method and the spectral element method, as discussed in Section 3.4.

3.2 Scattering relation from coupling conditions on surfaces and at interfaces

3.2.1 Local scattering relations of top and bottom surfaces

As depicted in Fig. 2, the response of state variables on the top and bottom surfaces, $\hat{\mathbf{v}}^{12}(0)$ and $\hat{\mathbf{v}}^{(N+1)N}(0)$, corresponding to various waves in the top and bottom layers, respectively,

should be in accordance with the external state variables $\hat{\mathbf{v}}_E^1$ $(=[(\hat{\mathbf{v}}_{uE}^1)^{\mathrm{T}},(\hat{\mathbf{v}}_{\sigma E}^1)^{\mathrm{T}}]^{\mathrm{T}})$ and $\hat{\mathbf{v}}_E^{(N+1)}$ $(=[(\hat{\mathbf{v}}_{uE}^{(N+1)})^{\mathrm{T}},(\hat{\mathbf{v}}_{\sigma E}^{(N+1)})^{\mathrm{T}}]^{\mathrm{T}})$, i.e.

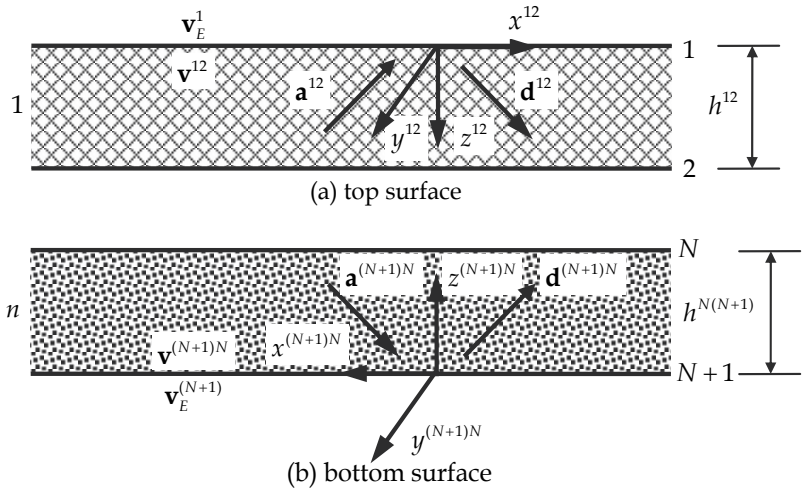


Fig. 2. The top and bottom surfaces of the multilayered anisotropic structure

$$\hat{\mathbf{v}}^{12}(0)=\mathbf{T}_E\hat{\mathbf{v}}_E^1,\ \mathbf{T}_v\hat{\mathbf{v}}^{(N+1)N}(0)=\mathbf{T}_E\hat{\mathbf{v}}_E^{(N+1)} \tag{22}$$

where $\mathbf{T}_E=<\mathbf{T}_{uE},\mathbf{T}_{\sigma E}>$ is a transformation matrix with $\mathbf{T}_{uE}=\mathbf{I}_{n_v/2}$ and $\mathbf{T}_{\sigma E}=-\mathbf{I}_{n_v/2}$; $\mathbf{T}_v=<\mathbf{T}_u,\mathbf{T}_\sigma>$ is also a transformation matrix, with $\mathbf{T}_u=<-1,1,-1>$ and $\mathbf{T}_\sigma=<1,-1,1>$ for elastic layers and $\mathbf{T}_u=<-1,1,-1,-1>$ and $\mathbf{T}_\sigma=<1,-1,1,1>$ for piezoelectric layers. Here $<\cdot>$ denotes the (block) diagonal matrix with elements (or sub-matrices) only on the main diagonal and $\mathbf{I}_{n_v/2}$ represents the identity matrix of order $n_v/2$.

By virtue of Eqs. (20) and (21), the solutions to $\hat{\mathbf{v}}^{12}(0)$ and $\hat{\mathbf{v}}^{(N+1)N}(0)$ can be obtained as

$$\hat{\mathbf{v}}^{12}(0)=\mathbf{\Phi}^{12}\mathbf{w}^{12}=\begin{bmatrix}\mathbf{\Phi}_-^{12} & \mathbf{\Phi}_+^{12}\end{bmatrix}\begin{Bmatrix}\mathbf{a}^{12}\\ \mathbf{d}^{12}\end{Bmatrix} \tag{23}$$

$$\hat{\mathbf{v}}^{(N+1)N}(0)=\mathbf{\Phi}^{(N+1)N}\mathbf{w}^{(N+1)N}=\begin{bmatrix}\mathbf{\Phi}_-^{(N+1)N} & \mathbf{\Phi}_+^{(N+1)N}\end{bmatrix}\begin{Bmatrix}\mathbf{a}^{(N+1)N}\\ \mathbf{d}^{(N+1)N}\end{Bmatrix} \tag{24}$$

where the exponential functions disappear since the thickness coordinates on the surfaces are always zero in the corresponding local coordinates. This is the main advantage of introducing the dual local coordinates. It should be noticed that half of the components of vectors $\hat{\mathbf{v}}_E^1$ and $\hat{\mathbf{v}}_E^{(N+1)}$ are known, which are denoted by vectors $\hat{\mathbf{v}}_K^1$ and $\hat{\mathbf{v}}_K^{(N+1)}$, respectively, while the remaining half are unknown, denoted by vectors $\hat{\mathbf{v}}_U^1$ and $\hat{\mathbf{v}}_U^{(N+1)}$, respectively.

Substituting Eqs. (23) and (24) into Eq. (22), we can derive

$$\mathbf{A}^{12}\mathbf{a}^{12}+\mathbf{D}^{12}\mathbf{d}^{12}=\mathbf{T}_K^1\hat{\mathbf{v}}_K^1=\mathbf{s}_0^1,\ \mathbf{A}^{(N+1)N}\mathbf{a}^{(N+1)N}+\mathbf{D}^{(N+1)N}\mathbf{d}^{(N+1)N}=\mathbf{T}_K^{(N+1)}\hat{\mathbf{v}}_K^{(N+1)}=\mathbf{s}_0^{(N+1)} \tag{25}$$

where $\mathbf{A}^{12}$, $\mathbf{D}^{12}$, $\mathbf{A}^{(N+1)N}$, $\mathbf{D}^{(N+1)N}$, $\mathbf{T}_K^1$ ($\mathbf{T}_K^{(N+1)}$) are the coefficient matrices with components extracted, in accordance with $\hat{\mathbf{v}}_K^1$ and $\hat{\mathbf{v}}_K^{(N+1)}$, from $\mathbf{\Phi}_-^{12}$, $\mathbf{\Phi}_+^{12}$, $\mathbf{\Phi}_-^{(N+1)N}$, $\mathbf{\Phi}_+^{(N+1)N}$ and $\mathbf{T}_E$ ($\mathbf{T}_E$ and $\mathbf{T}_v$) respectively. $\mathbf{s}_0^1$ and $\mathbf{s}_0^{(N+1)}$ are excitation source vectors with $n_v/2$ components of the top and bottom surfaces, respectively. Particularly as far as free waves are concerned, if the top surface is mechanically traction-free (and electrically open-circuit), we have

$$\hat{\mathbf{v}}_K^1 = \hat{\mathbf{v}}_{\sigma E}^1 = [\hat{\tau}_X^1, \hat{\tau}_Y^1, \hat{\sigma}_Z^1]^{\mathrm{T}} = \mathbf{0} \ (\text{or } \hat{\mathbf{v}}_K^1 = \hat{\mathbf{v}}_{\sigma E}^1 = [\hat{\tau}_X^1, \hat{\tau}_Y^1, \hat{\sigma}_Z^1, \hat{D}_Z^1]^{\mathrm{T}} = \mathbf{0}) \tag{26}$$

$$\mathbf{A}^{12} = \mathbf{\Phi}_{\sigma-}^{12}, \ \mathbf{D}^{12} = \mathbf{\Phi}_{\sigma+}^{12} \tag{27}$$

and when the top surface is mechanically fixed (and electrically closed-circuit) we have

$$\hat{\mathbf{v}}_K^1 = \hat{\mathbf{v}}_{uE}^1 = [\hat{u}_X^1, \hat{u}_Y^1, \hat{u}_Z^1]^{\mathrm{T}} = \mathbf{0} \ (\text{or } \hat{\mathbf{v}}_K^1 = \hat{\mathbf{v}}_{uE}^1 = [\hat{u}_X^1, \hat{u}_Y^1, \hat{u}_Z^1, \hat{\varphi}_Z^1]^{\mathrm{T}} = \mathbf{0}) \tag{28}$$

$$\mathbf{A}^{12} = \mathbf{\Phi}_{u-}^{12}, \ \mathbf{D}^{12} = \mathbf{\Phi}_{u+}^{12} \tag{29}$$

For mixed boundary conditions, the form of known quantities and coefficient matrices can also be worked out accordingly. The boundary conditions on the bottom surface can be similarly deduced and will not be discussed for brevity.

Eq. (25) can be further written in a form of local scattering relations on the top and bottom surfaces

$$\mathbf{A}^1\mathbf{a}^1 + \mathbf{D}^1\mathbf{d}^1 = \mathbf{s}_0^1, \ \mathbf{A}^{N+1}\mathbf{a}^{N+1} + \mathbf{D}^{N+1}\mathbf{d}^{N+1} = \mathbf{s}_0^{(N+1)} \tag{30}$$

where $\mathbf{a}^1 = \mathbf{a}^{12}$ ($\mathbf{a}^{N+1} = \mathbf{a}^{(N+1)N}$) and $\mathbf{d}^1 = \mathbf{d}^{12}$ ($\mathbf{d}^{N+1} = \mathbf{d}^{(N+1)N}$) are the arriving and departing wave vectors of the top (bottom) surface, $\mathbf{A}^1 = \mathbf{A}^{12}$ ($\mathbf{A}^{N+1} = \mathbf{A}^{(N+1)N}$) and $\mathbf{D}^1 = \mathbf{D}^{12}$ ($\mathbf{D}^{N+1} = \mathbf{D}^{(N+1)N}$) are $n_v/2 \times n_a^{12}$ ($n_v/2 \times n_a^{(N+1)N}$) and $n_v/2 \times n_d^{12}$ ($n_v/2 \times n_d^{(N+1)N}$) coefficient matrices corresponding to the arriving and departing wave vectors of the top (bottom) surface, respectively.

It should be pointed out that the form of local scattering relations at the boundaries given in Eq. (30) is also valid for surface waves in a multilayered structure.

3.2.2 Local scattering relations of a typical interface

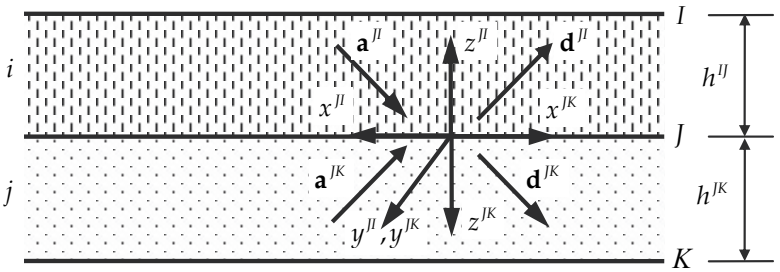


Fig. 3. A typical interface J of the multilayered anisotropic structure

Since the adjacent layers in the structure are perfectly bonded, the state variables should be continuous across the interfaces. Taking the typical interface J as shown in Fig. 2 for illustration, the compatibility of the generalized displacements and equilibrium of the generalized forces require

$$\mathbf{T}_v \hat{\mathbf{v}}^{Jl}(0) = \hat{\mathbf{v}}^{JK}(0) \tag{31}$$

This gives, according to the solutions in Eqs. (20) and (21),

$$\begin{bmatrix} \mathbf{T}_u \mathbf{\Phi}^{Jl}_{u-} & \mathbf{T}_u \mathbf{\Phi}^{Jl}_{u+} \\ \mathbf{T}_\sigma \mathbf{\Phi}^{Jl}_{\sigma-} & \mathbf{T}_\sigma \mathbf{\Phi}^{Jl}_{\sigma+} \end{bmatrix} \begin{Bmatrix} \mathbf{a}^{Jl} \\ \mathbf{d}^{Jl} \end{Bmatrix} = \begin{bmatrix} \mathbf{\Phi}^{JK}_{u-} & \mathbf{\Phi}^{JK}_{u+} \\ \mathbf{\Phi}^{JK}_{\sigma-} & \mathbf{\Phi}^{JK}_{\sigma+} \end{bmatrix} \begin{Bmatrix} \mathbf{a}^{JK} \\ \mathbf{d}^{JK} \end{Bmatrix} \tag{32}$$

It should be noticed once again that there is no exponential functions in the coupling equation (32) for interfaces. By grouping the arriving and departing wave vectors of relevant layers into the local arriving and departing wave vectors of the interface $\mathbf{a}^J = [(\mathbf{a}^{Jl})^{\mathrm{T}}, (\mathbf{a}^{JK})^{\mathrm{T}}]^{\mathrm{T}}$ and $\mathbf{d}^J = [(\mathbf{d}^{Jl})^{\mathrm{T}}, (\mathbf{d}^{JK})^{\mathrm{T}}]^{\mathrm{T}}$, Eq. (32) is reduced to the local scattering relation of the typical interface J

$$\mathbf{A}^J \mathbf{a}^J + \mathbf{D}^J \mathbf{d}^J = \mathbf{0} \tag{33}$$

where the $n_v \times (n_a^{Jl} + n_a^{JK})$ and $n_v \times (n_d^{Jl} + n_d^{JK})$ coefficient matrices $\mathbf{A}^J$ and $\mathbf{D}^J$, respectively, are

$$\mathbf{A}^J = \begin{bmatrix} \mathbf{T}_u \mathbf{\Phi}^{Jl}_{u-} & -\mathbf{\Phi}^{JK}_{u-} \\ \mathbf{T}_\sigma \mathbf{\Phi}^{Jl}_{\sigma-} & -\mathbf{\Phi}^{JK}_{\sigma-} \end{bmatrix}, \ \mathbf{D}^J = \begin{bmatrix} \mathbf{T}_u \mathbf{\Phi}^{Jl}_{u+} & -\mathbf{\Phi}^{JK}_{u+} \\ \mathbf{T}_\sigma \mathbf{\Phi}^{Jl}_{\sigma+} & -\mathbf{\Phi}^{JK}_{\sigma+} \end{bmatrix} \tag{34}$$

There are altogether N – 1 (n – 1) interfaces in the multilayered structure, so that we have $N-1$ $(n-1)$ local scattering equations like Eq. (33).

3.2.3 Global scattering relation of the structure

The local scattering relations of top surface, interfaces and bottom surface have respectively $n_v/2$, $n_v \times (N-1)$ and $n_v/2$ equations, which are grouped together from up to down to give the $n_v \times N$ global scattering relation

$$\mathbf{A}\mathbf{a} + \mathbf{D}\mathbf{d} = \mathbf{s}_0 \tag{35}$$

where the global arriving and departing wave vectors $\mathbf{a}$ and $\mathbf{d}$ are

$$\begin{aligned} \mathbf{a} &= \left[(\mathbf{a}^{12})^{\mathrm{T}}, (\mathbf{a}^{21})^{\mathrm{T}}, (\mathbf{a}^{23})^{\mathrm{T}}, \cdots, (\mathbf{a}^{Jl})^{\mathrm{T}}, (\mathbf{a}^{JK})^{\mathrm{T}}, \cdots, (\mathbf{a}^{N(N+1)})^{\mathrm{T}}, (\mathbf{a}^{(N+1)N})^{\mathrm{T}} \right]^{\mathrm{T}} \\ \mathbf{d} &= \left[(\mathbf{d}^{12})^{\mathrm{T}}, (\mathbf{d}^{21})^{\mathrm{T}}, (\mathbf{d}^{23})^{\mathrm{T}}, \cdots, (\mathbf{d}^{Jl})^{\mathrm{T}}, (\mathbf{d}^{JK})^{\mathrm{T}}, \cdots, (\mathbf{d}^{N(N+1)})^{\mathrm{T}}, (\mathbf{d}^{(N+1)N})^{\mathrm{T}} \right]^{\mathrm{T}} \end{aligned} \tag{36}$$

the corresponding $(n_v \times N) \times (n_v \times N)$ coefficient matrices $\mathbf{A}$ and $\mathbf{D}$ are

$$\mathbf{A} = < \mathbf{A}^1, \mathbf{A}^2, \cdots, \mathbf{A}^J, \cdots, \mathbf{A}^{N+1} >, \ \mathbf{D} = < \mathbf{D}^1, \mathbf{D}^2, \cdots, \mathbf{D}^J, \cdots, \mathbf{D}^{N+1} > \tag{37}$$

and $\mathbf{s}_0 = \left[(\mathbf{s}_0^1)^{\mathrm{T}}, \mathbf{0}^{\mathrm{T}}, \cdots, \mathbf{0}^{\mathrm{T}}, (\mathbf{s}_0^{(N+1)})^{\mathrm{T}} \right]^{\mathrm{T}}$ is the global excitation source vector. It should be noted that the forming process of scattering relations in Eqs. (30), (33) and (35) exclude matrix inversion as compared to that in the original formulation of MRRM (Pao et al, 2000, 2007; Su

et al., 2002; Tian et al., 2006), which guarantees the numerical stability and at the same time enables the inclusion of surface and interface wave modes, in the proposed formulation of MRRM (Guo & Chen, 2008a, 2008b; Guo, 2008; Guo et al., 2009).

3.3 Phase relation from compatibility conditions of layers

3.3.1 Local phase relation of a typical layer

Considering the formation of the dual local coordinates of a typical layer *JK* (*KJ*) as discussed in Section 3.1, we have the geometrical dual transformation relations

$$x^{JK} = -x^{KJ},\ y^{JK} = y^{KJ},\ z^{JK} = h^{JK} - z^{KJ},\ \mathrm{d}z^{JK} = -\mathrm{d}z^{KJ} \tag{38}$$

where h^{JK} $(= h^{KJ})$ represents the thickness of layer *JK* (*KJ*), and the physical dual transformation relations

$$\hat{\mathbf{v}}^{JK}\left(z^{JK}\right) = \mathbf{T}_v \hat{\mathbf{v}}^{KJ}\left(z^{KJ}\right) \tag{39}$$

By virtue of Eqs. (18) and (19), Eqs. (38) and (39), and the definitions of eigenvalue and eigenvector, it is derived that

$$\mathbf{A}^{KJ}(z^{KJ}) = -\mathbf{T}_v \mathbf{A}^{JK}(z^{JK})\mathbf{T}_v^{-1},\ \mathbf{\Lambda}^{JK} = -\mathbf{\Lambda}^{KJ},\ \mathbf{\Phi}^{JK} = \mathbf{T}_v \mathbf{\Phi}^{KJ} \tag{40}$$

It is interpreted that if λ^{JK} and $\boldsymbol{\varphi}_\lambda^{JK}$ are the eigenvalue and eigenvector of the coefficient matrix $\mathbf{A}^{JK}$, then $-\lambda^{JK}$ and $\mathbf{T}_v\boldsymbol{\varphi}_\lambda^{JK}$ must be the corresponding eigenvalue and eigenvector of the coefficient matrix $\mathbf{A}^{KJ}$. The equality relations between the numbers of arriving and departing waves in dual local coordinates, i.e. $n_a^{JK} = n_d^{KJ}$ and $n_d^{JK} = n_a^{KJ}$, are also implied.
Substituting Eqs. (20) and (21) into Eq. (39), and in view of Eq. (40) and $\mathbf{T}_v = \mathbf{T}_v^{-1}$, one obtains the local phase relation of a typical layer *JK* (*KJ*)

$$\begin{Bmatrix} \mathbf{a}^{JK} \\ \mathbf{a}^{KJ} \end{Bmatrix} = \begin{bmatrix} \exp\left(-\mathbf{\Lambda}_-^{JK} h^{JK}\right) & \mathbf{0} \\ \mathbf{0} & \exp\left(\mathbf{\Lambda}_+^{JK} h^{JK}\right) \end{bmatrix} \begin{Bmatrix} \mathbf{d}^{KJ} \\ \mathbf{d}^{JK} \end{Bmatrix} = \begin{bmatrix} \mathbf{P}^{JK} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}^{KJ} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{I}_a^{JK} \\ \mathbf{I}_d^{JK} & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \mathbf{d}^{JK} \\ \mathbf{d}^{KJ} \end{Bmatrix} \tag{41}$$

where the $n_a^{JK} \times n_a^{JK}$ and $n_d^{JK} \times n_d^{JK}$ diagonal matrices $\mathbf{P}^{JK} = \exp(-\mathbf{\Lambda}_-^{JK} h^{JK})$ and $\mathbf{P}^{KJ} = \exp(\mathbf{\Lambda}_+^{JK} h^{JK})$ are referred to as local phase matrices, and $\mathbf{I}_a^{JK}$ and $\mathbf{I}_d^{JK}$ are identity matrices of order n_a^{JK} and n_d^{JK}, respectively. It should be noted that the exponentially growing functions, which usually cause numerical instability (such as in the TMM) for large values of the frequency-thickness product, have been completely excluded from the phase matrices $\mathbf{P}^{JK}$ and $\mathbf{P}^{KJ}$, since we always have $\mathrm{Re}(\lambda_-^{JK} h^{JK}) > 0$ or $\mathrm{Re}(\lambda_-^{JK} h^{JK}) = 0,\ \mathrm{Im}(\lambda_-^{JK} h^{JK}) > 0$ $(\mathrm{Re}(\lambda_+^{JK} h^{JK}) < 0$ or $\mathrm{Re}(\lambda_+^{JK} h^{JK}) = 0,\ \mathrm{Im}(\lambda_+^{JK} h^{JK}) < 0)$. As indicated by Eq. (41), there are n_v equations in the local phase relation of each layer.

3.3.2 Global phase relation of the structure

Grouping together the local phase relations for all layers from up to down yields the global phase relation with $n_v \times N$ equations

$$\mathbf{a} = \mathbf{P}\bar{\mathbf{d}} = \mathbf{P}\mathbf{U}\mathbf{d} \tag{42}$$

where the $(n_v \times N) \times (n_v \times N)$ block diagonal matrices $\mathbf{P}$, named the global phase matrix, is composed of

$$\mathbf{P} =< \mathbf{P}^{12}, \mathbf{P}^{21}, \mathbf{P}^{23}, \cdots, \mathbf{P}^{Jl}, \mathbf{P}^{JK}, \cdots, \mathbf{P}^{N(N+1)}, \mathbf{P}^{(N+1)N} > \tag{43}$$

the variant of the global departing wave vector $\bar{\mathbf{d}}$ is related to the wave vector $\mathbf{d}$ by the $(n_v \times N)\times(n_v \times N)$ block diagonal matrix $\mathbf{U}$, which is referred to as the global permutation matrix, to account for the different sequence of components arrangement between $\bar{\mathbf{d}}$ and $\mathbf{d}$. The specific forms of $\mathbf{U}$ and $\bar{\mathbf{d}}$ are as follows

$$\mathbf{U} =< \mathbf{U}^{12}, \mathbf{U}^{23}, \cdots, \mathbf{U}^{JK}, \cdots, \mathbf{U}^{N(N+1)} >,\ \mathbf{U}^{JK}_{n_v \times n_v} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_a^{JK} \\ \mathbf{I}_d^{JK} & \mathbf{0} \end{bmatrix} \tag{44}$$

$$\bar{\mathbf{d}} = \left[(\mathbf{d}^{21})^{\mathrm{T}}, (\mathbf{d}^{12})^{\mathrm{T}}, (\mathbf{d}^{32})^{\mathrm{T}}, \cdots, (\mathbf{d}^{IJ})^{\mathrm{T}}, (\mathbf{d}^{KJ})^{\mathrm{T}}, \cdots, (\mathbf{d}^{(N+1)N})^{\mathrm{T}}, (\mathbf{d}^{N(N+1)})^{\mathrm{T}} \right]^{\mathrm{T}} \tag{45}$$

It is seen from Eq. (42) that the global arriving and departing wave vectors **a** and **d**, consisting of respectively the arriving and departing wave amplitudes in local dual coordinates of all layers and having the same forms as those in the global scattering relation in Eq. (36), are related directly through the global phase relation, which enables the dimension reduction of the system equation, making the final scale the same as the one in other analytical methods which are based on single local coordinates.

3.4 System equation and dispersion equation

The global scattering relation in Eq. (35) and global phase relation in Eq. (42) contain respectively $n_v \times N$ equations for the $n_v \times N$ unknown arriving wave amplitudes (in $\mathbf{a}$) and $n_v \times N$ unknown departing wave amplitudes (in $\mathbf{d}$). Thus the wave vectors can be determined. Substitution of Eq. (42) into Eq. (35) gives the system equation

$$(\mathbf{APU} + \mathbf{D})\mathbf{d} = \mathbf{Rd} = \mathbf{s}_0 \tag{46}$$

where $\mathbf{R} = \mathbf{APU} + \mathbf{D}$ is the system matrix.

If there is no surface excitation, i.e. $\mathbf{s}_0 = 0$ and the free wave propagation problem is considered, the vanishing of the system matrix determinant yields the following dispersion equation

$$\left|\mathbf{R}(k_x; k_y; \omega)\right| = \mathbf{0} \tag{47}$$

which may be solved numerically by a proper root searching technique (Guo, 2008). Thus, the complete dispersion curves of various waves can be obtained, as illustrated in Section 4 for multilayered anisotropic elastic structures.

If there is surface excitation, from Eq. (46) we have

$$\mathbf{d} = (\mathbf{APU} + \mathbf{D})^{-1}\mathbf{s}_0 = \mathbf{R}^{-1}\mathbf{s}_0 \tag{48}$$

Further making use of the global phase relation (42), the solution of the state vector in Eq. (20) and the inverse Fourier transform in Eq. (10) with respect to the wavenumbers, the steady-state response of state variables of a layer at circular frequency ω can be expressed as

$$\begin{aligned} \hat{\mathbf{v}}^{JK}(x^{JK}, y^{JK}, z^{JK}; \omega) &= \frac{1}{(2\pi)^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \hat{\mathbf{v}}^{JK}(k_x; k_y, z^{JK}; \omega)\mathrm{e}^{\mathrm{i}(k_x x^{JK} + k_y y^{JK})}\,\mathrm{d}k_x\,\mathrm{d}k_y \\ &= \frac{1}{(2\pi)^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \left\{ \mathbf{\Phi}_-^{JK} \exp(\mathbf{\Lambda}_-^{JK} z^{JK})\mathbf{E}^{KJ} + \mathbf{\Phi}_+^{JK} \exp(\mathbf{\Lambda}_+^{JK} z^{JK})\mathbf{E}^{JK} \right\} \mathbf{R}^{-1}\mathbf{s}_0\, \mathrm{e}^{\mathrm{i}(k_x x^{JK} + k_y y^{JK})}\,\mathrm{d}k_x\,\mathrm{d}k_y \end{aligned} \tag{49}$$

and the transient response should be obtained from the corresponding steady-state response by means of inverse Fourier transform with respect to frequency as

$$\begin{aligned}\mathbf{v}^{JK}(x^{JK},y^{JK},z^{JK},t) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\hat{\mathbf{v}}^{JK}(x^{JK},y^{JK},z^{JK};\omega)\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{d}\omega \\ &= \frac{1}{(2\pi)^3}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left\{\boldsymbol{\Phi}_{-}^{JK}\exp(\boldsymbol{\Lambda}_{-}^{JK}z^{JK})\mathbf{E}^{KJ}+\boldsymbol{\Phi}_{+}^{JK}\exp(\boldsymbol{\Lambda}_{+}^{JK}z^{JK})\mathbf{E}^{JK}\right\}\mathbf{R}^{-1}\mathbf{s}_0\,\mathrm{e}^{\mathrm{i}(k_x x^{JK}+k_y y^{JK}+\omega t)}\,\mathrm{d}k_x\,\mathrm{d}k_y\,\mathrm{d}\omega\end{aligned} \tag{50}$$

Eqs. (49) and (50) can be solved numerically by the FFT technique (Guo, 2008).

4. Numerical examples

To verify the proposed method, we consider laminated composites with reinforced fibers alternately aligning along the *x*-axis (0°) and the *y*-axis (90°). The material properties of a lamina with fibers in 0° and 90° directions are given in Table 1.

Lamina	ρ	c_{11}	c_{12}	c_{13}	c_{22}	c_{23}	c_{33}	c_{44}	c_{55}	c_{66}
0°	1200	1.6073	0.0644	0.0644	0.1392	0.0692	0.1392	0.035	0.0707	0.0707
90°	1200	0.1392	0.0644	0.0692	1.6073	0.0644	0.1392	0.0707	0.035	0.0707

Note: The unit of density is kg/m^3 and that of stiffness is 10^{11} N/m^2.

Table 1. Material properties of 0° and 90° laminas

In the previous work (Guo & Chen, 2008a), dispersion curves have been calculated for a single layer of 0° lamina and a triple-layered structure with a 0°/90°/0° configuration with equal thickness of each lamina. The obtained results were compared with those obtained by the finite element stiffness method (Datta et al., 1988) (Figs. 4 and 8 for single layer and Figs. 5 and 9 for triple layers therein). The excellent agreement validates our derivation and the computer codes. Nevertheless, it should be pointed out that the proposed MRRM is analytical, based on continuous (distributed-parameter) model. Thus, it can give more accurate results but at less computational expense especially in the high-frequency range. Recently, the characteristics of free waves in single PZT-4 and/or barium sodium niobate (BSN) layers have been discussed (Guo et al., 2009) for different boundary conditions, and dispersion curves of bi-layered, triple-layered and ten-layered piezoelectric structures composed of alternate PZT-4 and BSN layers with equal thickness were also presented (Guo & Chen, 2008b; Guo et al., 2009).

Here in this chapter, a four-layered composite with a 0°/90°/0°/90° configuration is considered. The waves are assumed to propagate in the X-direction for illustration, i.e. $k_x = k$ and $k_y = 0$. Note that the formulations established in previous sections are valid for waves propagating in any direction in the XOY plane. For the sake of presentation, we define the dimensionless quantities, including frequency Ω, wavenumber γ, wavelength L and phase velocity V by $\Omega = \omega H/(2\pi c_s)$, $\gamma = kH/(2\pi)$, $L = 2\pi\lambda/H = 4\pi^2/(kH)$ and $V = c/c_s$, respectively, where H is the total thickness of the composite plate, k, λ and c are respectively the wavenumber, wavelength and phase velocity, $c_s = \sqrt{(c_{55}/\rho)_{0°}}$ is the shear wave velocity constant with $(c_{55})_{0°}$ and $(\rho)_{0°}$ the stiffness coefficient and material density of 0° lamina. The thickness of the 0° and 90° laminas are denoted as h_1 and h_2, respectively.

4.1 Dispersion curves of multilayered anisotropic structures with free surfaces

First, the laminas of the four-layered composite structure are assumed to have equal thicknesses and the top and bottom surfaces of the composite are assumed to be traction-free. The dispersion curves, in terms of frequency-wavenumber spectra, wavelength-frequency spectra and phase velocity-frequency spectra, are presented in Figs. 4(a), 4(b) and 4(c) respectively.

The sub-figures (a) to (c) in Fig. 4 show similar dispersion properties of free waves in the four-layered composite as compared with those for single and triple layers. The quasi P-SV and SH bulk modes, surface and interface modes and characteristic asymptotic line are obtained all at once from the dispersion equation by a root searching algorithm.

4.2 Dispersion curves in high frequency range

The frequency-wavenumber, wavelength-frequency and phase velocity-frequency spectra with dimensionless frequency Ω in the range of 100 to 102 are given in Figs. 5(a), 5(b) and 5(c) respectively, which indicate the proposed formulation of MRRM can assure a good numerical stability in the high-frequency range. Since the wave modes at small values of wavelength and phase velocity are relatively intensive and difficult to differentiate within this frequency range, as implied in Figs. 4(b) and 4(c), the dimensionless wavelength L and phase velocity V are specified within 0.75~2.00 and 10~50 in Figs. 5(b) and 5(c), respectively.

4.3 Effects of configuration on the dispersion curves

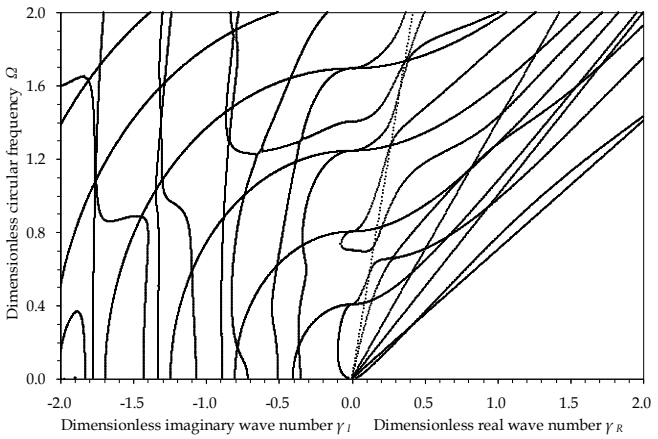
Next, the thickness of the 0° and 90° laminas of the four-layered composite are assumed to be unequal in order to study the effect of configuration on the characteristics of free waves. The dispersion curves for cases $h_1 / h_2 = 1/4$ and $h_1 / h_2 = 4/1$ as well as their comparison with those for the equal thickness case (denoted as $h_1 / h_2 = 1/1$) are depicted in Fig. 6, with the frequency-wavenumber, wavelength-frequency and phase velocity-frequency spectra given in the sub-figures (a), (b) and (c) respectively.

It is seen from Fig. 6 that the dispersion curves of a specified wave mode corresponding to the case $h_1 / h_2 = 1/1$ locate in between those for cases $h_1 / h_2 = 1/4$ and $h_1 / h_2 = 4/1$. As also indicated in Fig. 6, the thickness ratio has a distinct effect on the characteristics of all free wave modes. The effect is however somehow larger for the higher-order wave modes than the lower-order ones.

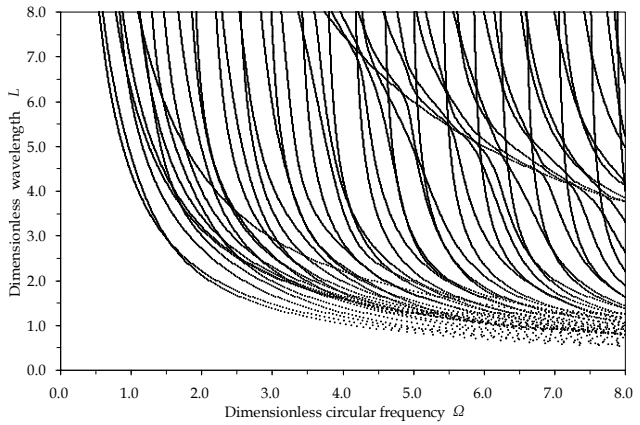
4.4 Effects of boundary conditions on the dispersion curves

In order to show the effects of boundary conditions on the dispersion characteristics, the same 0°/90°/0°/90° laminated composite with equal layer thicknesses is considered for two different boundary conditions: one is that both the top and bottom surfaces are fixed and the other is that the top surface is traction-free while the bottom surface is fixed. The dispersion curves for the two cases are given in Fig. 7 and compared with those for a laminate with free surfaces.

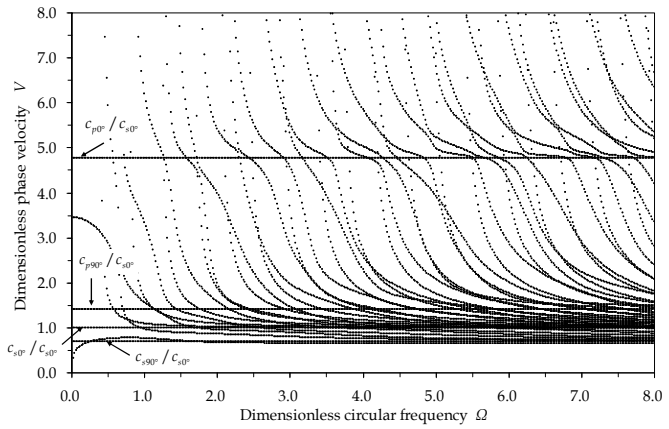
It is seen from Fig. 7 that some parts of the dispersion curves of certain specified modes for the four-layered composite with different surface conditions may coincide, but they may be completely different at other parts or for other modes. Fig. 7 indicates that the boundary conditions have a complex effect on the dispersion characteristics of free waves in multilayered anisotropic structures. In-depth study is needed.



(a) Frequency-wavenumber spectra

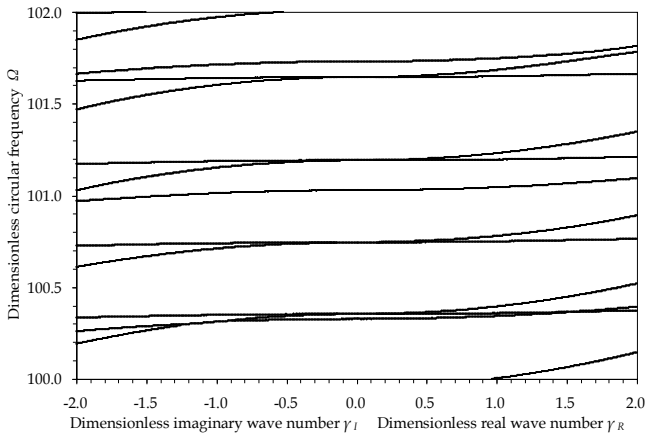


(b) Wavelength-frequency spectra

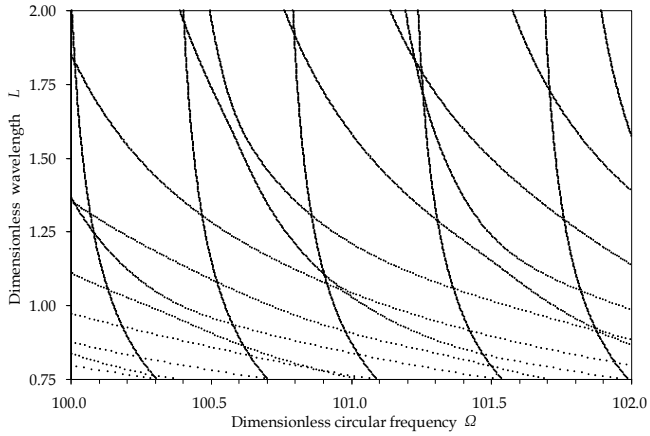


(c) Phase velocity-frequency spectra

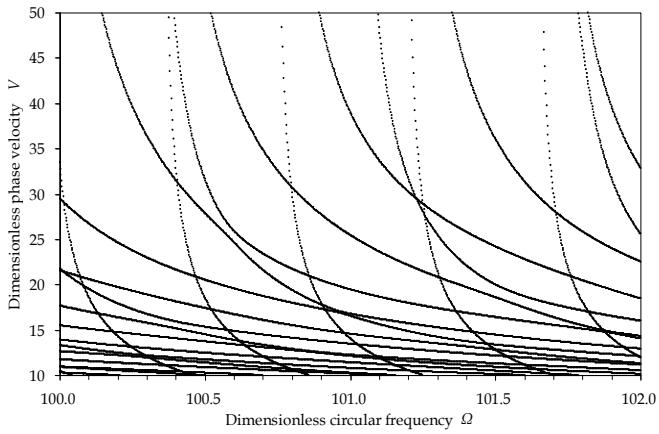
Fig. 4. Dispersion curves of the 0°/90°/0°/90° laminated composite with free surfaces



(a) Frequency-wavenumber spectra

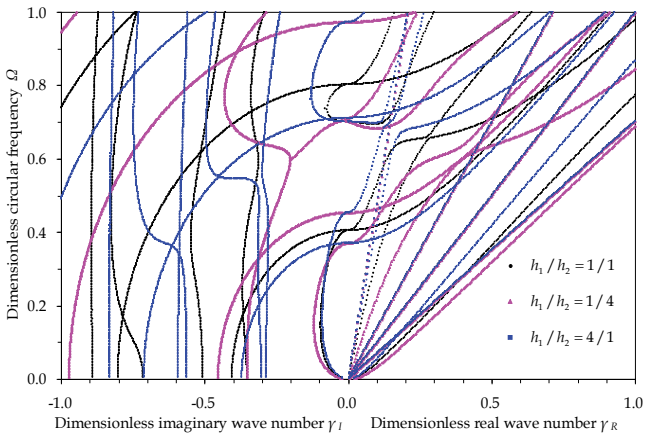


(b) Wavelength-frequency spectra

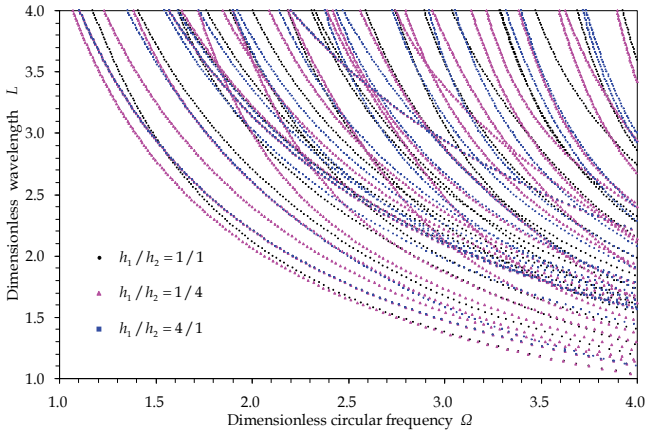


(c) Phase velocity-frequency spectra

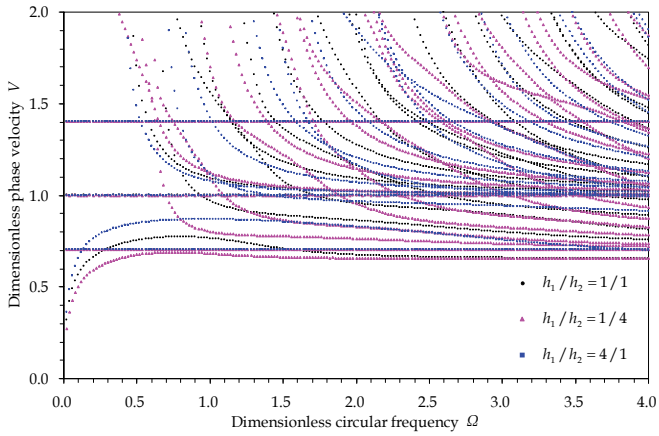
Fig. 5. Dispersion curves at high frequency of the four-layered composite with free surfaces



(a) Frequency-wavenumber spectra

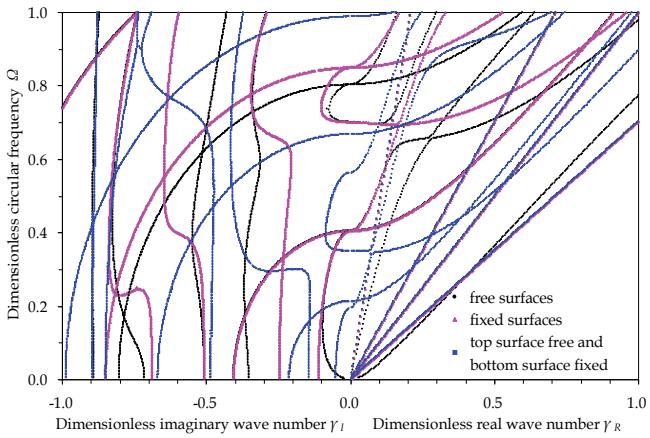


(b) Wavelength-frequency spectra

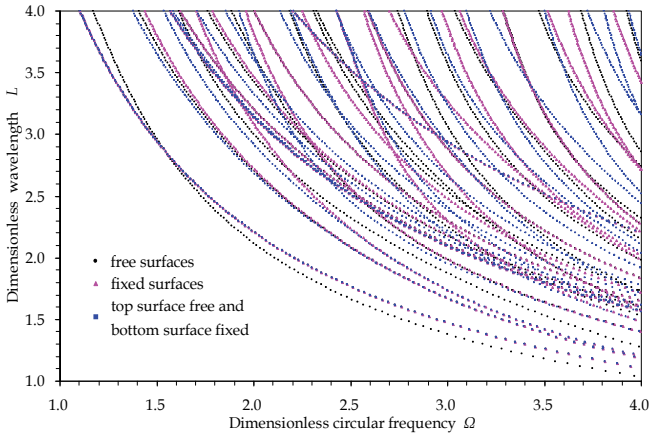


(c) Phase velocity-frequency spectra

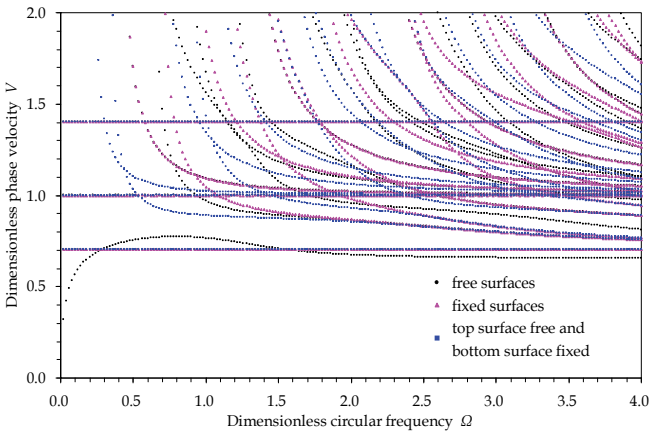
Fig. 6. Comparisons of dispersion curves of the composite with different configurations



(a) Frequency-wavenumber spectra



(b) Wavelength-frequency spectra



(c) Phase velocity-frequency spectra

Fig. 7. Comparison of dispersion curves of the composite with different boundary conditions

5. Conclusion

We present a unified formulation of the method of reverberation-ray matrix (MRRM) for the analysis of acoustic wave propagation in multilayered anisotropic elastic/piezoelectric structures based on the state space formalism and Fourier transforms in the framework of three-dimensional elasticity or piezoelectricity. The proposed formulation of MRRM includes all wave modes in the structure and possesses good numerical stability by properly excluding exponentially growing function and matrix inversion operation. It is therefore suitable for the accurate analysis of acoustic waves in complex multilayered anisotropic structures by a uniform computer program. In comparison with the well-known traditional transfer matrix method, the present MRRM is unconditionally numerically stable, irrespective of the total number of layers, the thickness of individual layers and the frequency. Besides, in comparison with the numerical methods based on discrete models, the present MRRM is based on a continuous model (distributed-parameter model) and gives accurate results at a much smaller computational cost especially in the high-frequency range. Numerical results indicate a high accuracy and broad versatility of the proposed formulation of MRRM for wave propagation in multilayered anisotropic structures with various configurations and boundary conditions in any frequency range. The obtained dispersion curves and their dependence on the structural configurations and boundary conditions shall be useful in the design and optimization of laminated composites and acoustic wave devices.

6. Acknowledgements

This study was financially supported by the National Natural Science Foundation of China (No. 10902045 and No. 10725210) and the Postdoctoral Science Foundation of China (No. 20090460155) and the Fundamental Research Funds for the Central Universities of China.

7. References

Achenbach, J. D. (1973). *Wave Propagation in Elastic Solids*, North-Holland, Amsterdam.

Coddington, E. A., Levinson, N. (1955). *Theory of Ordinary Differential Equations*, McGraw-Hill, New York.

Adler, E. L. (1990). Matrix methods applied to acoustic waves in multilayers, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 37(6): 485-490.

Adler, E. L. (2000). Bulk and surface acoustic waves in anisotropic solids, *International Journal of High Speed Electronics and Systems*, 10(3): 653-684.

Alshits V. I. & Maugin, G. A. (2008). Dynamics of anisotropic multilayers, *Wave Motion*, 45(5): 629-640.

Auld, B. A. (1990). *Acoustic Fields and Waves in Solids*, Second Edition, Vol. I & II, Robert E. Krieger Publishing Company, Malabar, Florida.

Brekhovskikh, L. M. (1980). *Waves in Layered Media*, Second Edition. Academic Press, New York.

Chakraborty, A. & Gopalakrishnan, S. (2006). A spectral finite element model for wave propagation analysis in laminated composite plate, *ASME Journal of Vibration and Acoustics*, 128: 477-488.

Chimenti, D. E. (1997). Guided waves in plates and their use in materials characterization, *Applied Mechanics Reviews*, 50(5): 247-284.

Collet, B. (2004). Recursive surface impedance matrix methods for ultrasonic wave propagation in piezoelectric multilayers, *Ultrasonics*, 42: 189-197.

Datta, S. K., Shah, A. H., Bratton, R. L. & Chakraborty, T. (1988). Wave propagation in laminated composite plates, *Journal of the Acoustical Society of America*, 83(6): 2020-2026.

Degettekin, F. L., Honein, B. V., & Khuri-Yakub, B. T. (1996). Application of surface impedance approach to ultrasonic wave propagation in layered anisotropic media, *IEEE Ultrasonics Symposium*, 559-562.

Ding, H. J. & Chen, W. Q. (2001). *Three Dimensional Problems of Piezoelasticity*. Nova Science Publishers, New York.

Ewing, W. M., Jardetzky, W. S. & Press, F. (1957). *Elastic Waves in Layered Media*. McGraw-Hill, New York.

Fedosov, V. I., Aniiimkin, V. I., Kotelyanskii, I. M., Caliendo, C., Verardi, P. & Verona, E. (1996). Analysis of acoustic waves in multilayers using compound matrices, *IEEE Ultrasonics Symposium*, 207-212.

Guo, Y. Q. (2008). *The Method of Reverberation-Ray Matrix and its Applications*, Doctorial dissertation. Zhejiang University, Hangzhou, China. (in Chinese)

Guo, Y. Q. & Chen, W. Q. (2008a). On free wave propagation in anisotropic layered media, *Acta Mechanica Solida Sinica*, 21(6): 500-506.

Guo, Y. Q. & Chen, W. Q. (2008b) Modeling of multilayered acoustic wave devices with the method of reverberation-ray matrix, *2008 Symposium on Piezoelectricity, Acoustic Waves, and Device Applications* (SPAWDA 2008), 105-110.

Guo, Y. Q., Chen, W. Q. & Zhang, Y. L. (2009). Guided wave propagation in multilayered piezoelectric structures, *Science in China*, *Series G*: *Physics*, *Mechanics and Astronomy*, 52(7): 1094-1104.

Haskell, N. A. (1953). The dispersion of surface waves on multilayered media, *Bulletin of the Seismological Society of America*, 43: 17-34.

Honein, B., Braga, A. M. B., Barbone, P. & Herrmann, G. (1991). Wave propagation in piezoelectric layered media with some applications, *Journal of Intelligent Material Systems and Structures*, 2: 542-557.

Hosten, B. & Castaings, M. (2003). Surface impedance matrices to model the propagation in multilayered media, *Ultrasonics*, 41: 501-507.

Igel, H., Mora, P. & Riollet, B. (1995). Anisotropic wave propagation through finite-difference grids, *Geophysics*, 60(4): 1203-1216.

Ingebrigtsen, K. A. & Tonning, A. (1969). Elastic surface waves in crystal, *Physical Review*, 184: 942-951.

Kausel, E. & Roesset, J. M. (1981). Stiffness matrices for layered soils, *Bulletin of the Seismological Society of America*, 71: 1743-1761.

Kennett, B. L. N. (1983). *Seismic Wave Propagation in Stratified Media*, Cambridge University Press, Cambridge.

Lowe, M. J. S. (1995). Matrix techniques for modeling ultrasonic waves in multilayered media, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 42: 525-542.

Makkonen, T. (2005). *Numerical Simulations of Microacoustic Resonators and Filters*, Doctoral Dissertation. Helsinki University of Technology, Espoo, Finland.

Nayfeh, A. H. (1995). *Wave Propagation in Layered Anisotropic Media with Applications to Composites*. Elsevier, Amsterdam.

Pao, Y. H., Chen, W. Q. & Su, X. Y. (2007). The reverberation-ray matrix and transfer matrix analyses of unidirectional wave motion, *Wave Motion*, 44: 419-438.

Pao, Y. H., Su, X. Y. & Tian, J. Y. (2000). Reverberation matrix method for propagation of sound in a multilayered liquid, *Journal of Sound and Vibration*, 230(4): 743-760.

Pastureaud, T., Laude, V. & Ballandras S. (2002). Stable scattering-matrix method for surface acoustic waves in piezoelectric multilayers, *Applied Physics Letters*, 80: 2544-2546.

Rizzi, S. A. & Doyle, J. F. (1992). A spectral element approach to wave motion in layered solids, *ASME Journal of Vibration and Acoustics*, 114: 569-577.

Rokhlin, S. I. & Wang, L. (2002a). Stable recursive algorithm for elastic wave propagation in layered anisotropic media: Stiffness matrix method, *Journal of the Acoustical Society of America*, 112: 822-834.

Rokhlin, S. I. & Wang, L. (2002b). Ultrasonic waves in layered anisotropic media: characterization of multidirectional composites, *International Journal of Solids and Structures*, 39(16): 4133-4149.

Rose, J. L. (1999). *Ultrasonic Waves in Solid Media*, Cambridge University Press, Cambridge.

Shen, S. P., Kuang, Z. B. & Hu, S.L. (1998). Wave propagation in multilayered anisotropic media, *Mechanics Research Communications*, 25(5): 503-507.

Stroh, A. N. (1962). Steady state problems in anisotropic elasticity, *Journal of Mathematics and Physics*, 41: 77-103.

Su, X. Y., Tian, J. Y. & Pao, Y. H. (2002). Application of the reverberation-ray matrix to the propagation of elastic waves in a layered solid, *International Journal of Solids and Structures*, 39: 5447-5463.

Synge, J. L. (1956). Flux of energy for elastic waves in anisotropic media, *Proceedings of the Royal Irish Academy*, 58: 13-21.

Tan, E. L. (2005). Stiffness matrix method with improved efficiency for elastic wave propagation in layered anisotropic media, *Journal of the Acoustical Society of America*, 118(6): 3400-3403.

Tan, E. L. (2006). Hybrid compliance-stiffness matrix method for stable analysis of elastic wave propagation in multilayered anisotropic media, *Journal of the Acoustical Society of America*, 119(1): 45-53.

Tan, E. L. (2007). Matrix Algorithms for modeling acoustic waves in piezoelectric multilayers, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 54: 2016-2023.

Tarn, J. Q. (2002a). A state space formalism for anisotropic elasticity. Part I: Rectilinear anisotropy, *International Journal of Solids and Structures*, 39: 5143-5155

Tarn, J. Q. (2002b). A state space formalism for piezothermoelasticity, *International Journal of Solids and Structures*, 39: 5173-5184

Thomson, T. (1950). Transmission of elastic waves through a stratified solid medium, *Journal of Applied Physics*, 21: 89-93.

Tian, J. Y., Yang, W. X. & Su, X. Y. (2006). Transient elastic waves in a transversely isotropic laminate impacted by axisymmetric load, *Journal of Sound and Vibration*, 289: 94-108.

Wang, L. & Rokhlin, S. I. (2001). Stable reformulation of transfer matrix method for wave propagation in layered anisotropic media, *Ultrasonics*, 39: 413-424.

Wang, L. & Rokhlin, S. I. (2002a). An efficient stable recursive algorithm for elastic wave propagation in layered anisotropic media, in Thompson, D. O. & Chimenti, D. E. (ed.), *Review of Quantitative Nondestructive Evaluation* 21, pp. 115-122.

Wang, L. & Rokhlin, S. I. (2002b). Recursive asymptotic stiffness matrix method for analysis of surface acoustic wave devices on layered piezoelectric media, *Applied Physics Letters*, 81: 4049-4051.

Wang, L. & Rokhlin, S. I. (2004a). A compliance/stiffness matrix formulation of general Green's function and effective permittivity for piezoelectric multilayers, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 51: 453-463.

Wang, L. & Rokhlin, S. I. (2004b). A simple method to compute ultrasonic wave propagation in layered anisotropic media, in Thompson, D. O. & Chimenti, D. E. (ed.), *Review of Quantitative Nondestructive Evaluation* 23, pp. 59-66.

Wang, L. & Rokhlin, S. I. (2004c). Modeling of wave propagation in layered piezoelectric media by a recursive asymptotic method, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 51: 1060-1071.

Zhang, V. Y., Lefebvre, J. E., Bruneel, C. & Gryba, T. (2001). A unified formalism using effective surface permittivity to study acoustic waves in various anisotropic and piezoelectric multilayers, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 48: 1449-1461.

3

Rectifying Acoustic Waves

Yukihiro Tanaka, Takahiro Murai, and Norihiko Nishiguchi
Department of Applied Physics, Hokkaido University
Sapporo 060-8628, Hokkaido,
Japan

1. Introduction

Diodes play an essential and important role in controlling current flow in the electric circuits. The rectification mechanisms utilize thermo electrons in vacuum tubes and the difference in energy band structures between p – n junctions of semiconductors. Ever since the advent of nanotechnology, efforts were made to realize a nano diode, and rectification of current is achieved by using geometric effects on electron scattering by asymmetric scatterers.(Song et al., 1998; Fleischmann & Geisel, 2002; Linke et al., 1998). In contrast, the rectification mechanism of acoustic waves is not established, and then rectification mechanisms of acoustic waves attract much attention from both theoretical (Liang et al., 2009) and experimental viewpoints (Chang et al., 2006).

Recent theoretical work proposed a rectification mechanism of acoustic waves utilizing mode conversion of acoustic waves owing to elastic anharmonicity in the constituent materials (Liang et al., 2009), and then the efficiency of rectification of the proposed model depends on the amplitude of acoustic waves. In addition, there is difficulty in finding suitable materials for fabricating the system.

We have proposed a rectification mechanism for acoustic waves (Krishnan et al., 2007;Shirota et al., 2007) based on geometric effects on scattering of acoustic waves by asymmetric scatterers, and have confirmed, with numerical simulations, the rectification effects of acoustic waves for bulk acoustic waves (Krishnan et al., 2007; Shirota et al., 2007), and Rayleigh waves in the surface (Tanaka et al., n.d.). In this chapter, we review the rectification mechanism, and illustrate the capability of rectification for bulk acoustic waves, and Rayleigh waves, respectively. Surface acoustic waves (SAWs) are exploited in a variety of devices, and then the proposed rectification mechanism is expected to provide novel functions to the acoustic wave devices.

The chapter is organized as follows; we introduce a model of rectifier for acoustic waves, and provide the methodology in §2. The numerical results for bulk transverse and longitudinal waves and Rayleigh waves in the surface are given in §3. A summary and future prospects are given in §4.

2. Model and methodology

The acoustic wave rectifier to be discussed consists of an elastically isotropic material containing a one-dimensional array of isosceles-triangular holes with summit angle α in the y direction, whose axis is in the z direction as shown in Fig. 1(a). The distance between the

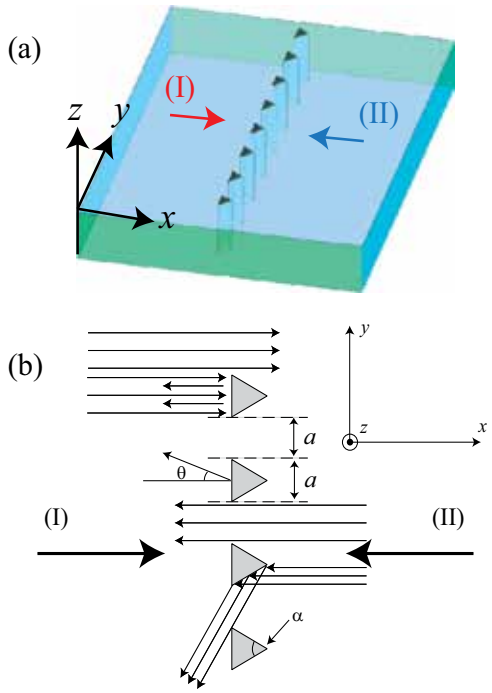


Fig. 1. (a) Schematic view of our model that shows the rectification effect for acoustic waves. (b) z–polarized transverse acoustic wave reflection from isosceles-triangular holes drilled in an elastically isotropic material. The triangles are separated by the same distance as the base, and are aligned in the y direction. The thick arrows indicate the directions (I) and (II) of incident waves. The thin arrows show reflection and transmission of waves in the very short wavelength limit, where the diffraction is ignored.

neighboring triangles is the same as the base length a of the triangles (thus the periodicity in the y direction is $D = 2a$). The holes are left empty in order to get strong reflection of acoustic waves.

Rectification for acoustic waves with very short wavelength in this system is apparent. Considering case (I) that acoustic rays propagate from the left region toward a matrix with equilateral-triangular holes ($\alpha = \pi/3$) of periodically aligned in the y direction as shown in Fig. 1(b), one half of them are scattered backward, and the rest passes between the scatterers. The resultant transmission rate becomes 0.5. On the other hand, the transmission rate becomes 1 for case (II) where the rays impinge on the summits of the triangular voids from the right region, since the rays reflected from the surface are transmitted to the left region through the passes between the holes. Thus the acoustic waves are rectified in the very short wavelength limit.

The prediction by the ray-acoustics cannot be immediately applied to the wave propagation at finite wavelength because of decay in the geometric effects and of interference effects. The waves peculiar to the geometry of scatterers decay near the scatterers and only the azimuthally symmetric waves propagate in the asymptotic field as a cylindrical wave in a two-dimensional (2D) system or a spherical wave in a 3D one. Thus the rectification seems impossible for finite wavelength, however, the present system works as a rectifier for acoustic waves as shown below.

Here we consider acoustic waves propagating through an array of triangular holes drilled in an isotropic material. In this case the equations of motion governing the displacement vectors $\mathbf{u}(\mathbf{r}, t)$ are given by

$$\rho(\mathbf{x})\ddot{u}_i(\mathbf{r},t) = \partial_j \sigma_{ij}(\mathbf{r},t) \quad (i=1,2,3) \tag{1}$$

$$\sigma_{ij}(\mathbf{r},t) = c_{ijmn}(\mathbf{x})\partial_n u_m(\mathbf{r},t) \tag{2}$$

where $\mathbf{r} = (\mathbf{x}, z) = (x, y, z)$ and the summation convention over repeated indices is assumed in Eqs. (1) and (2). $\rho(\mathbf{x})$ and $c_{ijmn}(\mathbf{x})$ are the position-dependent mass density and elastic stiffness tensor of the system, and $\sigma_{ij}(\mathbf{r}, t)$ is the stress tensor. Furthermore, we need to impose proper boundary conditions for SAWs. SAWs should satisfy the stress-free boundary condition at the surface $z = 0$, or

$$\sigma_{i3}\big|_{z=0} = c_{i3mn}\partial_n u_m\big|_{z=0} = 0 \quad (i=1,2,3). \tag{3}$$

Solving the equations with finite-difference time-domain (FDTD) method numerically, we can obtain the time evolutions of displacement vectors $\mathbf{u}(\mathbf{r}, t)$ and stress tensors $\sigma_{ij}(\mathbf{r}, t)$ at each point in the system. To calculate the transmission rate through the periodic array of triangle holes, we define the acoustic Poynting vector $J_i(\mathbf{r}, t) = -\dot{u}_j(\mathbf{r}, t)\sigma_{ji}(\mathbf{r}, t)$ from the continuity of energy flow. In terms of the Fourier components of the displacement $\hat{\mathbf{u}}(\mathbf{r},\omega)$ and the stress tensor $\hat{\sigma}_{ij}(\mathbf{r},\omega)$, the energy flow at frequency ω in the x direction at the position x is expressed by

$$\hat{J}_x(x,\omega) = -4\pi \int \mathrm{Im}\left[\omega \hat{u}_j(\mathbf{r},\omega)\hat{\sigma}^*_{jx}(\mathbf{r},\omega)\right] dydz. \tag{4}$$

Hence we can determine the transmission rate $T(\omega)$ by the ratio of the element of the acoustic Poynting vector in the x direction $\hat{J}_x(x_D,\omega)$ to that in the absence of scatterers $\hat{J}^0_x(x_D,\omega)$, which is given by

$$T(\omega) = \frac{\hat{J}_x(x_D,\omega)}{\hat{J}^0_x(x_D,\omega)}, \tag{5}$$

where x_D is the detecting position which is on the right side of the scatterers for case (I) and on the left side of the scatterers for case (II). We introduce an efficiency $\eta(\omega)$ to quantify rectification by

$$\eta(\omega) = \left|\frac{T_{\mathrm{I}}(\omega) - T_{\mathrm{II}}(\omega)}{T_{\mathrm{I}}(\omega) + T_{\mathrm{II}}(\omega)}\right| \tag{6}$$

where T_{I} and T_{II} are the transmission rates for cases (I) and (II), respectively.

3. Numerical results

3.1 Bulk acoustic waves

In the section we illustrate the rectifying effects of bulk acoustic waves propagating in the x-direction. Because of the homogeneity in the z direction, the governing equations (1) and (2) are decoupled into two independent sets; one is expressed by

$$\rho(\mathbf{x})\frac{\partial^2 u_z(\mathbf{x},t)}{\partial t^2} = \frac{\partial \sigma_{zx}(\mathbf{x},t)}{\partial x} + \frac{\partial \sigma_{zy}(\mathbf{x},t)}{\partial y},$$

$$\sigma_{zx}(\mathbf{x},t) = C_{44}(\mathbf{x})\frac{\partial u_z(\mathbf{x},t)}{\partial x},$$

$$\sigma_{zy}(\mathbf{x},t) = C_{44}(\mathbf{x})\frac{\partial u_z(\mathbf{x},t)}{\partial y},$$

where the acoustic wave is z-polarized transverse wave, referred to as *single mode*. Another is the acoustic waves termed *mixed modes* with polarization in the x – y plane, which obey

$$\rho(\mathbf{x})\frac{\partial^2 u_x(\mathbf{x},t)}{\partial t^2} = \frac{\partial \sigma_{xx}(\mathbf{x},t)}{\partial x} + \frac{\partial \sigma_{xy}(\mathbf{x},t)}{\partial y},$$

$$\rho(\mathbf{x})\frac{\partial^2 u_y(\mathbf{x},t)}{\partial t^2} = \frac{\partial \sigma_{xy}(\mathbf{x},t)}{\partial x} + \frac{\partial \sigma_{yy}(\mathbf{x},t)}{\partial y},$$

$$\sigma_{xx}(\mathbf{x},t) = C_{11}(\mathbf{x})\frac{\partial u_x(\mathbf{x},t)}{\partial x} + C_{12}(\mathbf{x})\frac{\partial u_y(\mathbf{x},t)}{\partial y},$$

$$\sigma_{yy}(\mathbf{x},t) = C_{12}(\mathbf{x})\frac{\partial u_x(\mathbf{x},t)}{\partial x} + C_{11}(\mathbf{x})\frac{\partial u_y(\mathbf{x},t)}{\partial y},$$

$$\sigma_{xy}(\mathbf{x},t) = C_{44}(\mathbf{x})\left(\frac{\partial u_x(\mathbf{x},t)}{\partial y} + \frac{\partial u_y(\mathbf{x},t)}{\partial x}\right),$$

The mixed modes consist of longitudinal and transverse waves due to the scattering by the triangular voids since the mode conversion between the longitudinal and transverse waves takes place for scattering.

3.1.1 Single mode

Figure 2 shows the transmission rates versus frequency in the case of equilateral-triangular holes ($\alpha = \pi/3$) and isosceles-triangular holes ($\alpha = 2\pi/9$) for two opposite incident directions (I) and (II). For both types of triangular holes, there is not noticeable difference in the transmission rates between the two incident directions at low frequency; $\omega a/v_t < \pi$. On the other hand, we find remarkable dependence in the transmission rates on the incident directions at $\omega a/v_t > \pi$. Above the threshold frequency $\omega_{th} a/v_t = \pi$, the transmission rate for (I) is approximately $T = 0.5$ that is the same as the magnitude predicted from the ray-acoustics, showing small dips in magnitude at the multiples of the threshold frequency. The transmission rate for (II) is larger than that for (I), and also shows periodic dips in magnitude with the same period as (I). The obvious difference in the transmission rates above the threshold frequency between (I) and (II) indicates that the rectification occurs at the wavelength comparable to the dimension of scatterers, i. e. $a/\lambda > 1/2$ due to the linear dispersion relation $\omega = kv_t = 2\pi v_t/\lambda$. Although the periodic dips, which appear when $\omega a/v_t = n\pi$ (n = 1, 2, 3, . . .), are common to both the equilateral- and isosceles-triangular holes, the latter system has advantageous properties for rectification of acoustic waves; the transmission rates for (II) of $\alpha = 2\pi/9$ are larger than those for $\alpha = \pi/3$. This indicates that the rectification is enhanced with decreasing α.

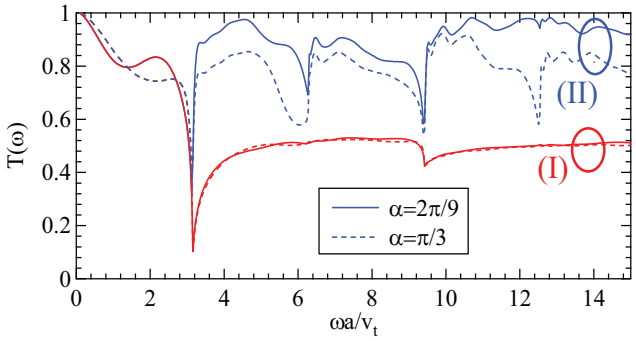


Fig. 2. Transmission rate versus frequency. The dashed and solid lines indicate the transmission rate for $\alpha = \pi/3$ and $2\pi/9$, respectively. Each case of (I) and (II) is bundled with an ellipse.

Within the ray acoustics approximation, the transmission rate of (II) is expected to be 1 for $\alpha < \pi/3$ and 0.5 for $\alpha > \pi/2$, and varies as $T = (1/2) + \cos\alpha$ for $\pi/3 < \alpha < \pi/2$. On the other hand, the transmission rate of (I) becomes 0.5, independent of α. For finite wavelength, the transmission rate changes with α as shown in Fig. 2, showing a larger transmission rate at $\alpha = 2\pi/9$ than that at $\alpha = \pi/3$. From the results, we expect that the rectification effects decay with increasing α. To investigate the angle dependence, we examine the change in the transmission rates for variation of α. Generating a wave packet having a Gaussian spectral distribution of central frequency $\omega_C = (5\pi/2)(v_t/a)$ with $\Delta\omega = (\pi/2)(v_t/a)$, we evaluate the transmission rate for the wave packet, defining

$$T = \frac{\int\int_{\omega_C-\Delta\omega}^{\omega_C+\Delta\omega} \hat{J}_x(x_D,\omega)\,d\omega dy dz}{\int\int_{\omega_C-\Delta\omega}^{\omega_C+\Delta\omega} \hat{J}_x^0(x_D,\omega)\,d\omega dy dz}. \tag{7}$$

Figure 3 plots the transmission rates defined by Eq. (7) versus α. The difference in the transmission rates decreases with increasing α. However, the rectification effects survive for $\alpha > \pi/2$. We also find that the transmission for (I) is slightly larger than 0.5. We can regard these deviations from the predictions based on the ray acoustics as diffraction effects.

The threshold frequency for the rectification and the periodic change in the transmission rate originate from the interference effects. Because of the periodic structure in the y-direction, the wavenumber in the y direction is discretized in unit of $\frac{n\pi}{a}$, so that the dispersion relation of the acoustic waves becomes subband structure given by

$$\omega = v_t k = v_t\sqrt{k_x^2 + \left(\frac{n\pi}{a}\right)^2}, \tag{8}$$

where n is an integer. Figure 4 shows the dispersion relation of single mode for the summit angle $\alpha = \pi/3$ together with the corresponding transmission rate. When the incident waves with $k_y = 0$ are elastically scattered, the waves are transited to the waves with finite k_y. However, below the threshold frequency $\omega_{th}a/v_t$, there is no waves with finite k_y, so that the incident waves in the x-direction are scattered only forward or backward, even if the waves are scattered from the legs of triangles, resulting in the transmission rates independent of the incident-wave directions.

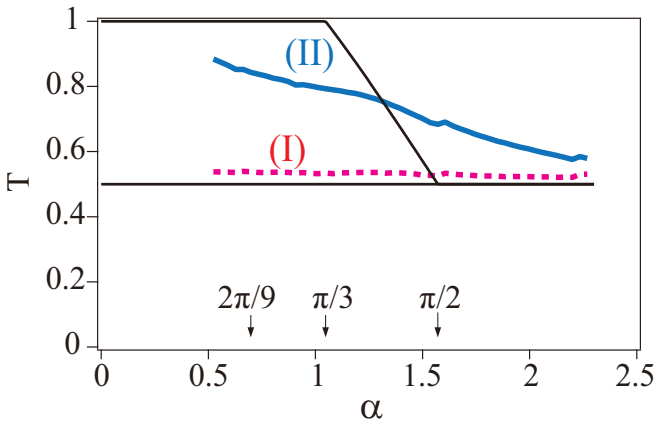


Fig. 3. Transmission rate defined by Eq. (7) versus the summit angle α. The thick dashed and solid lines indicate the transmission rate for cases (I) and (II), respectively. The thin solid lines indicate the transmission rates for (I) and (II) based on the ray acoustics.

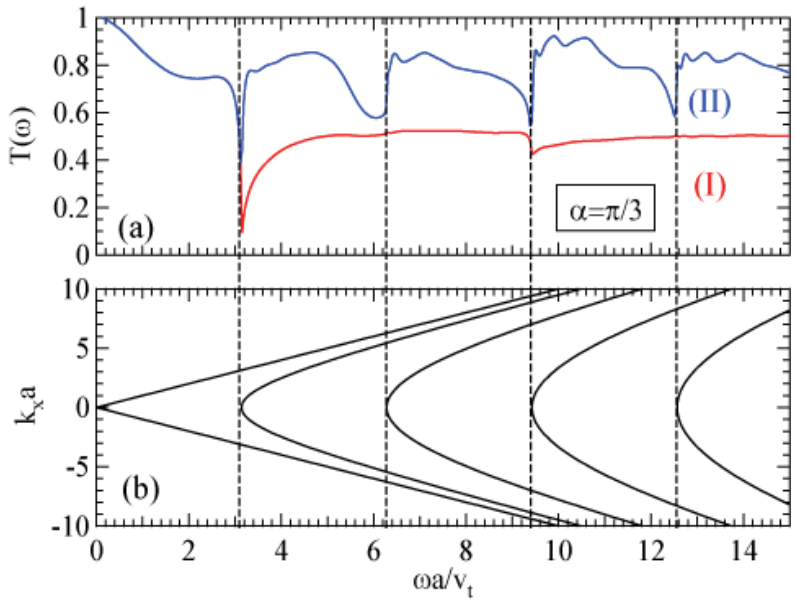


Fig. 4. (a) Transmission rate versus frequency for single modes through single-array of triangular holes with $\alpha = \pi/3$. The labels (I) and (II), designated by red and blue solid line, respectively, indicate the incident direction of acoustic waves. The vertical dashed lines indicate the positions of $n\pi v_t/a$, where $n = 1, 2, \ldots$ where v_t are the velocity of transverse waves. (b) Dispersion relation of single modes within the empty-lattice approximation.

Redirection of the incident waves for scattering occurs only in the frequency region above the threshold frequency. Since each dispersion relation of the waves with finite k_y becomes minimum at $k_x = 0$, the density of states diverges, resulting in remarkable scattering into the waves with finite k_y and $k_x = 0$ when the frequency matches the subband bottoms. The

geometry of the scatterers enhances or suppresses the redirection depending on the incident directions of the wave. Hence, the rectification occurs only above the threshold frequency and the dips in the transmission rates take place.

3.1.2 Mixed mode

The transmission rates versus frequency for mixed modes are shown in Fig. 5, when the longitudinal waves are transmitted. We assumed the matrix made of tungsten, whose the mass density ρ and the elastic stiffness tensors C_{11}, C_{44} are 19.317g cm^{-3} and 5.326×10^{12}dyn cm^{-2}, 1.631 × 10^{12}dyn cm^{-2}, respectively. The velocities of bulk longitudinal and transverse waves are v_l = 5.25 × 10^{5}cm s^{-1} and v_t = 2.906 × 10^{5}cm s^{-1}. (Kittel, 2004)
The red and blue solid lines indicate two different incident directions (I) and (II), respectively. The two transmission rates agree for $\omega a/v_t < \pi$, and we can see the difference between the transmissions for $\omega a/v_t > \pi$, although it is not as large as that for the single modes, manifesting rectification of the mixed modes. Unlike the single modes, we can see two kinds of periodic changes in transmission rates above $\omega a/v_t = \pi$, one is periodic modulation with period $\Delta\omega a/v_t = \pi$, indicated by the black dashed vertical lines, and another is periodic variations with period $\Delta\omega a/v_t = \pi \times v_l/v_t \sim 1.807\pi$, indicated by the green ones. In addition, some aperiodic dips in the transmission rate indicated by the arrows appear above $\omega a/v_t = \pi$. These dips shift when the shape of the triangular hole changes. Very interestingly there is no rectification in high frequency regions ($\omega a/v_t > 13$) because, for the waves impinging on the summit, the mode conversion from longitudinal waves to transverse ones is strongly caused and the scattered transverse waves return to the incident direction.

3.2 Surface acoustic waves

Figure 6(a) shows the frequency dependences of the transmission rates for SAWs with the incident-wave directions (I) and (II) which are denoted by red and blue solid lines, respectively. For numerical evaluation, the matrix is assumed to be polycrystalline silicon regarded as an isotropic medium, where the mass density ρ and the stiffness tensors C_{11}, C_{44} are 2.33g cm^{-3} and 1.884 × 10^{12}dyn cm^{-2}, 0.680 × 10^{12}dyn cm^{-2}, respectively. (Tamura, 1985) Then the velocities of bulk longitudinal and transverse waves are v_l = 8.99 × 10^{5}cm s^{-1} and v_t = 5.40 × 10^{5}cm s^{-1}, respectively. The equation for the velocity of a Rayleigh wave in an isotropic medium with a surface is given by

$$\xi^6 - 8\xi^4 + 8\xi^2\left(3 - 2\frac{v_t^2}{v_l^2}\right) - 16\left(1 - \frac{v_t^2}{v_l^2}\right) = 0, \tag{9}$$

where $\xi = v_R/v_t$ (v_R is the velocity of Rayleigh wave). (Graff, 1991) Solving Eq. (9), we obtain ξ = 0.914. A wave packet with z-polarized vector is used as an incident wave in order to excite SAWs in the system. Below the threshold frequency corresponding to the wavelength of SAWs equivalent to the periodicity of the array, both the transmission rates are coincident because the waves with long wavelength cannot recognize the geometrical difference. However, above the threshold frequency, the transmission rate shows obvious rectification of SAWs as well as periodic dips with respect to frequency, resulting from the strong interference effects of scattered SAWs. We also find the periodic structure of the transmission rate of case (II) is more pronounced than that of case (I) because the former makes the mode conversion more accessible than the latter due to the geometry of the scatterers.

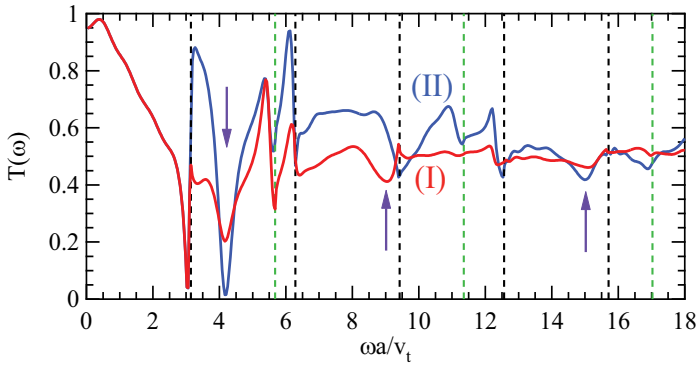


Fig. 5. Transmission rate versus frequency for mixed modes through single-array of triangular holes with $\alpha = \pi/3$. (I) and (II), designated by red and blue solid line, respectively, indicate the incident direction of acoustic waves. The vertical black dashed lines (as shown in Fig. 4) and green ones indicate the positions of $n\pi v_t/a$ and $n\pi v_l/a$, where v_l and v_t are the velocity of longitudinal and transverse waves, respectively, and n is a positive integer (n = 1, 2, 3, . . .). The arrows indicate the dips whose positions depend on the geometry of triangular holes such as the summit angle α and the length of base a.

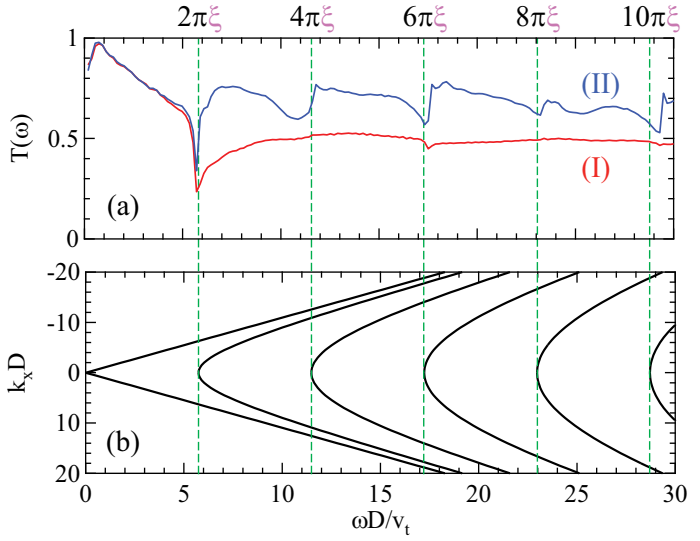


Fig. 6. (a) Transmission rate versus frequency for SAWs through single-array of triangular holes with $\alpha = \pi/3$. The labels (I) and (II), designated by red and blue solid line, respectively, indicate the incident directions of SAWs. The vertical dashed lines indicate the positions of $2n\xi\pi v_t/D$, where n = 1, 2, . . . where $\xi = v_R/v_t$ (v_R and v_t are the velocity of Rayleigh and transverse waves, respectively). (b) Dispersion relation of SAWs within the framework of empty-lattice approximation.

Figure 6(b) shows the dispersion relation of SAWs within the framework of empty-lattice approximation to reveal the origin of the periodic dips in Fig. 6(a). Within the empty-lattice approximation the subband structures due to the periodicity of the y direction appear in the dispersion relation, which are given by

$$\frac{\omega D}{v_t} = \xi\sqrt{\left(k_x D\right)^2 + \left(2\pi n\right)^2}\,, \left(n = 0, \pm 1, \pm 2, \cdots\right). \tag{10}$$

The dispersion relation can be obtained by replacing the wavevector k in the dispersion relation of Rayleigh waves $\omega = v_R k$ by $\sqrt{k_x^2 + \left(k_y + 2\pi n / D\right)^2}$ where $2\pi n/D$ is the reciprocal lattice vector in the y direction. The dips in Fig. 6(a) correspond to the band edges of the subband structure, manifesting that the periodic dips in the transmission are due to the Bragg reflection of SAWs in the y direction. It should be noted that the shift of the band edges for the SAWs is modified by a factor of ξ as much as that for bulk transverse waves.
Figure 7 shows the efficiency for the rectification of SAWs which is denoted by black solid line. The thin blue lines indicate the efficiency for bulk transverse waves as reference. The efficiency for SAWs is lower than that for bulk waves because of the mode conversion from SAWs to bulk waves due to the triangular scatterers.
Figure 8 (a) shows the transmission rate versus frequency for shear horizontal (SH) modes through the single-array of triangular holes. For excitation of SH waves in the system, we use a wave packet with y-polarization vector as an incident wave. The threshold frequency above which the rectifying effect occurs becomes exactly $2\pi v_t/D$ where v_t is the velocity of SH waves. Above the threshold frequency, the transmission rates exhibit dips periodically at multiples of the threshold frequency due to the same mechanism as the SAWs and bulk waves. However, the SH waves are inefficient compared to the SAWs as shown in Fig. 8(b). The inversions between the transmissions of cases (I) and (II) occur around $\omega D/v_t \approx 18$.

4. Summary and future prospects

We proposed an acoustic-wave rectifier and numerically demonstrated the rectification effects on bulk waves as well as SAWs above the threshold frequencies. The rectification mechanism is due to the geometric effects of the asymmetric scatterers on acoustic wave scattering, which is enhanced by interference among the scattered waves. The threshold frequency for the rectification results from the periodic arrangement of scatterers. Hence, it is possible to tune the rectifier by adjusting the position of the scatterers. The findings of this work can be applied not only to sound waves in solids or liquids but also to optical waves, leading to new devices in wave engineering.

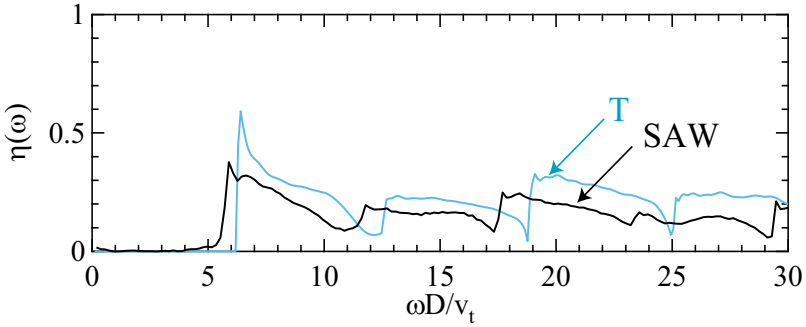


Fig. 7. Efficiency for the rectification of SAWs ($\alpha = \pi/3$). The solid black and thin blue lines indicate the efficiencies for SAWs and bulk single modes (T), respectively. The efficiency of the SAW rectifiers is slightly lower than that of the bulk single mode.

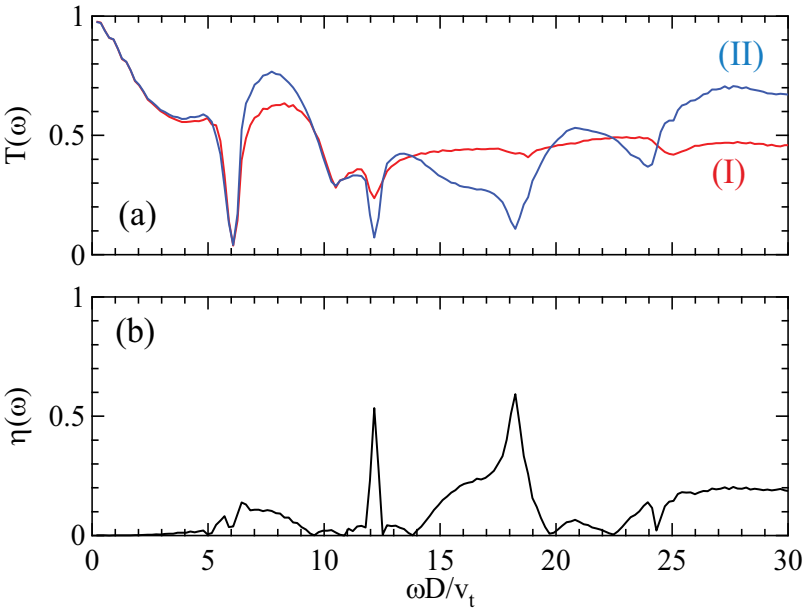


Fig. 8. (a) Transmission rate versus frequency for SH modes through single-array of triangular holes with $\alpha = \pi/3$. The labels, (I) and (II), designated by red and blue solid line, respectively, indicate the incident direction of acoustic waves. (b) Efficiency for the rectification of SH waves.

5. References

Chang, C.W., Okawa, D., Majumdar, A. & Zettl, A. (2006). Solid-state thermal rectifier, *Science* 314(5802): 1121–1124.

Fleischmann, R. & Geisel, T. (2002). Mesoscopic rectifiers based on ballistic transport, *Phys. Rev. Lett.* 89(1): 016804.

Graff, K. F. (1991). *Wave Motion in Elastic Solids*, Dover Publications.

Kittel, C. (2004). *Introduction to Solid State Physics, 8 edition*, Wiley.

Krishnan, R., Shirota, S., Tanaka, Y. & Nishiguchi, N. (2007). High-efficient acoustic wave rectifier, *Solid State Communications* 144(5-6): 194–197.

Liang, B., Yuan, B. & Cheng, J.-C. (2009). Acoustic diode: Rectification of acoustic energy flux in one-dimensional systems, *Phys. Rev. Lett.* 103(10): 104301.

Linke, H., Sheng, W., Löfgren, A., Xu, H.-G., Omling, P. & Lindelof, P. E. (1998). A quantum dot ratchet: Experiment and theory, *Europhys. Lett.* 44(3): 341.

Shirota, S., Krishnan, R., Tanaka, Y. & Nishiguchi, N. (2007). Rectifying acoustic waves, *Japanese Journal of Applied Physics* 46(42): L1025–L1027.

Song, A. M., Lorke, A., Kriele, A., Kotthaus, J. P., Wegscheider, W. & Bichler, M. (1998). Nonlinear electron transport in an asymmetric microjunction: A ballistic rectifier, *Phys. Rev. Lett.* 80(17): 3831–3834.

Tamura, S. (1985). Spontaneous decay rates of la phonons in quasi-isotropic solids, *Phys. Rev. B* 31(4): 2574–2577.

Tanaka, Y., Murai, T. & Nishiguchi, N. (n.d.). in preparation.

4

Dispersion Properties of Co-Existing Low Frequency Modes in Quantum Plasmas

S. A. Khan[1,2,3] and H. Saleem[1,2]
[1]*Department of Physics,*
COMSATS Institute of Information Technology (CIIT), Islamabad 44000
[2]*National Centre for Physics,*
Quaid-i-Azam University Campus, Islamabad 43520
[3]*Department of Physics,*
Government College Bagh 12500, Azad Jammu & Kashmir,
Pakistan

1. Introduction

The underlying physics of nonconventional quantum plasmas has been introduced long ago. Analytical investigations of collective interactions between an ensemble of degenerate electrons in a dense quantum plasma dates back to early fifties. The general kinetic equations for quantum plasmas were derived and the dispersion properties of plasma waves were studied (Klimotovich & Silin, 1952). It was thought that the quantum mechanical behaviour of electrons, in the presence of heavier species modifies the well known properties of plasma. The dynamics of quantum plasmas got particular attention in the framework of relationship between individual particle and collective behavior. Emphasizing the excitation spectrum of quantum plasmas, theoretical investigations describe the dispersion properties of electron plasma oscillations involving the electron tunneling (Bohm & Pines, 1953; Pines, 1961). A general theory of electromagnetic properties of electron gas in a quantizing magnetic field and many particle kinetic model of non-thermal plasmas was also developed treating the electrons quantum mechanically (Zyrianov et al., 1969; Bezzerides & DuBois, 1972). Since the pioneering work of these authors which laid foundations of quantum plasmas, many theoretical studies have been done in the subsequent years. The rapidly growing interest in quantum plasmas in the recent years has several different origins but is mainly motivated by its potential applications in modern science and technology (e.g. metallic and semiconductor micro and nanostructures, nanoscale plasmonic devices, nanotubes and nanoclusters, spintronics, nano-optics, etc.). Furthermore, quantum plasmas are ubiquitous in planetary interiors and in compact astrophysical objects (e.g., the interior of white dwarfs, neutron stars, magnetars, etc.) as well as in the next generation intense laser-solid density plasma interaction experiments. Such plasmas also provide promises of important futuristic developments in ultrashort pulsed lasers and ultrafast nonequilibrium phenomena (Bonitz, 1898; Lai, 2001; Shukla & Eliasson, 2009).

Contrary to classical plasmas, the number density of degenerate electrons, positrons/holes in quantum plasmas is extremely high and they obey Fermi-Dirac statistics whereas the temperature is very low. Plasma and quantum mechanical effects co-exist in such systems and many unusual effects like tunneling of electrons, quantum destabilization, pressure ionization, Bose-Einstein condensation and crystallization etc. may be equally important (Bonitz et al., 2009). Their properties in equilibrium and nonequilibrium are governed by many-body effects (collective and correlation effects) which require quantum statistical theories and versatile computational techniques. The average inter-particle distance $n^{-1/3}$ (where n is the particle density) is comparable with electron thermal de Broglie wavelength λ_{Be} (= $\hbar/mv_{te}$, where $\hbar$ is Planck's constant divided by 2π, m is the electronic mass and v_{te} is thermal speed of electron). The overlapping of wave functions associated with electrons or positrons take place which leads to novel quantum effects.

It was recognized long ago that the governing quantum-like equations describing collective behavior of dense plasmas can be transformed in the form of hydrodynamic (or fluid) equations which deals with macroscopic variables only (Madelung, 1926). Here, the main line of reasoning starts from Schrodinger description of electron. The N-body wave function of the system can be factored out in N one-body wave functions neglecting two-body and higher order correlations. This is justified by weak coupling of fermions at high densities. The coupling parameter of quantum plasmas decreases with increase in particle number density. For hydrodynamic representation, the electron wave function is written as $\psi = \sqrt{n}$ $\exp(iS/\hbar)$ where n is amplitude and S is phase of the wave function. Such a decomposition of ψ was first presented by Bohm and de Broglie in order to understand the dynamics of electron wave packet in terms of classical variables. It introduces the Bohm-de Broglie potential in equation of motion giving rise to dispersion-like terms. In the recent years, a vibrant interest is seen in investigating new aspects of quantum plasmas by developing non-relativistic quantum fluid equations for plasmas starting either from real space Schrodinger equation or phase space Wigner (quasi-) distribution function. (Haas et al., 2003, Manfredi & Haas, 2001; Manfredi, 2005). Such approaches take into account the quantum statistical pressure of fermions and quantum diffraction effects involving tunneling of degenerate electrons through Bohm-de Broglie potential. The hydrodynamic theory is also extended to spin plasmas starting from non-relativistic Pauli equation for spin-$\frac{1}{2}$ particles (Brodin & Marklund, 2007; Marklund & Brodin, 2007). Generally, the hydrodynamic approach is applicable to unmagnetized or magnetized plasmas over the distances larger than electron Fermi screening length λ_{Fe} (= v_{Fe}/ω_{pe}, where v_{Fe} is the electron Fermi velocity and ω_{pe} is the electron plasma frequency). It shows that the plasma effects at high densities are very short scaled.

The present chapter takes into account the dispersive properties of low frequency electrostatic and electromagnetic waves in dense electron-ion quantum plasma for the cases of dynamic as well as static ions. Electrons are fermions (spins=1/2) obeying Pauli's exclusion principle. Each quantum state is occupied only by single electron. When electrons are added, the Fermi energy of electrons ϵ_{Fe} increases even when interactions are neglected ($\epsilon_{Fe} \propto n^{2/3}$). This is because each electron sits on different step of the ladder according to Pauli's principle which in turn increases the statistical (Fermi) pressure of electrons. The de Broglie wavelength associated with ion as well as its Fermi energy is much smaller as compared to electron due to its large mass. Hence the ion dynamics is classical. Quantum

diffraction effects (quantum pressure) of electrons are significant only at very short length scales because the average interparticle distance is very small. This modifies the collective modes significantly and new features of purely quantum origin appear. The quantum ion-acoustic type waves in such system couples with shear Alfven waves. The wave dispersion due to gradient of Bohm-de Broglie potential is weaker in comparison with the electrons statistical/Fermi pressure. The statistical pressure is negligible only for wavelengths smaller than the electron Fermi length. For plasmas with density greater than the metallic densities, the statistical pressure plays a dominant role in dispersion.

2. Basic description

The coupling parameter for a traditional classical plasma is defined as

$$\Gamma^{C} = \frac{\left|\langle U \rangle\right|}{\langle K \rangle} \propto \frac{\sqrt[3]{n}}{T}, \tag{1}$$

where $\langle U \rangle = \frac{1}{2}\sum_{i \neq j} \frac{e_i e_j}{\bar{r}}$ is the two-particle Coulomb interaction (potential) energy, $\langle K \rangle = \frac{3}{2} k_B T$ is the average kinetic energy, k_B is the Boltzmann constant and T is the system's temperature. The average interparticle distance $\bar{r}$ is given by

$$\bar{r} = \langle r_i - r_j \rangle \propto \frac{1}{\sqrt[3]{n}}. \tag{2}$$

The parameter Γ^{C} may also be written in the form $\left(\frac{1}{n\lambda_D^3}\right)^{2/3}$ where $\lambda_D = \left(\frac{k_B T}{4\pi n e^2}\right)^{1/2}$ is the Debye screening length. Such a plasma obeys Boltzmann distribution function in which the ordering $\Gamma^{C} << 1$ corresponds to collisionless and $\Gamma^{C} \simeq 1$ to collisional regime. So, a classical plasma can be said collisionless (ideal) when long-range self-consistent interactions (described by the Poisson equation) dominate over short-range two-particle interactions (collisions).

When the density is very high , $\bar{r}$ become comparable to thermal de Broglie wavelength of charged particles defined by

$$\lambda_B = \frac{h}{\sqrt{2\pi m k_B T}}, \tag{3}$$

where h is the Planck's constant. Here, the degeneracy effects cannot be neglected i.e., $1 \lesssim n\lambda_B^3$ and the quantum mechanical effects along with collective (plasma) effects become important at the same time. Such plasmas are also referred to as quantum plasmas. Some common examples are electron gas in an ordinary metal, high-density degenerate plasmas in white dwarfs and neutron stars, and so on. From quantum mechanical point of view, the state of a quantum particle is characterized by the wave function associated with the particle instead of its trajectory in phase space. The Heisenberg uncertainty principle leads to fundamental modifications of classical statistical mechanics in this case. The de Broglie wavelength has no role in classical plasmas because it is too small compared to the average

interparticle distances. There is no overlapping of the wave functions and consequently no quantum effects. So the plasma particles are considered to be point like and treated classically.

However, in quantum plasmas, the overlapping of wave functions takes place which introduces novel quantum effects. It is clear from (3) that the de Broglie wavelength depends upon mass of the particle and its thermal energy. That is why, the quantum effects associated with electrons are more important than the ions due to smaller mass of electron which qualifies electron as a true quantum particle. The behavior of such many-particle system is now essentially determined by statistical laws. The plasma particles with symmetric wave functions are termed as Bose particles and those with antisymmetric wave function are called Fermi particles. We can subdivide plasmas into (*i*) quantum (degenerate) plasmas if $1 < n\lambda_B^3$ and (*ii*) classical (nondegenerate) plasmas if $n\lambda_B^3 < 1$. The border between the degenerate and the non-degenerate plasmas is roughly given by

$$n\lambda_B^3 = n\left(\frac{h}{\sqrt{2\pi m k_B T}}\right)^3 = 1. \tag{4}$$

For quantum plasmas, the Boltzmann distribution function is strongly modified to Fermi-Dirac or Bose-Einstein distribution functions in a well known manner, i.e.,

$$f(\epsilon) = \frac{1}{\left[e^{\beta(\epsilon-\mu)} \pm 1\right]}, \tag{5}$$

where $\beta = 1/k_B T$; ϵ is the particle energy and μ is the chemical potential. The '+' sign corresponds to Fermi-Dirac distribution function (for fermions with spin 1/2, 3/2, 5/2, . . .) and '_' sign to Bose-Einstein distribution function (for bosons with spin 0, 1, 2, 3, . . .). The different signs in the denominators of (5) are of particular importance at low temperatures. For fermions, this leads to the existence of Fermi energy (Pauli principle), and for bosons, to the possibility of macroscopic occupation of the same quantum state which is the well known phenomenon of Bose-Einstein condensation.

Let us consider a degenerate Fermi gas of electrons at absolute zero temperature. The electrons will be distributed among the various quantum states so that the total energy of the gas has its least possible value. Since each state can be occupied by not more than one electron, the electrons occupy all the available quantum states with energies from zero (least value) to some largest possible value which depends upon the number of electrons present in the gas. The corresponding momenta also starts from zero to some limiting value (Landau & Lifshitz, 1980). This limiting momentum is called the Fermi momentum p_F given by

$$p_F = \hbar\left(3\pi^2 n\right)^{1/3}. \tag{6}$$

Similarly, the limiting energy is called the Fermi energy ϵ_F which is

$$\epsilon_F = \frac{p_F^2}{2m} = \frac{\hbar^2}{2m}\left(3\pi^2 n\right)^{2/3}. \tag{7}$$

The Fermi-Dirac distribution function becomes a unit step function in the limit $T \to 0$. It is zero for $\mu < \epsilon$ and unity for $\epsilon < \mu$. Thus the chemical potential of the Fermi gas at $T = 0$ is same as the limiting energy of the fermions ($\mu = \epsilon_F$). The statistical distribution of plasma particles changes from Maxwell-Boltzmann $\propto \exp(-\epsilon/k_BT)$ to Fermi-Dirac statistics $\propto \exp[(\beta(\epsilon - \epsilon_F) + 1)]^{-1}$ whenever T approaches the so-called Fermi temperature T_F, given by

$$k_B T_F \equiv \epsilon_F = \frac{\hbar^2}{2m}\left(3\pi^2 n\right)^{2/3}. \tag{8}$$

Then the ratio $\chi = T_F / T$ can be related to the degeneracy parameter $n\lambda_B^3$ as,

$$\chi = T_F / T = \frac{1}{2}\left(3\pi^2\right)^{2/3}\left(n\lambda_B^3\right)^{2/3} \tag{9}$$

It means that the quantum effects are important when $1 \lesssim T_F / T$. In dense plasmas, the plasma frequency $\omega_p = (4\pi ne^2/m)^{1/2}$ becomes sufficiently high due to very large equilibrium particle number density. Consequently, the typical time scale for collective phenomena $(\omega_p)^{-1}$ becomes very short. The thermal speed $v_T = (k_BT/m)^{1/2}$ is sufficiently smaller than the Fermi speed given by

$$v_F = \left(\frac{2\epsilon_F}{m}\right)^{1/2} = \frac{\hbar^2}{m}\left(3\pi^2 n\right)^{1/3}. \tag{10}$$

With the help of plasma frequency and Fermi speed, we can define a length scale for electrostatic screening in quantum plasma i.e., the Fermi screening length λ_F (= v_{Fe}/ω_{pe}) which is also known as the quantum-mechanical analogue of the electron Debye length λ_{De}. The useful choice for equation of state for such dense ultracold plasmas is of the form (Manfredi, 2005)

$$P = P_0\left(\frac{n}{n_0}\right)^{\gamma}, \tag{11}$$

where the exponent $\gamma = (d + 2)/d$ with $d = 1, 2, 3$ denoting the dimensionality of the system, and P_0 is the equilibrium pressure. In three dimensions, $\gamma = 5/3$ and $P_0 = (2/5)\, n_0\epsilon_F$ which leads to

$$P = \left(\hbar^2 / 5m\right)(3\pi^2)^{2/3} n^{5/3}. \tag{12}$$

For one dimensional case, $\gamma = 3$ and $P = \left(mv_F^2 / 3n_0^2\right)n_e^3$. It shows that the electrons obeying Fermi-Dirac statistics introduce a new pressure at zero temperature called the Fermi pressure, which is significant in dense low temperature plasmas. The Fermi pressure increases with increase in number density and is different from thermal pressure.

Like classical plasmas, a coupling parameter can be defined in a quantum plasma. For strongly degenerate plasmas, the interaction energy may still be given by $\langle U \rangle$, but the kinetic energy is now replaced by the Fermi energy. This leads to the quantum coupling parameter

$$\Gamma^Q = \frac{\langle U \rangle}{\epsilon_F} \propto \frac{\sqrt[3]{n}}{T_F}, \tag{13}$$

which shows that $\Gamma^Q \propto n^{-1/3}$. So the peculiar property of quantum plasma is that it increasingly approaches the more collective (ideal) behavior as its density increases. Quantum plasma is assumed to be collisionless when $\Gamma^Q << 1$ because the two body correlation can be neglected in this case. This condition is fulfilled in high density plasmas since $\epsilon_F = \epsilon_F(n)$. In the opposite limit of high temperature and low density, we have $1 >> n\lambda_B^3$ and the system behaves as a classical ideal gas of free charge carries. Another useful form of Γ^Q is as follows

$$\Gamma^Q \approx \left(\frac{1}{n\lambda_F^3}\right)^{2/3} = \left(\frac{\hbar\omega_p}{\epsilon_F}\right)^{2/3}, \tag{14}$$

which shows the resemblance with classical coupling parameter which may be recovered in the limit $\lambda_F \to \lambda_D$.

3. Governing equations

Suppose that the N-particle wave function of the system can be factorized into N one-particle wave functions as $\Psi(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N) = \psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2)\cdots\psi_N(\mathbf{x}_N)$. Then, the system is described by statistical mixture of N states ψ_α, $\alpha = 1, 2, ..., N$ where the index α sums over all particles independent of species. We then take each ψ_α to satisfy single particle Schrodinger equation where the potentials ($\mathbf{A}$, ϕ) is due to the collective charge and current densities. For each ψ_α, we have corresponding probability p_α such that $\sum_{\alpha=1}^{N} p_\alpha = 1$ and all types of entanglements are neglected in the weak coupling limit. To derive the quantum fluid description, we define $\psi_\alpha = \sqrt{n_\alpha}\exp(iS_\alpha/\hbar)$ where n_α and S_α are real and the velocity of αth particle is $v_\alpha = \nabla S_\alpha / m_\alpha - (q_\alpha / m_\alpha c)\mathbf{A}$. Next, defining the global density $n = \sum_{\alpha=1}^{N} p_\alpha n_\alpha$, the global velocity $v = \sum_{\alpha=1}^{N} p_\alpha \frac{n_\alpha}{n} v_\alpha \equiv \langle v_\alpha \rangle$, and separating the real and imaginary parts in Schrodinger equation, the resulting continuity and momentum balance equation take the form

$$\frac{\partial n}{\partial t} + \nabla\cdot(n\mathbf{v}) = 0, \tag{15}$$

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v} = \frac{q}{m}\left(\mathbf{E} + \frac{1}{c}\mathbf{v}\times\mathbf{B}\right) + \frac{\hbar^2}{2m^2}\nabla\left(\frac{\nabla^2\sqrt{n}}{\sqrt{n}}\right) - \frac{1}{mn}\nabla P. \tag{16}$$

The last term in (16) is the statistical pressure term. For high temperature plasmas, it is simply thermal pressure. However, in low temperature and high density regime, the Fermi pressure is significant which corresponds to fermionic nature of electrons and P is given by equation (12). In the model (15)-(16), it is assumed that pressure $P = P(n)$ which leads to the appropriate

equation of state to obtain the closed system of equations. For typical length scales larger than λ_{Fe}, we have approximated the $\hbar$-term as $\sum_{\alpha=1}^{N} p_\alpha \left(\nabla^2\sqrt{n_\alpha}/\sqrt{n_\alpha}\right) \approx \left(\nabla^2\sqrt{n}/\sqrt{n}\right)$. This term is gradient of the so called Bohm-de Broglie potential. The equations (15)-(16) are commonly known as the quantum fluid equations which are coupled to the Poisson's equation and Ampere's law to study the self -consistent dynamics of quantum plasmas (Manfredi 2005; Brodin & Marklund, 2007). This model has obtained considerable attention of researchers in the recent years to study the behaviour of collisionless plasmas when quantum effects are not negligible. Starting from simple cases of electrostatic linear and nonlinear modes in two component and multicomponent plasmas, e.g., linear and nonlinear quantum ion waves in electron-ion (Haas et al., 2003, Khan et al., 2009), electron-positron-ion (Khan et al., 2008, 2009) and dust contaminated quantum plasmas (Khan et al., 2009), the studies are extended to electromagnetic waves and instabilities (Ali, 2008). Some particular developments have also appeared in spin-1/2 plasmas (Marklund & Brodin, 2007; Brodin & Marklund, 2007; Shukla, 2009), quantum electrodynamic effects (Lundin et al., 2007) and quantum plasmadynamics (Melrose, 2006). It is to mention here that the inclusion of simple collisional terms in such model is much harder and the exclusion of interaction terms is justified by small value of Γ^Q.

4. Fermionic pressure and quantum pressure

For dense electron gas in metals with equilibrium density $n_{e0} \simeq 10^{23}\text{cm}^{-3}$, the typical value of Fermi screening length is of the order of Angstrom while the plasma oscillation time period (ω_{pe}^{-1}) is of the order of femtosecond. The electron-electron collisions can been ignored for such short time scales. The Fermi temperature of electrons is very large in such situations i.e., $T_{Fe} \simeq 9 \times 10^4\text{K}$ which shows that the electrons are degenerate almost always (Manfredi & Haas, 2001). The Fermi energy, which increases with the plasma density, becomes the kinetic energy scale. The quantum criterion of ideality has the form

$$\Gamma^Q \approx \frac{e^2\sqrt[3]{n}}{\epsilon_{Fe}} << 1. \tag{17}$$

The parameter Γ^Q decreases with increasing electron density, therefore, a degenerate electron plasma becomes even more ideal with compression. So, even in the fluid approximation, it is reasonable to compare the statistical pressure term arising due to the fermionic character of electrons and the quantum Bohm-de Broglie potential term in the ultracold plasma.

Let us consider two-component dense homogenous plasma consisting of electrons and ions. The plasma is embedded in a very strong uniform magnetic field $B_0\hat{\mathbf{z}}$; where B_0 is the strength of magnetic field and $\hat{\mathbf{z}}$ is the unit vector in z-axis direction. However, plasma anisotropies, collisions and the spin effects are not considered in the model for simplicity. The low frequency (in comparison with the ion cyclotron frequency $\Omega_{ci} = eB_0/m_ic$, where e, m_i and c are the magnitude of electron charge, ion mass and speed of light in vacuum, respectively) electric and magnetic field perturbations are defined as $\mathbf{E} = -\nabla\phi - c^{-1}\left(\partial A_z/\partial t\right)\hat{\mathbf{z}}$ and $\mathbf{B}_\perp = \nabla_\perp A_z \times \hat{\mathbf{z}}$, respectively, where ϕ is the electrostatics wave

potential and A_z is the component of vector potential along z-axis. For very low temperature plasma by assuming that the ions behave classically in the limit $T_{Fi} << T_{Fe}$ (where T_{Fi} is the ion Fermi temperature), the pressure effects of quantum electrons are relevant only. In this situation the Fermi pressure which is contribution of the electrons obeying the Fermi-Dirac equilibrium is of most significance. In linearized form, the gradient of Fermi pressure for spin- $\frac{1}{2}$ electrons from (12) leads to

$$\nabla P_{Fe1} \simeq \frac{\hbar^2}{3m_e}\left(3\pi^2 n_{e0}\right)^{2/3}\nabla n_{e1}, \tag{18}$$

where the perturbation is assumed to be proportional to $exp[i\ (\mathbf{k}\cdot\mathbf{r} - \omega t)]$. The index 0 and 1 is used to denote the equilibrium, and perturbation, respectively. The $\hbar$-term in expression (16) i.e., the gradient of the Bohm-de Broglie potential in the linear limit may be written as

$$\nabla P_{q1} \simeq \frac{\hbar^2 k^2}{4m_e}\nabla n_{e1}, \tag{19}$$

where P_q has the dimensions of pressure. Notice that $(3\pi^2)^{2/3} \simeq 9.6$ and $n_{e0}^{2/3} = \frac{1}{\bar{r}^2}$ where $\bar{r}$ is the average interparticle distance. If $k \sim 10^6 \text{cm}^{-1}$ is assumed, then near metallic electron densities i.e., $n_{e0} \sim 10^{23}\text{cm}^{-3}$, we have, $\bar{r} \sim 10^{-8}$cm; which shows that

$$k^2 << \frac{1}{\bar{r}^2}. \tag{20}$$

The inequality (20) shows that the variation of quantities should be on length scales that are larger than λ_{Fe} and the fluid model is useful on such scales (Khan & Saleem, 2009).

5. Dynamics of ions and electrons

Starting from the fluid equations (15)-(16), it is assumed that the quantum effects of ions are neglected due to their larger mass in the limit $T_{Fi} << T_{Fe}$. The equation of motion for jth species may be written as,

$$m_j n_j\left(\frac{\partial}{\partial t} + \mathbf{v}_j.\nabla\right)\mathbf{v}_j = q_j n_j\left(\mathbf{E} + \frac{1}{c}\mathbf{v}_j \times \mathbf{B}\right) - \nabla P_j + \frac{\hbar^2 n_j}{2m_j}\nabla\left(\frac{\nabla^2\sqrt{n_j}}{\sqrt{n_j}}\right), \tag{21}$$

where $j = e, i$ for electron, and ion. Furthermore, the ions are assumed to be singly charged i.e., $q_j = e\ (-e)$ for electrons (ions) and the steady state is defined as $n_{e0} = n_{i0}$. The linearized ion velocity components in the perpendicular and parallel directions are

$$\mathbf{v}_{i\perp 1} \simeq \frac{c}{B_0}\left(\mathbf{z}\times\nabla_\perp\phi_1 - \frac{1}{\Omega_{ci}}\frac{\partial\nabla_\perp\phi_1}{\partial t}\right) \simeq \mathbf{v}_E + \mathbf{v}_{Pi}, \tag{22}$$

$$\partial_t v_{iz1} \simeq -\frac{e}{m_i}\left(\frac{\partial\phi_1}{\partial z} + \frac{1}{c}\frac{\partial A_{z1}}{\partial t}\right), \tag{23}$$

with v_E and v_{pi} being the electric and polarization drifts, respectively. Similarly, the components of electron velocities in perpendicular and parallel directions can be written, respectively, as,

$$\begin{aligned}\mathbf{v}_{e\perp 1} &\simeq \frac{c}{B_0}\hat{\mathbf{z}}\times\nabla_\perp\left(\phi_1+\frac{\hbar^2}{4m_e n_0 e}\nabla^2 n_{e1}-\frac{2k_B T_{Fe}}{3n_0 e}n_{e1}\right),\\ &\simeq \mathbf{v}_E+\mathbf{v}_{qe}+\mathbf{v}_{De},\end{aligned} \tag{24}$$

$$\partial_t v_{ez1} \simeq \frac{e}{m_e}\frac{\partial}{\partial z}\left(\phi_1+\frac{\hbar^2}{4m_e n_0 e}\nabla^2 n_{e1}-\frac{2k_B T_{Fe}}{3n_0 e}n_{e1}\right)+\frac{e}{m_e c}\frac{\partial A_{z1}}{\partial t}, \tag{25}$$

where v_{qe} and v_{De} are defined as the quantum and diamagnetic type drifts, respectively, $|\partial_t| \ll \omega_{pe}$, ck, and $n_{e0} = n_{i0} = n_0$. The continuity equation for jth species can be expressed as

$$\partial_t n_{j1}+n_0\left(\nabla.\mathbf{v}_{E1}+\nabla.\mathbf{v}_{pj}+\partial_z v_{jz1}\right)=0. \tag{26}$$

The Poisson's equation is

$$\nabla^2\phi_1 = 4\pi e\left(n_{e1}-n_{i1}\right), \tag{27}$$

and the Ampere's law can be written as,

$$\nabla_\perp^2 A_{z1} = \frac{4\pi n_0 e}{c}\left(v_{ez1}-v_{iz1}\right). \tag{28}$$

5.1 Mobile ions

Let us consider that the ions as well as electrons are mobile. The electron and ion continuity equations lead to

$$\frac{\partial}{\partial t}\left(n_{e1}-n_{i1}\right)-n_0\nabla.\mathbf{v}_{pi}+n_0\frac{\partial}{\partial z}\left(v_{ez1}-v_{iz1}\right)=0. \tag{29}$$

Using expressions (22), (27) and (28) in the above equation, we obtain

$$\frac{\partial}{\partial t}\left(\nabla^2+\frac{c^2}{v_A^2}\nabla_\perp^2\right)\phi_1+\frac{c\partial\nabla_\perp^2 A_{z1}}{\partial z}=0, \tag{30}$$

where $v_A = B_0/\sqrt{4\pi n_0 m_i}$ is the speed of Alfven wave, and we have defined the current as $J_{z1} \simeq en_0(v_{iz1}-v_{ez1})$. Ion continuity equation along with (22) and (23) yields,

$$\frac{\partial^2 n_{i1}}{\partial t^2}-\frac{n_0 c}{B_0\Omega_{ci}}\frac{\partial^2\nabla_\perp^2\phi_1}{\partial t^2}-\frac{n_0 e}{m_i}\left(\frac{\partial^2\phi_1}{\partial z^2}+\frac{1}{c}\frac{\partial}{\partial z}\frac{\partial A_{z1}}{\partial t}\right)=0. \tag{31}$$

Eliminating A_{z1} from (30) and (31) and Fourier analyzing, we obtain,

$$\frac{n_{i1}}{n_0} \simeq \frac{1}{\omega^2}\left[-\rho_q^2 k_\perp^2 \omega^2 + \frac{c_q^2 k_z^2}{\omega_A^2}\left\{\omega_A^2 - \omega^2\left(1+\frac{v_A k^2}{c^2 k_\perp^2}\right)\right\}\right]\Phi_1, \tag{32}$$

where we have defined the quantum ion-acoustic type speed as $c_q = \sqrt{T_q / m_i}$, the ion Larmor radius at effective electron temperature as $\rho_q = c_q/\Omega_{ci}$, the Alfven wave frequency as $\omega_A = k_z v_A$ and $\Phi_1 = e\phi_1/T_q$. The effective temperature of electrons (in energy units) is defined as $T_q = (\hbar^2 k^2/4m_e + 2k_B T_{Fe}/3)$, which is a pure quantum mechanical effect. The first term in T_q corresponds to quantum Bohm-de Broglie potential, and the second term represents the electron Fermi energy. So the parameters c_q and ρ_q contain the contribution of both the terms. The Poisson's equation (27) gives

$$\frac{n_{e1}}{n_0} = \frac{n_{i1}}{n_0} - \frac{c_q^2 k^2}{\omega_{pi}^2}\Phi_1, \tag{33}$$

with $\omega_{pi}\left(=\sqrt{4\pi n_0 e^2 / m_i}\right)$ being the ion plasma frequency. Using (32) and (33), we obtain

$$\frac{n_{e1}}{n_0} = \frac{1}{\omega^2}\left[-\rho_q^2 k_\perp^2 \omega^2 + \frac{c_q^2 k_z^2}{\omega_A^2}\left(\omega_A^2 - \omega^2\right) - \frac{c_q^2 k^2}{\omega_{pi}^2}\omega^2\right]\Phi_1. \tag{34}$$

The electron parallel equation of motion leads to,

$$\frac{\partial v_{ez1}}{\partial t} \simeq \frac{e}{m_e}\left(\frac{\partial \phi_1}{\partial z} + \frac{1}{c}\frac{\partial A_{z1}}{\partial t}\right) - \frac{T_q}{m_e n_0}\frac{\partial n_{e1}}{\partial z}. \tag{35}$$

From Ampere's law, we find $v_{ez1} = \frac{c}{4\pi n_0 e}\nabla_\perp^2 A_{z1} + v_{iz1}$. Equation (35) along with (30) leads to

$$\frac{n_{e1}}{n_0} \simeq \frac{1}{\omega_A^2}\left[\omega_A^2 - \left(1+\lambda_e^2 k_\perp^2\right)\left(1+\frac{v_A^2 k^2}{c^2 k_\perp^2}\right)\omega^2 + \frac{m_e}{m_i}\left\{\omega_A^2 - \left(1+\frac{v_A^2 k^2}{c^2 k_\perp^2}\right)\omega^2\right\}\right]\Phi_1. \tag{36}$$

where $\lambda_e = c/\omega_{pe}$ is the electron collisionless skin depth and the small term in the curly brackets appears from ion parallel velocity component. From (34) and (36), we obtain the linear dispersion relation of low frequency coupled electrostatic and electromagnetic modes in dense cold magnetoplasma as,

$$\begin{aligned}\left[\left(1+\frac{v_A^2 k^2}{c^2 k_\perp^2}\right)\omega^2 + \frac{m_e}{m_i\left(1+\lambda_e^2 k_\perp^2\right)}\left\{\left(1+\frac{v_A^2 k^2}{c^2 k_\perp^2}\right)\omega^2 - \omega_A^2\right\} - \frac{c_q^2 k_z^2}{\left(1+\lambda_e^2 k_\perp^2\right)}\left(1+\frac{v_A^2 k^2}{c^2 k_\perp^2}\right)\right]\omega^2 - \\ \frac{\omega_A^2}{\left(1+\lambda_e^2 k_\perp^2\right)}\left[\left(1+\frac{c_q^2 k^2}{\omega_{pi}^2}\right)\omega^2 - c_q^2 k_z^2\right] = \frac{\rho_q^2 k_\perp^2 \omega_A^2}{\left(1+\lambda_e^2 k_\perp^2\right)}\omega^2.\end{aligned} \tag{37}$$

Since $m_e/m_i << 1$, therefore relation (37) reduces to

$$\left[\left(1+\frac{v_A^2k^2}{c^2k_\perp^2}\right)\omega^2-\frac{c_q^2k_z^2}{\left(1+\lambda_e^2k_\perp^2\right)}\left(1+\frac{v_A^2k^2}{c^2k_\perp^2}\right)\right]\omega^2- \frac{\omega_A^2}{\left(1+\lambda_e^2k_\perp^2\right)}\left[\left(1+\frac{c_q^2k^2}{\omega_{pi}^2}\right)\omega^2-c_q^2k_z^2\right]=\frac{\rho_q^2k_\perp^2\omega_A^2}{\left(1+\lambda_e^2k_\perp^2\right)}\omega^2, \tag{38}$$

In the limit $\lambda_e^2k_\perp^2 << 1$, (38) can be written as

$$\left[\left(1+\frac{v_A^2k^2}{c^2k_\perp^2}\right)\omega^2-c_q^2k_z^2\left(1+\frac{v_A^2k^2}{c^2k_\perp^2}\right)\right]\omega^2- \omega_A^2\left[\left(1+\frac{c_q^2k^2}{\omega_{pi}^2}\right)\omega^2-c_q^2k_z^2\right]=\rho_q^2k_\perp^2\omega_A^2\omega^2. \tag{39}$$

If we assume $v_A << ck_\perp / k$ in a quasi-neutral limit $n_{e1} \approx n_{e1}$, (39) leads to

$$\left(\omega^2-c_q^2k_z^2\right)\left(\omega^2-\omega_A^2\right)=\rho_q^2k_\perp^2\omega_A^2\omega^2. \tag{40}$$

The above relation can be found in the recent paper of Saleem et al. (Saleem et al., 2008) if the density inhomogeniety is neglected there and we assume $k_BT_{Fe}n_{e1} << (\hbar^2/4m_e)\nabla^2n_{e1}$. However, the pressure effects of dense ultracold electron plasma are negligible only for wavelengths smaller than the electron Fermi wavelength (Manfredi & Haas, 2001). The plasmas found in the compact astrophysical objects (e.g., white dwarfs and neutron stars) have very high densities and correspondingly, very small interparticle distances. Then the electron statistical pressure effects are significant over the length scales larger than λ_{Be} and it plays a dominant role in wave dispersion.
In case of small parallel ion current, equation (38) yields,

$$\omega^2=\frac{k_z^2v_A^2}{\left(1+\lambda_e^2k_\perp^2\right)}\left(\frac{1}{\left(1+\frac{v_A^2k^2}{c^2k_\perp^2}\right)}+\rho_q^2k_\perp^2\right). \tag{41}$$

Expression (41) shows the effects of electron inertia on shear Alfven wave frequency at quantum scale lengths of electrons in a dense ultracold magnetoplasma. If the electron inertia is neglected, (41) may be written as

$$\omega^2=k_z^2v_A^2\left(\frac{1}{\left(1+\frac{v_A^2k^2}{c^2k_\perp^2}\right)}+\rho_q^2k_\perp^2\right). \tag{42}$$

For $v_A << \frac{ck_\perp}{k}$, we have

$$\omega^2 = k_z^2 v_A^2 \left(1 + \rho_q^2 k_\perp^2\right), \tag{43}$$

which shows the dispersion of shear Alfven wave frequency due to quantum effects associated with electrons in a dense quantum magnetplasma.

5.2 Immobile ions

Now we assume that the ions are immobile. Then Ampere's law leads to

$$v_{ez1} = \frac{c}{4\pi n_0 e}\nabla_\perp^2 A_{z1}. \tag{44}$$

The perpendicular component of electron fluid velocity from (21) becomes

$$\begin{aligned}\mathbf{v}_{e\perp 1} \simeq \frac{c}{B_0}\hat{\mathbf{z}}\times\nabla_\perp\left(\phi_1 + \frac{\hbar^2}{4m_e n_0 e}\nabla^2 n_{e1} - \frac{2k_B T_{Fe}}{3n_0 e}n_{e1}\right) + \\ \frac{c\partial\nabla_\perp}{B_0\Omega_{ce}\partial t}\left(\phi_1 + \frac{\hbar^2}{4m_e n_0 e}\nabla^2 n_{e1} - \frac{2k_B T_{Fe}}{3n_0 e}n_{e1}\right).\end{aligned} \tag{45}$$

Using (27), (28) and (45) in the electron continuity equation, we obtain

$$\frac{\partial}{\partial t}\left[\nabla^2\phi_1 + \frac{\omega_{pe}^2}{\Omega_{ce}}\nabla_\perp^2\left(\phi_1 + \frac{\hbar^2}{4m_e n_0 e}\nabla^2 n_{e1} - \frac{2k_B T_{Fe}}{3n_0 e}n_{e1}\right)\right] + c\frac{\partial}{\partial z}\nabla_\perp^2 A_{z1} = 0, \tag{46}$$

where $\Omega_{ce} = eB_0/m_e c$ is the electron cyclotron frequency. Then using (25), (44) and (46) along with electron continuity equation, we have

$$\begin{aligned}\left(1 - \lambda_e^2\nabla_\perp^2\right)\frac{\partial^2}{\partial t^2}\left[\nabla^2\phi_1 + \frac{\omega_{pe}^2}{\Omega_{ce}}\nabla_\perp^2\left\{\phi_1 + \frac{1}{n_0 e}\left(\frac{\hbar^2\nabla^2}{4m_e} - \frac{2T_{Fe}}{3}\right)n_{e1}\right\}\right] - \\ c^2\nabla_\perp^2\frac{\partial^2}{\partial z^2}\left\{\phi_1 + \frac{1}{n_0 e}\left(\frac{\hbar^2\nabla^2}{4m_e} - \frac{2T_{Fe}}{3}\right)n_{e1}\right\} = 0.\end{aligned} \tag{47}$$

Using (47) with Poisson's equation, and Fourier transforming the resulting expression, we obtain the linear dispersion relation as follows

$$\omega^2 = \frac{v_{Ae}^2 k_z^2\left(1 + \Lambda_q^2 k^2\right)}{\left(1 + \lambda_e^2 k_\perp^2\right)\left[1 + \Lambda_q^2 k^2 + \dfrac{v_{Ae}^2 k^2}{c^2 k_\perp^2}\right]}, \tag{48}$$

where $v_{Ae} = B_0/\sqrt{4\pi n_0 m_e}$, $\Lambda_q = \frac{v_q}{\omega_{pe}}$ and $v_q = \sqrt{T_q/m_e}$. The above equation shows that the wave frequency strongly depends on the quantum nature of electrons which gives rise to the dispersion due to the fermionic pressure and diffraction effects. The last term in square brackets in denominator is negligibly small in general. In an earlier paper, such a relation has been derived in the absence of electron statistical effects (Shukla & Stenflo, 2007).

However, it is seen that the contribution of Fermi pressure is dominant in the wave dispersion as compared with the quantum pressure arising due to Bohm-de Broglie potential.

6. Parametric analysis

Plasmas found in the interior of Jovian planets (Jupiter, Saturn), in compact astrophysical objects e.g., white dwarf stars and neutron stars (Lai, 2001; Bonitz, et al., 2009) as well as dense electron Fermi gas near metallic densities (Manfredi 2005) are typical examples of what is known as quantum plasmas. Here, we numerically analyze the quantum effects on such dense plasmas arising due to Fermi pressure and Bohm-de Broglie potential of electron using typical parameters. The hydrodynamic model is useful for understanding the properties of long wavelength perturbations (> λ_{Fe}) in such systems. The density in the interior of white dwarf stars is of the order of 10^{26}cm^{-3}. For such densities, we have $v_{Fe} \cong 10^8$cm/s, $\lambda_{Fe} \cong 4 \times 10^{-9}$cm and $\bar{r} \cong 2 \times 10^{-9}$cm. For $v_{te} < v_{Fe}$, we have $\lambda_{Be} > \lambda_{Fe}$ and $n_{e0}\lambda_{Be}^3 > 1$. The choice of $k_\perp \sim 10^6$cm^{-1} shows that the wavelength of perturbation is much larger than λ_{Be} and $\bar{r}$. Since we have assumed $k_z << k_\perp$ in deriving the dispersion relation, therefore we take $k_z/k_\perp \sim 0.002$. These parameters are used to numerically analyze the dispersion relation (38). Figs. (1) and (2) shows the dispersive contribution of electrons quantum effects on shear Alfven waves and electrostatic waves, respectively. We use the high magnetic field of the order 10^8G which is within the limits of dense astrophysical plasmas (Lai, 2001). It leads to $\lambda_e^2 k_\perp^2 \cong 0.003$, $\rho_q \cong 0.3 \times 10^{-5}$cm, $v_A \cong 2\times10^6$cm/s and $ck_\perp/k >> v_A$. The overall contribution of the quantum effects in wave dispersion is weak but the effect of electron Fermi pressure is more important as compared with corresponding quantum diffraction term. The dispersion is predominantly due to the Fermi pressure of electrons. It may be mentioned here that T_{Fe} is a function of density and assumes very large values. The dispersion relation (48) is plotted in Fig. (3) for relatively less dense plasmas with $n_{e0} \simeq 10^{24}$cm^{-3}. Such densities are relevant to

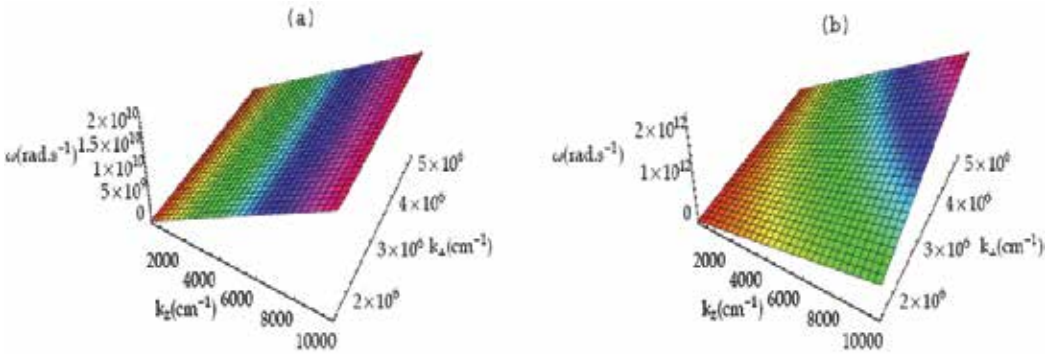


Fig. 1. (Color online). The Alfven wave frequency ω is plotted vs wave numbers k_z and $k_\perp$ for mobile ions using typical parameters of white dwarf stars i.e., $n_{e0} \cong 10^{26}$cm^{-3} and $B_0 \cong 10^8 G$. Case (a) corresponds to the frequency when the effect of Fermionic pressure is not considered and case (b) when considered. It is evident from (b) that the wave frequency is increased by two orders of magnitude when electron fermionic pressure is taken into account.

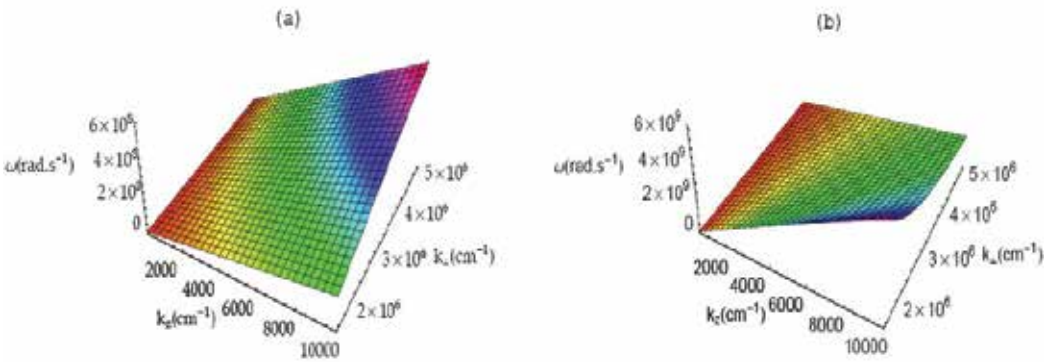


Fig. 2. (Color online). The quantum ion-acoustic wave frequency ω is plotted from (38) against the wave numbers k_z and $k_\perp$ using the same parameters as in Fig. 1. Case (a) refers to the wave frequency when the effect of Fermi pressure is not included, whereas case (b) when included.

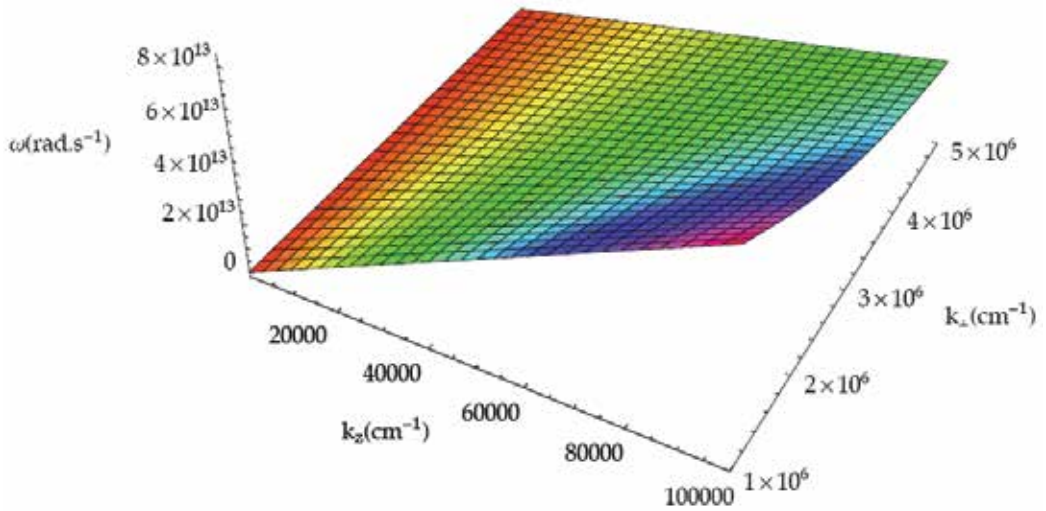


Fig. 3. (Color online). The linear dispersion relation (48) for immobile ions is plotted with $n_{e0} \cong 1 \times 10^{24}\text{cm}^{-3}$ and $B_0 \cong 1 \times 10^8 G$. The wave frequency is much higher in the present case becase of the dynamics of electrons with the background of stationary ions.

the dense Fermi gas of electrons with the background of stationary positive ions (Manfredi 2005) as well as dense plasmas in the outer layers of white dwarfs. For $\omega << \Omega_{ce}$, and $k_z \sim 0.002 k_\perp$, it is found that $\Lambda_q^2 k_\perp^2 \cong 0.03$ which is due to the dominant contribution of electron Fermi pressure since the dispersive effects due to quantum diffraction term are negligibly small. The approximations and the assumptions made in deriving the dispersion relation (48) are satisfied in the parameteric range used for numerical work. The fluid model may be used for physical understanding of the waves in dense plasmas even if $\rho_q^2 k_\perp^2 > 1$ since we have $T << T_{Fe}$. The results also indicate that the well known electron and ion plasma wave spectra in dense plasmas is significantly modified by quantum nature of electrons.

7. Concluding remarks

The self-consistent dynamics of low frequency linear modes in a dense homogenous electron ion quantum magnetoplasma has been investigated. Using quantum hydrodynamic approach, the generalized dispersion relation is obtained to investigate the properties of electrostatic and electromagnetic waves. The dispersion relation for coupled electrostatic and shear Alfven waves reveals the dispersive effects associated with Fermi pressure of electrons and quantum Bohm-de Broglie potential. For illustration, we have applied our results to the dense magnetoplasma with $n_0 \sim 10^{26}\text{cm}^{-3}$ which seems to be possible locally in the dense astrophysical objects e.g., white dwarf stars (Lai, 2001; Bonitz et al., 2009). The dispersion relation has also been obtained for dense plasma with the background of stationary ions in the presence of very high magnetic field. In the case of immobile ions, only electromagnetic part survives and the dispersion enters through the electron quantum effects. This case may be of interest for the self-consistent dynamics of electrons in dense plasma systems with background of immobile ions. The analysis of the waves in different limiting cases shows that the contribution of electron Fermi pressure is dominant over the quantum pressure due to Bohm-de Broglie potential for mobile as well as immobile ions. It is also shown that the the quantum behavior of electrons in high density low temperature plasmas modifies wave frequency at short length scales.

The field of quantum plasmas is extremely rich and vibrant today. The investigation of quantum plasma oscillations in unmagnetized and magnetized charged particle systems of practical importance has been a subject of interest in the recent years. In dense astrophysical plasmas such as in the atmosphere of neutron stars or the interior of massive white dwarfs, the magnetic field and density varies over a wide range of values. For instance, the magnetic field is estimated to be varying from kilogauss to gigagauss (petagauss) range in white dwarfs (neutron stars) and the density n_0 lies in the range 10^{23} – 10^{28}cm^{-3}. The quantum corrections to magnetohydrodynamics can be experimentally important in such systems (Haas, 2005). Similarly, in magnetars, and in the next generation intense laser-solid density plasma interaction experiments, one would certainly have degenerate positrons along with degenerate electrons. Apart from theoretical perspective, such plasmas also holds promises of providing future technologies.

8. Acknowledgement

S. A. K. acknowledges National Centre for Physics (NCP), Quaid-i-Azam University Campus, Islamabad, Pakistan for associate membership as well as Higher Education Commission Pakistan for providing support through Indigenous PhD fellowship scheme.

9. References

Ali, S. (2008). *Waves and Instabilities in Quantum Plasmas*, Ph.D. Thesis (Umea University Sweden); see also http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva 1520

Bezzerides, B. and DuBois, D.F. (1972). Quantum electrodynamics of nonthermal relativistic plasmas: Kinetic theory, *Ann. Phys.* (N. Y.) 70, 10

Bohm, D. and Pines, D. (1953). A collective description of electron interactions: III. Coulomb interactions in a degenerate electron gas, *Phys. Rev.* 92 609

Bonitz, M. (1998). "*Quantum Kinetic Theory*" (B. G. Teubner, Stuttgart-Leipzig).

Bonitz, M., Filinov, A., Boning, J. and Dufty, J.W. (2009). *Introduction to quantum plasmas*, in: *Introduction to Complex Plasmas*, edited by Bonitz, M., Horing, N., Meichsner, J. and Ludwig, P. (Springer Series "*Atomic, Optical and Plasma Physics*", Springer, Berlin).

Brodin G. and Marklund, M. (2007). Spin magnetohydrodynamics, *New J. Phys.* 9, 277

Brodin G. and Marklund M. (2008) *New Aspects of Plasma Physics* in: Proceedings of *the AS-ICTP Summer College on Plasma Physics*, Eds. Shukla, P.K., Stenflo, L. and Eliasson, B.(World Scientific, Singapore) pp. 26-34

Haas, F., Garcia, L.G., Goedert, J. and Manfredi, G. (2003). Qunatum ion-acoustic waves, *Phys. Plasmas* 10, 3858

Haas, F. (2005). A magnetohydrodynamic model for quantum plasmas, *Phys. Plasmas* 12, 062117

Khan, S. A. and Saleem, H. (2009). Linear coupling of Alfven waves and acoustic-type modes in dense quantum magnetoplasmas, *Phys. Plasmas* 16, 052109

Khan, S.A., Mahmood, S. and Saleem, H. (2008). Linear and nonlinear ion-acoustic waves in very dense magnetized plasmas, *Phys. Plasmas* 15, 082303

Khan, S.A., Mushtaq, A. and Waqas, M. (2008). Dust ion-acoustic waves in magnetized quantum dusty plasmas with polarity effect, *Phys. Plasmas* 15, 013701

Khan, S.A., Mahmood, S. and Ali, S. (2009). Quantum ion-acoustic double layers in unmagnetized dense electron-positron-ion plasmas, *Phys. Plasmas* 16, 044505

Khan, S.A., Mahmood, S. and Mirza, Arshad. M. (2008). Cylindrical and spherical dust ion-acoustic solitary waves in quantum plasmas, *Phys. Lett. A.* 372, 148-153

Klimontovich, Y. and Silin, V.P. (1952). On the spectra of interacting particles, *Zh. Eksp. Teor. Fiz.* 23, 151

Lai, D. (2001). Matter in strong magnetic fields, *Rev. Mod. Phys.* 73, 629

Landau, L.D. and Lifshitz, E.M. (1980). *Satistical Physics Part 1* (Butterworth-Heinemann, Oxford)

Lunden, J., Zamanian, J., Marklund, M. and Brodin, J. (2007). Short wavelength electromagnetic propagation in magnetized quantum plasmas, *Phys. Plasmas* 14, 062112

Madelung, E.Z., (1926). Quantentheorie in hydrodynamischer Form, *Physik 40*, 322-326 (In German)

Marklund, M. and Brodin, G. (2007). Dynamics of spin-1/2 quantum plasmas, *Phys. Rev. Lett* 98, 025001

Manfredi, G. and Haas, F. (2001). Self-consistent fluid model for a quantum electron gas, *Phys. Rev. B* 64, 075316

Manfredi, G. (2005). How to model quantum plasmas, *Fields Inst. Commun.* 46, 263

Melrose, D. B. (2007). *Quantum Plasmdynamics-Unmagnetized Plasmas* (Springer-Verlag New York)

Pines, D. (1961). Quantum Plasma Physics; Classical and quantum plasmas, *J. Nucl. Energy,* Part C 2, 5

Saleem, H, Ahmad Ali and Khan, S.A. (2008). Low frequency electrostatic and electromagnetic modes of ultracold magnetized nonuniform dense plasmas, *Phys. Plasmas* 15, 094501

Shukla, P.K. and Eliasson, B. (2010). Nonlinear Aspects of Quantum Plasma Physics, Phys. Usp. 53, 51

Shukla, P.K. and Stenflo, L. (2006). Shear Alfven modes in ultracold quantum magnetoplasmas, *New J. Phys.* 8, 111

Shukla, P. K. (2009). A new spin on quantum plasmas, *Nature Physics* 5, 92

Zyrianov, P.S., Okulov, V.I. and Silin, V. P. (1969). Quantum electron spin-acoustic waves, *JETP Lett.* 9, 371

5

Research of the Scattering of Non-linearly Interacting Plane Acoustic Waves by an Elongated Spheroid

Iftikhar B. Abbasov
Taganrog Technological Institute, Southern Federal University
Russia

1. Introduction

For the first time the problem of acoustic wave scattering on elongated spheroids was stated in works [Cpence & Ganger, 1951], [Burke, 1966], [Kleshchyov & Sheiba, 1970]. Work [Cpence & Ganger, 1951] considers the problem of sound scattering on a elongated spheroid with various boundary conditions. Work [Burke, 1966] considers the problem of sound scattering on a rigid spheroid in the long-wave approximation. Work [Kleshchyov & Sheiba, 1970] considers the problems of sound wave scattering on a elongated spheroid where angular characteristics for sound wave scattering on a soft and rigid elongated spheroid were found.

The studies of acoustic field of spheroidal radiators were considered in works [Chertock, 1961], [Andebura, 1969], where acoustic field, radiation impedance of arbitrary elongated spheroid were defined. Work [Andebura, 1976] considers integral characteristics of the interaction between spheroid and incident sound wave with different spheroid orientations relating to propagation direction of incident wave.

The diffraction problem of plane sound wave on elongated rigid revolution bodies within the field of small values of wave rate is considered in work [Fedoryuk, 1981], where scattering amplitude asymptotics are found. Work [Tetyuchin & Fedoryuk, 1989] describe plane sound wave diffraction on a elongated rigid revolution body in liquid, give calculation scattering diagrams on a steel and aluminium spheroid with lateral incidence of a plane wave.

Work [Boiko, 1983] considers the case of plane wave scattering on a thin revolution body that differs from medium with its contractiveness and density. The principal term of evanescent field asymptotics was found, angular characteristics for plane wave scattering by rigid elongated spheroid in geometrical scattering field were given.

The questions of sound scattering by gas-filled spheroidal fish-maw are considered in works [Haslett, 1962], [Babailov & Kanevskyi, 1988]. A fish-maw is given as a elongated soft spheroid, frequency-angle characteristics of inverse scattering are given as well as resonance characteristic spheroidal maw.

In recent decades a number of works devoted to sound scattering on spheroids were published by Kleshchyov A.A. [Kleshchyov, 1986; 1992; 2004]. These works are devoted to studies of sound scattering on fish and fish flock maws near surface and bottom. A gas maw

is approximated by soft elongated spheroid; the flock is divided to a definite number of scatterers. The scattered sound is formed shape at the cost of signal interference, scattered by separate scatterers, on condition that the distance between scatterers is more than scatterer's dimensions. The fish flock is near one of the two boundarys: either water-air or water-bottom. At that it is assumed that the boundarys are plane and ideal, the air is replaced by vacuum, the bottom is ideally firm. Scattering characteristics of separate maws in the form of soft elongated spheroid are calculated with ranges from angular and radially spheroidal functions.

The problem of plane acoustic wave scattering on spheroidal shells was considered in works [Werby & Green, 1987], [Weksler et al., 1999]. These works study the surface waves directly on scatterers themselves, they describe frequency dependence of inverse scattering in farfield.

Some questions of experimental studies of acoustic wave scattering on elongated form bodies are considered in works [Stanton, 1989], [Lebedev & Salin, 1997].

Last years some works appeared, written by [Belkovich et al., 2002], [Kuzkin, 2003], devoted to acoustic wave scattering on spheroids in waveguides. This problem appears when active acoustic monitoring of Cetacea population in world's oceans and seas. The model problem of sound wave diffraction on elongated soft spheroids (Cetacea) is considered, when locating them in ocean waveguide. Diffuse sound field is analysed as a parameter function: spheroid dimensions, its position relative to sound source and detector, vertical profile of sound speed in waveguide, acoustic parameters of bottom boundary.

Except the works above devoted to linear scattering with spheroids, there are several works devoted to nonlinear acoustic spectroscopy [Guyer & Johnson, 1999], [Lebedev et al., 2005]. Work [Guyer & Johnson, 1999] considers the problem of nonlinear acoustic defect diagnosis in materials and formations. Work [Lebedev et al., 2005] is devoted to solving the problem of nonlinear defect acoustic spectroscopy in geomaterials. A cavity model in the form of oblate spheroid is taken as a defect. The crack on a thin bar is shown as an example of defect isolation problem solving.

However the problem of interacting acoustic wave scattering on elongated spheroid has not been under study before. It becomes one of current interest when using parametric acoustic array for remote diagnostics of aquatic medium. This problem can also appear within biological environment diagnostics, where high nonlinear nature and nonhomogeneities are in the near field of a radiating unit. This chapter studies the scattering problem of nonlinearly interacting plane acoustic waves on rigid elongated spheroid.

2. Wave problems in elongated spheroidal coordinates

When solving the problems of wave diffraction on elongated form bodies, confocal coordinates, spheroidal in particular, are often used. These coordinates are used within studying acoustic wave radiating and scattering by ellipsoids, cigar-shaped bodies, as well as within studying diffraction by circular apertures [Skudrzyk, 1971].

When studying diffraction on cigar-shaped bodies, the elongated spheroidal coordinates system is used. Coordinate surfaces are spheroids $\xi = const$ and two-sheeted hyperboloids $\eta = const$. The elongated spheroid is formed by ellipse rotation round its longer axis (Fig.1). Within ellipse rotation round shorter axis, the oblate spheroid is formed. A great number of revolution body surfaces can be described with the help of spheroidal coordinate systems. Orb and cylinder can be considered as special cases of spheroidal surfaces, a continued thin bar and disks are confluent spheroids.

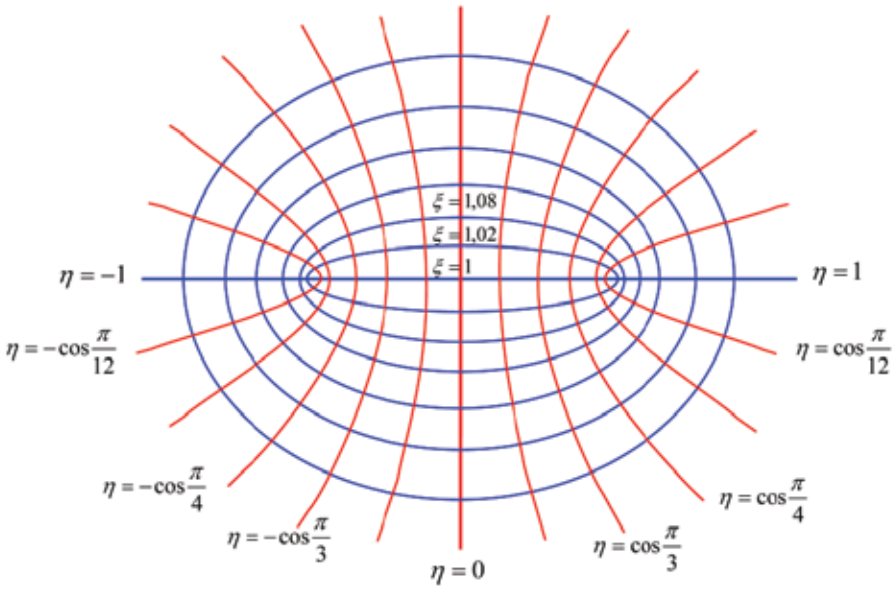


Fig. 1. Elliptic coordinate system

Ellipse is a geometrical locus, their sum of the distances r_1 and r_2 from two given points (focal points) is constant [Abramovitz & Stegun, 1971]:

$$r_1 + r_2 = const = \xi d\,, \qquad \text{or} \quad \frac{r_1 + r_2}{d} = \xi\,,$$

where d - distance between ellipse focuses, ξ - radial coordinate.
The length of ellipse longer axis L is (Fig.1) — $L = r_1 + r_2$, or $L = \xi d$, shorter axis D is defined from formula $D = d\sqrt{(\xi^2 - 1)}$.
Parameter ξ is an ellipse eccentricity measure - $e = d/L$, it follows that $\xi = 1/e$. With $\xi = 1$ ellipse degenerates into the interval with length d, with $\xi = \infty$ ellipse grades into circle of infinite radius. For long distances product ξd is equal in practice to duplicated distance from the origin of axis system.
The coordinate, equivalent to coordinate θ in a polar system, is obtained with the help of confocal hyperboloids (Fig.1)

$$\frac{r_1 - r_2}{d} = \eta = \cos\theta\,,$$

where η - angular coordinate.
The hyperbolic curve is a geometrical locus, their difference of the distances r_1 and r_2, from two given points (focal points F_1 and F_2) is constant (Fig.2). In spherical coordinates angle θ is an angle between radius-vector of observation point $M(\xi,\eta,\varphi)$ and coordinate axis x (Fig.3).
With larger coordinate value ξ spheroidal coordinates grade into spherical ones, and angle θ in formula $\eta = \cos\theta$ corresponds to asymptote angle for hyperbolic curve η.

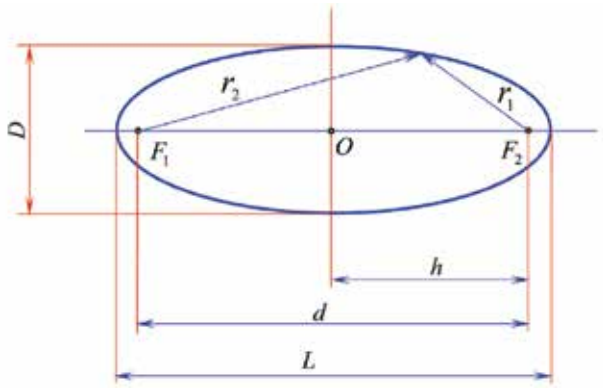


Fig. 2. Ellipse basic parameters

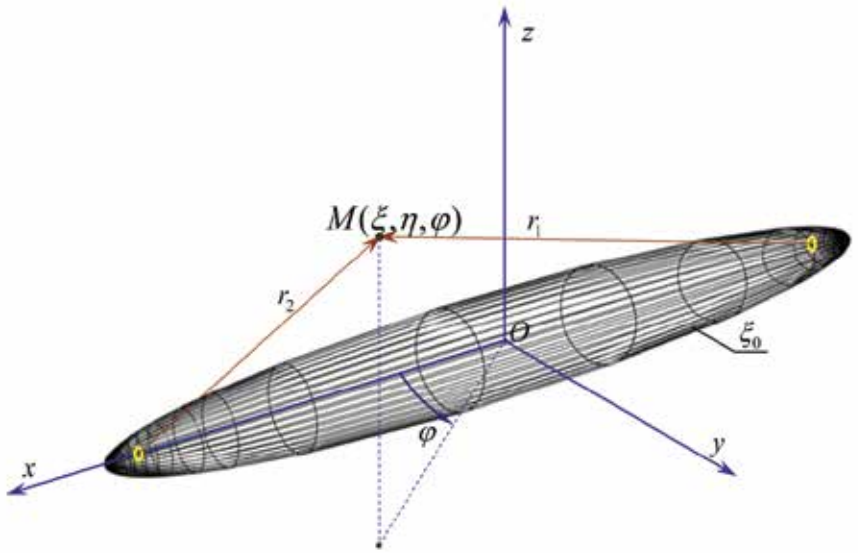


Fig. 3. Scatterer in elongated spheroidal coordinates system

3. Theory

To present the problem, the system of elongated spheroidal coordinates ξ, η, φ was chosen. The foci of the spheroid coincide with the foci of the spheroidal coordinate system. The spheroid is formed by the ellipse ξ_0 rotated about a major axis, which coincides with the x - axis of the Cartesian system. The geometry of the problem is presented in Fig.4. The coordinate surfaces are: for the spheroids - ξ=const and for the two-sheeted hyperboloids - η=const.

Elongated spheroidal coordinates are related to Cartesian coordinates by the following expressions [Tikhonov & Samarskyi, 1966]:

$$x = h_0 \xi \eta\,, \quad y = h_0 \sqrt{(\xi^2 - 1)(1 - \eta^2)}\cos\varphi\,, \quad z = h_0 \sqrt{(\xi^2 - 1)(1 - \eta^2)}\sin\varphi\,,$$

where $h_0 = d/2$, and d is the interfocal distance. Spheroidal coordinates ξ, η, φ are considered within the limits: $1 \le \xi < \infty$; $-1 \le \eta \le 1$; $0 \le \varphi \le 2\pi$.

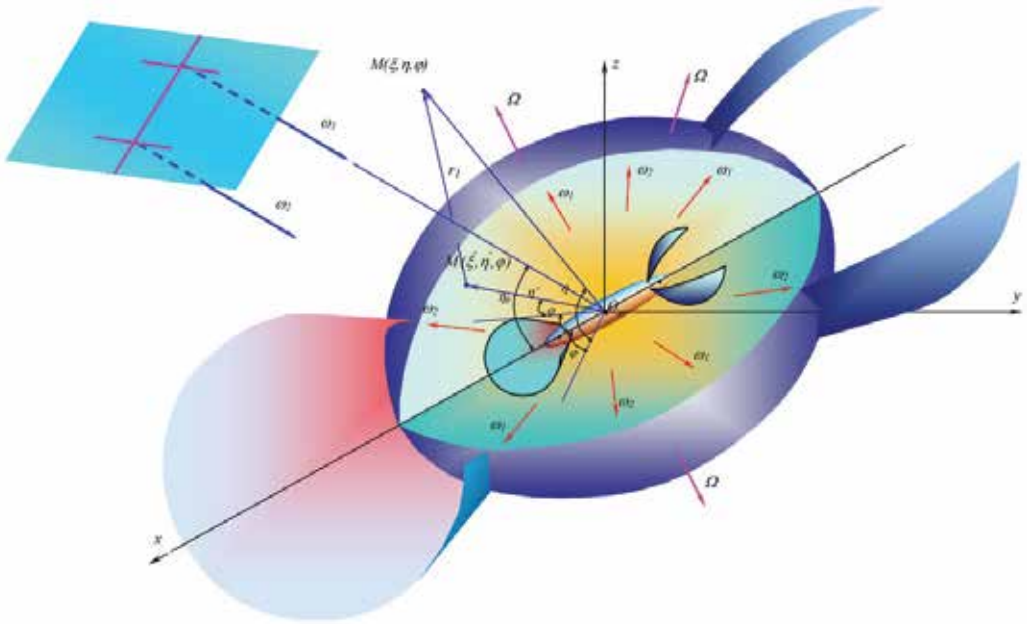


Fig. 4. Geometry of the problem

The perfect spheroid was put into homogeneous medium. The spheroid's surface is characterized by the coordinate ξ_0. Assuming that interacting plane high-frequency acoustic waves of the unit pressure amplitude falls on the spheroid at an arbitrary polar angle θ_0 (θ_0=arccosη_0) and an azimuthal angle φ_0, we express the acoustic pressure as:

$$p_{ni} = \exp\left[-i(k_n r_0 \cos\theta_0 - \omega_n t)\right], \tag{1}$$

where k_n - is the wave number, $n = 1,2$ according to the waves with frequencies ω_1 and ω_2, and r_0 is the radius-vector of the polar coordinate system.
Consider an incident plane wave in the spheroidal coordinate system [Skudrzyk, 1971]:

$$\exp\left[i(\omega_n t - k_n r_0 \cos\theta_0)\right] = -2\exp(i\omega_n t)\sum_{m=0}^{\infty}\sum_{l\ge m}^{\infty} i^{-l}\,\overline{S_{ml}}(k_n h_0, \eta_0)\overline{S_{ml}}(k_n h_0, \eta) R_{ml}^{(1)}(k_n h_0, \xi)\cos m(\varphi - \varphi_0),$$

where $\overline{S_{ml}}(k_n h_0, \eta)$ is the normalized angular first-order function and $R_{ml}^{(1)}(k_n h_0, \xi)$ is the radial spheroidal first-order function.
After the plane wave scattering on the spheroid, the scattered spheroidal wave of pressure will propagate as an outgoing wave [Kleshchyov & Klyukin, 1987]

$$p_{ns}(\xi, \eta, \varphi) = 2\exp(i\omega_n t)\sum_{m=0}^{\infty}\sum_{l\ge m}^{\infty} A_{ml}(k_n h_0, \xi_0)\overline{S_{ml}}(k_n h_0, \eta) R_{ml}^{(3)}(k_n h_0, \xi)\cos m\varphi, \tag{2}$$

where the coefficient $A_{ml}(k_n h_0, \xi_0)$ is dependent on boundary conditions on the spheroid surface, and $R_{ml}^{(3)}(k_n h_0, \xi)$ is the radial spheroidal third-order function.
In this case the spheroid is considered to be acoustically rigid, so the Neumann boundary condition must be applied on the surface:

$$\left(\frac{\partial p_{ni}}{\partial n} + \frac{\partial p_{ns}}{\partial n} \right)\Bigg|_{\xi=\xi_0} = 0, \tag{3}$$

and the coefficient $A_{ml}(k_n h_0, \xi_0)$ will be determined by the following expression:

$$A_{ml}(k_n h_0, \xi_0) = -i^l \varepsilon_m \overline{S_{ml}}(k_n h_0, \eta_0) \frac{R_{ml}^{(1)'}(k_n h_0, \xi_0)}{R_{ml}^{(3)'}(k_n h_0, \xi_0)}$$

where $R_{ml}^{(1)'}(k_n h_0, \xi_0)$ and $R_{ml}^{(3)'}(k_n h_0, \xi_0)$ are the derivatives of the first- and third-order functions, $\varepsilon_m = 1$, for $m = 0$, $\varepsilon_m = 2$, for $m > 0$.
With the appearance of the scattered spheroidal wave, the total acoustic pressure of the primary field around the spheroid will have the form:

$$p^{(1)} = p_{ni} + p_{ns} = \left[\sum_{m=0}^{\infty} \sum_{l \geq m}^{\infty} B_{ml}(k_n h_0) \exp\left[i(\omega_n t - l\pi/2) \right] + \sum_{m=0}^{\infty} \sum_{l \geq m}^{\infty} D_{ml}(k_n h_0) \exp\left[i(\omega_n t - m\varphi) \right] \right] +$$

$$+ \left[\sum_{m=0}^{\infty} \sum_{l \geq m}^{\infty} B_{ml}(k_n h_0) \exp\left[-i(\omega_n t - l\pi/2) \right] + \sum_{m=0}^{\infty} \sum_{l \geq m}^{\infty} D_{ml}(k_n h_0) \exp\left[-i(\omega_n t - m\varphi) \right] \right] \tag{4}$$

where
$B_{ml}(k_n h_0) = 2\overline{S_{ml}}(k_n h_0, \eta_0) \overline{S_{ml}}(k_n h_0, \eta) R_{ml}^{(1)}(k_n h_0, \xi) \cos m(\varphi - \varphi_0)$,
$D_{ml}(k_n h_0) = 2A_{ml}(k_n h_0, \xi_0) \overline{S_{ml}}(k_n h_0, \eta_0) R_{ml}^{(3)}(k_n h_0, \xi) \cos m\varphi$.
To solve the problem of the non-linear interaction of the primary high-frequency waves, we combine expression (4) with its complex-conjugate part.
Nonlinear wave processes between incident and scattered waves surrounding the spheroid can be described with the inhomogeneous wave equation [Novikov et al., 1987]:

$$\nabla^2 p^{(2)} - \frac{1}{c_0^2} \frac{\partial^2 p^{(2)}}{\partial t^2} = -Q = -\frac{\varepsilon}{c_0^4 \rho_0} \frac{\partial^2 p^{(1)2}}{\partial t^2}, \tag{5}$$

where Q is the volume density of the sources of secondary waves, c_0 is the sound velocity in the medium, ε is the quadratic nonlinearity parameter, ρ_0 is the density of the unperturbed medium, and $p^{(1)}$ and $p^{(2)}$ are the total acoustic pressures of the primary and secondary fields.
It is important to note that the waves of the primary field are the high frequency waves: incident plane waves p_i and scattered spheroidal waves p_s with angular frequencies ω_1 and ω_2. The waves of the secondary field are the waves that appear as a result of the non-linear interaction of initial high frequency waves. This includes the difference frequency

wave $\omega_2 - \omega_1 = \Omega$, the summation frequency wave $\omega_2 + \omega_1$, and the second harmonic waves $2\omega_1$, $2\omega_2$.

The wave equation (5) is solved by the method of successive approximations. In the first approximation, the solution is represented by the expression (4) for the total acoustic pressure of the primary field $p^{(1)}$. To determine solution in the second approximation $p^{(2)}$, the right-hand side of equation (5) should feature four frequency components: second harmonics of the incident waves ($2\omega_1$, $2\omega_2$) and ($\omega_1 + \omega_2$, $\omega_2 - \omega_1 = \Omega$).

The expression for the volume density of secondary waves sources at the difference frequency Ω is:

$$Q_- = \frac{2\Omega^2\varepsilon}{c_0^4\rho_0}\left[\sum_{m=0}^{\infty}\sum_{l\geq m}^{\infty} B_{ml}(k_1h_0)B_{ml}(k_2h_0)\cos\Omega t + \sum_{m=0}^{\infty}\sum_{l\geq m}^{\infty} B_{ml}(k_1h_0)D_{ml}(k_2h_0)\cos(\Omega t + l\pi/2 - m\varphi) + \right.$$

$$\left. + \sum_{m=0}^{\infty}\sum_{l\geq m}^{\infty} B_{ml}(k_2h_0)D_{ml}(k_1h_0)\cos(\Omega t + m\varphi - l\pi/2) + \sum_{m=0}^{\infty}\sum_{l\geq m}^{\infty} D_{ml}(k_1h_0)D_{ml}(k_2h_0)\cos\Omega t\right]. \quad (6)$$

To solve the inhomogeneous wave equation (5) with the right-hand side given by equation (6) in the second approximation, we seek the solution in the complex form

$$p_-^{(2)} = \frac{1}{2}P_-^{(2)}\exp(i(\Omega t + \delta) + (c.c.)). \quad (7)$$

Substitution of the expression (7) into the inhomogeneous wave equation (5) gives the inhomogeneous Helmholtz equation:

$$\nabla^2 P_-^{(2)} + k_-^2 P_-^{(2)} = -q_-(\xi,\eta,\varphi), \quad (8)$$

where k_- is the wave number of the difference frequency Ω, and

$$q_-(\xi,\eta,\varphi) = \frac{2\Omega^2\varepsilon}{c_0^4\rho_0}\left[\sum_{m=0}^{\infty}\sum_{l\geq m}^{\infty} B_{ml}(k_1h_0)B_{ml}(k_2h_0)\exp(i\Omega t) + \right.$$

$$+ \sum_{m=0}^{\infty}\sum_{l\geq m}^{\infty} B_{ml}(k_1h_0)D_{ml}(k_2h_0)\exp\left[i(\Omega t + l\pi/2 - m\varphi)\right] +$$

$$+ \sum_{m=0}^{\infty}\sum_{l\geq m}^{\infty} B_{ml}(k_2h_0)D_{ml}(k_1h_0)\exp\left[i(\Omega t + m\varphi - l\pi/2)\right] +$$

$$\left. + \sum_{m=0}^{\infty}\sum_{l\geq m}^{\infty} D_{ml}(k_1h_0)D_{ml}(k_2h_0)\exp(i\Omega t)\right].$$

The solution to the inhomogeneous Helmholts equation (8) has the form of a volume integral of the product of the Green function with the density of the secondary wave sources [Novikov et al., 1987] [Lyamshev & Sakov, 1992]:

$$P_{-}^{(2)}(\xi,\eta,\varphi)=\int_{V} q_{-}(\xi',\eta',\varphi')G(r_1)h_{\xi'}h_{\eta'}h_{\varphi'}d\xi' d\eta' d\varphi' \ , \tag{9}$$

where $G(r_1)$ is the Green function, r_1 is the distance between the current point of the volume $M'(\xi',\eta',\varphi')$ and the observation point $M(\xi,\eta,\varphi)$ (Fig.4), and $h_{\xi'}$, $h_{\eta'}$, $h_{\varphi'}$ are the scale factors [Corn & Corn, 1968]:

$$h_{\xi'}=h_0\sqrt{\frac{\xi'^2-\eta'^2}{\xi'^2-1}}\ ,\ h_{\eta'}=h_0\sqrt{\frac{\xi'^2-\eta'^2}{1-\eta'^2}}\ ,\ \ h_{\varphi'}=h_0\sqrt{(\xi'^2-1)(1-\eta'^2)}\ .$$

In the far field $r' << r$, the Green function is determined by the asymptotic expression

$$G(r_1)=\exp(-ik\_r_1)/r_1\approx\exp\left[-ik\_\left(h_0\xi-h_0\xi'\eta\eta'-h_0\xi'\sqrt{(1-\eta^2)(1-\eta'^2)}\right)\right]\Big/h_0\xi\ .$$

The integration in equation (9) is performed over the volume V occupied by the second wave sources and bounded in the spheroidal coordinates by the relations

$$\xi_0\le\xi'\le\xi_S\ ,\qquad -1\le\eta'\le 1\ ,\qquad 0\le\varphi'\le 2\pi\ .$$

This volume has the form of a spheroidal layer of the medium, stretching from the spheroid's surface to the non-linear interaction boundary (Fig.4). An external spheroid with coordinate ξ_S appears to be the boundary of this area. Coordinate ξ_S is defined by the size of the non-linear interaction area between the initial high-frequency waves. This size is inversely proportional to the coefficient of viscous sound attention associated with the corresponding pumping frequency. Beyond this area, the initial waves are assumed to attenuate linearly.

After the integration with respect to coordinates φ' and η' (considering the high-frequency approximation), equation (9) takes the form

$$P_{-}^{(2)}(\xi,\eta,\varphi)=P_{-1}^{(2)}(\xi,\eta,\varphi)+P_{-2}^{(2)}(\xi,\eta,\varphi)+P_{-3}^{(2)}(\xi,\eta,\varphi)+P_{-4}^{(2)}(\xi,\eta,\varphi)=$$

$$=C_{-}\frac{1}{k\_h_0\eta}\left[\int_{\xi_0}^{\xi_S}T\xi'\sin(k\_h_0\xi'\eta)d\xi'-\int_{\xi_0}^{\xi_S}T\frac{\sin(k\_h_0\xi'\eta)}{\xi'}d\xi'\right], \tag{10}$$

where

$$C_{-}=\frac{8h_0^2\pi\Omega^2\varepsilon\exp(-ik\_h_0\xi)}{c_0^4\rho_0\xi},$$

$$T=\left[\sum_{m=0}^{\infty}\sum_{l\ge m}^{\infty}B_{ml}(k_1h_0)B_{ml}(k_2h_0)+\sum_{m=0}^{\infty}\sum_{l\ge m}^{\infty}B_{ml}(k_1h_0)D_{ml}(k_2h_0)\exp\left[i(l\pi/2-m\varphi)\right]+\right.$$

$$\left.+\sum_{m=0}^{\infty}\sum_{l\ge m}^{\infty}B_{ml}(k_2h_0)D_{ml}(k_1h_0)\exp\left[i(m\varphi-l\pi/2)\right]+\sum_{m=0}^{\infty}\sum_{l\ge m}^{\infty}D_{ml}(k_1h_0)D_{ml}(k_2h_0)\right]$$

(from here on, the time factor $\exp(i\Omega t)$ is omitted).

The expression (10) for the total acoustic pressure of the difference-frequency wave $P_{-}^{(2)}(\xi,\eta,\varphi)$ consists of four spatial components. The first component $P_{-1}^{(2)}(\xi,\eta,\varphi)$ corresponds to the part of the acoustic pressure of the difference-frequency wave, that is formed in the spheroidal layer of the non-linear interaction area by the incident high-frequency plane waves ω_1 and ω_2. The second component $P_{-2}^{(2)}(\xi,\eta,\varphi)$ describes the interaction of the incident plane wave of frequency ω_1 with the scattered spheroidal wave of frequency ω_2. The third component $P_{-3}^{(2)}(\xi,\eta,\varphi)$ corresponds to the interaction of the scattered plane wave of frequency ω_2 with the scattered spheroidal wave of ω_1. The fourth component $P_{-4}^{(2)}(\xi,\eta,\varphi)$ characterises the interaction of two scattered spheroidal waves with frequencies ω_1 and ω_2.

4. Results

To obtain the final expression of the total acoustic pressure of the difference-frequency wave $P_{-}^{(2)}(\xi,\eta,\varphi)$, consider the first spatial component $P_{-1}^{(2)}(\xi,\eta,\varphi)$ from equation (10), which characterises the non-linear interaction between incident plane waves of highfrequency:

$$P_{-1}^{(2)}(\xi,\eta,\varphi)=\frac{C_{-}}{k_{-}h_0\eta}\left[\int_{\xi_0}^{\xi_S}\sum_{m=0}^{\infty}\sum_{l\geq m}^{\infty}B_{ml}(k_1h_0)B_{ml}(k_2h_0)\xi'\sin(k_{-}h_0\xi'\eta)d\xi'-\right.$$

$$\left.-\int_{\xi_0}^{\xi_S}\sum_{m=0}^{\infty}\sum_{l\geq m}^{\infty}B_{ml}(k_1h_0)B_{ml}(k_2h_0)\frac{\sin(k_{-}h_0\xi'\eta)}{\xi'}d\xi'\right]. \qquad (11)$$

It should be noted that this is the only component that gives no information about the scatterer. The boundaries of the integration layer are directly defined by the elongated spheroid shape.

Using representation of the plane wave in the spheroidal coordinate system and substituting $B_{ml}(k_nh_0)$, the expression (11) takes the form

$$P_{-1}^{(2)}(\xi,\eta,\varphi)=\frac{C_{-}}{k_{-}h_0\eta}\left[\int_{\xi_0}^{\xi_S}\exp\left[-ik_{-}h_0\xi'\eta\right]\xi'\sin(k_{-}h_0\xi'\eta)d\xi'-\int_{\xi_0}^{\xi_S}\exp\left[-ik_{-}h_0\xi'\eta\right]\frac{\sin(k_{-}h_0\xi'\eta)}{\xi'}d\xi'\right]. \qquad (12)$$

After the final integration with respect to the coordinate ξ', the expression for the first component (12) has the form

$$P_{-1}^{(2)}(\xi,\eta,\varphi)=P_{-11}^{(2)}+P_{-12}^{(2)}+P_{-13}^{(2)}+P_{-14}^{(2)}, \qquad (13)$$

where

$$P_{-11,-12}^{(2)}\approx\mp\frac{C_{-}}{2k_{-}^2h_0^2\eta(\eta_0\mp\eta)}\left[\xi_S\exp\left[ik_{-}h_0(\eta_0\mp\eta)\xi_S\right]-\xi_0\exp\left[ik_{-}h_0(\eta_0\mp\eta)\xi_0\right]\right],$$

$$P_{-13,-14}^{(2)}\approx\mp\frac{C_{-}}{2i}\left[-\mathrm{Ei}\left[-ik_{-}h_0(\eta_0\mp\eta)\xi_S\right]+\mathrm{Ei}\left[-ik_{-}h_0(\eta_0\mp\eta)\xi_0\right]\right],$$

and $\mathrm{Ei}(ax)=\int\frac{\exp(ax)}{x}dx$ is the integral exponential function.

From the expression (13) for the first component $P_{-1}^{(2)}(\xi,\eta,\varphi)$ of the total acoustic pressure of the difference-frequency wave, it follows that the scattering diagram of this component is determined by the function $1/(\eta_0 \pm \eta)$. This function depends on the coordinate η_0 or, the polar coordinate system, equivalent to the angle of incidence θ_0 of the highfrequency plane waves. The scattering diagram of the first component $P_{-1}^{(2)}(\xi,\eta,\varphi)$ are shown in Fig.5 for angle of incidence of the high-frequency plane waves $\theta_0 = 30^0$ $(k\_h_0 = 5)$.

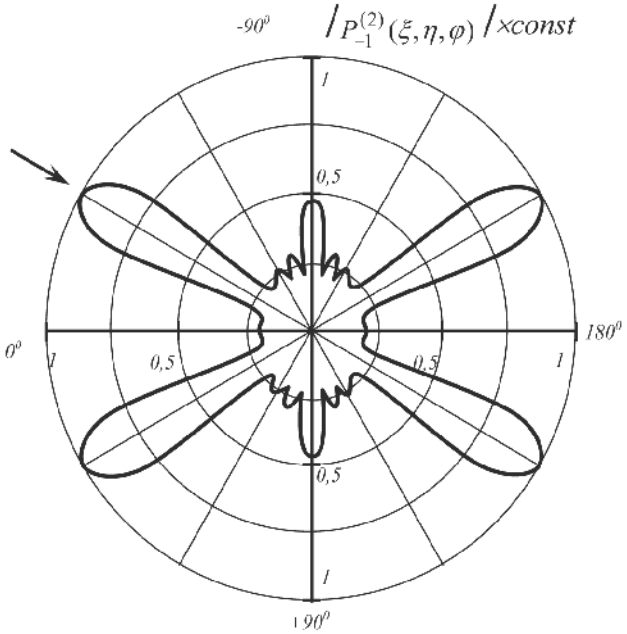


Fig. 5. Scattering diagram of the spatial component $P_{-1}^{(2)}(\xi,\eta,\varphi)$ of the total acoustic pressure produced by the difference-frequency wave by a rigid elongated spheroid for: f_2 =1000 kHz, f_1 =880 kHz, $F\_$ =120 kHz, k_h_0 =5, $\theta_0 = 30^0$, $k_{1,2}h_0 \approx 40, h_0$ =0,01 м, ξ_0 =1,005 (relations axis - 1:10), ξ =7.

In the direction of the angle of incidence (with respect to the *z*-axis), the scattering diagrams have major maximums. Increase of the amplitude of the spheroidal wave produced by the scatterer leads to additional maximums in lateral directions (irrespective of the angle of incidence). This result is connected with the increase of the function $1/\eta$. Increasing the extent of the interaction region (the coordinate ξ_S) results in the narrowing of the scattering lobes; this scenario corresponds to increasing the size of the re-radiating volume around the scatterer.

The elongated spheroid has radial dimension $\xi_0 = 1{,}005$ with the semi-axes correlation 1:10. Acoustic pressure of the difference frequency wave has been calculated in the far field of the scattering spheroid, i.e. in the Fraunhofer region.

Therefore, the scattering field can be considered as being shaped by. Shadowing of the secondary waves sources by the scatterer itself can occur in the Rayleigh region. Here it is necessary to take into account wave dimensions of the scatterer as well as the distance to the point of observation $M(\xi,\eta,\varphi)$. In the cases presented in this contribution, the point of

observation was at radial distances $\xi = 7$ and 15, which exceeded the length of the elongated spheroid by an order magnitude.

Now consider the second $P_{-2}^{(2)}(\xi,\eta,\varphi)$ and third $P_{-3}^{(2)}(\xi,\eta,\varphi)$ components from the equation (10) for the total acoustic pressure of the difference-frequency wave, these components characterise the non-linear interaction of the incident plane waves with the scattered spheroidal ones waves:

$$P_{-2}^{(2)}(\xi,\eta,\varphi)=\frac{C_-}{k_-h_0\eta}\left[\int_{\xi_0}^{\xi_S}\sum_{m=0}^{\infty}\sum_{l\geq m}^{\infty}B_{ml}(k_1h_0)D_{ml}(k_2h_0)\exp\left[i(l\pi/2-m\varphi)\right]\xi'\sin(k_-h_0\xi'\eta)d\xi'-\right.$$

$$\left.-\int_{\xi_0}^{\xi_S}\sum_{m=0}^{\infty}\sum_{l\geq m}^{\infty}B_{ml}(k_1h_0)D_{ml}(k_2h_0)\exp\left[i(l\pi/2-m\varphi)\right]\frac{\sin(k_-h_0\xi'\eta)}{\xi'}d\xi'\right]. \qquad (14)$$

Values of $B_{ml}(k_nh_0)$ and $D_{ml}(k_nh_0)$ are substituted into equation (14) and the plane wave expansion is used. For the axially symmetrical scattering problem (perfect spheroid), the high-frequency asymptotic forms the angular spheroidal 1st- order function $S_{ml}(k_nh_0,\eta)$ and the radial spheroidal 3rd - order function $R_{ml}^{(3)}(k_nh_0,\xi')$ [Kleshchyov & Klyukin, 1987], [Abramovitz & Stegun, 1971]:

$$R_{ml}^{(3)}(k_nh_0,\xi')\underset{k_nh_0\xi'\to\infty}{\approx}\frac{i^{-l-1}}{k_nh_0\xi'}\exp\left[ik_nh_0\xi'\right].$$

Then equation (11) takes the form

$$P_{-2}^{(2)}(\xi,\eta,\varphi)\approx\frac{2iC_-A(k_2h_0)}{k_-k_2h_0^2\eta\sqrt{2(1-\eta)}}\left[\int_{\xi_0}^{\xi_S}\exp\left[-i(k_2h_0-k_1h_0\eta_0)\xi'\right]\sin(k_-h_0\xi'\eta)d\xi'-\right.$$

$$\left.-\int_{\xi_0}^{\xi_S}\exp\left[-i(k_2h_0-k_1h_0\eta_0)\xi'\right]\frac{\sin(k_-h_0\xi'\eta)}{\xi'^2}d\xi'\right]. \qquad (15)$$

After the final integration [Prudnikov et al., 1983], the expression for the 2nd component of the total acoustic pressure of the difference-frequency wave takes the form

$$P_{-2}^{(2)}(\xi,\eta,\varphi)=P_{-21}^{(2)}+P_{-22}^{(2)}+P_{-23}^{(2)}+P_{-24}^{(2)}, \qquad (16)$$

where

$$P_{-21,-22}^{(2)}\approx\mp\frac{iC_-A(k_2h_0)}{2k_-k_2h_0^2\eta\sqrt{(1-\eta_0)(1-\eta)}}\left[\frac{\exp(iu_2\xi_S)-\exp(iu_2\xi_0)}{u_2}\right],$$

$$P_{-23,-24}^{(2)}\approx\mp\frac{C_-A(k_2h_0)}{2k_-k_2h_0^2\eta\sqrt{(1-\eta_0)(1-\eta)}}\left[\frac{\exp(-iu_2\xi_S)}{\xi_S}-\frac{\exp(iu_2\xi_0)}{\xi_0}-u_2\left[\mathrm{Ei}(-iu_2\xi_S)-\mathrm{Ei}(-iu_2\xi_0)\right]\right],$$

$$u_2=(k_2h_0-k_1h_0\eta_0\mp k_-h_0\eta).$$

The expression for the 3rd component $P_{-3}^{(2)}(\xi,\eta,\varphi)$ is similar to the expression (15). An analysis of equation (15) shows that the behaviour of scattering diagrams for the components $P_{-2}^{(2)}(\xi,\eta,\varphi)$ and $P_{-3}^{(2)}(\xi,\eta,\varphi)$ is determined mainly by the function $1/\eta\sqrt{(1-\eta_0)(1-\eta)}$, where the dependence on the angle of incident θ_0 (that is η_0) is not clear. The scattering diagram of these components are shown in Fig.6, for $\theta_0 = 30^0$ ($k\_h_0 = 5$). These diagrams have maximums in the backward and side directions (0^0 and $\pm 90^0$). The increase of the wave size of the spheroidal scatterer leads to additional maximums, which depend on the angle of incident of the high-frequency plane waves.

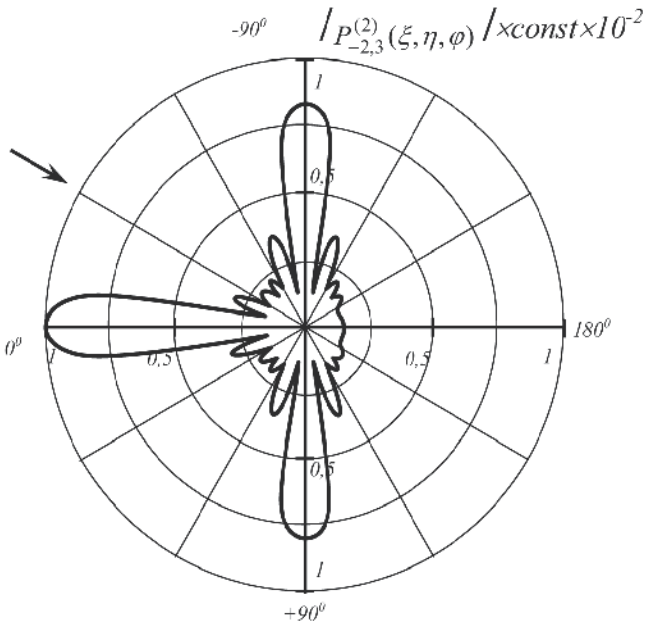


Fig. 6. Scattering diagram of the spatial components $P_{-2}^{(2)}(\xi,\eta,\varphi)$, $P_{-3}^{(2)}(\xi,\eta,\varphi)$ by a rigid elongated spheroid for: f_2 = 1000 kHz, f_1 =880 kHz, $F\_$ =120 kHz, k_h_0 =5, $\theta_0 = 30^0$, ξ_0 =1.005, ξ =7.

Now, we consider the fourth component $P_{-4}^{(2)}(\xi,\eta,\varphi)$ of the total acoustic pressure of the difference-frequency wave. This component characterises the non-linear interaction of the scattered spheroidal waves with frequencies ω_1 and ω_2:

$$P_{-4}^{(2)}(\xi,\eta,\varphi) = \frac{C\_}{k\_h_0\eta}\left[\int_{\xi_0}^{\xi_S}\sum_{m=0}^{\infty}\sum_{l\geq m}^{\infty} D_{ml}(k_1h_0)D_{ml}(k_2h_0)\xi' \sin(k\_h_0\xi'\eta)d\xi' - \right.$$

$$\left. - \int_{\xi_0}^{\xi_S}\sum_{m=0}^{\infty}\sum_{l\geq m}^{\infty} D_{ml}(k_1h_0)D_{ml}(k_2h_0)\frac{\sin(k\_h_0\xi'\eta)}{\xi'}d\xi'\right]. \qquad (17)$$

After some algebraic manipulations, equation (17) takes the form

$$P_{-4}^{(2)}(\xi,\eta,\varphi) = P_{-41}^{(2)} + P_{-42}^{(2)} + P_{-43}^{(2)} + P_{-44}^{(2)}, \tag{18}$$

where

$$P_{-41,-42}^{(2)} \approx \mp \frac{C_{-}A(k_2h_0)A(k_2h_0)}{2ik_{-}k_2k_1h_0^2\eta(1-\eta_0)(1-\eta)}\left[-u_4\left[\mathrm{Ei}(-iu_4\xi_S)-\mathrm{Ei}(-iu_4\xi_0)\right]\right],$$

$$P_{-43,-44}^{(2)} \approx \mp \frac{C_{-}A(k_2h_0)A(k_2h_0)}{4ik_{-}k_2k_1h_0^2\eta(1-\eta_0)(1-\eta)}\left[iu_4\left(\frac{\exp(-iu_4\xi_S)}{\xi_S}-\frac{\exp(iu_4\xi_0)}{\xi_0}\right)+u_4^2\left[\mathrm{Ei}(-iu_4\xi_S)-\mathrm{Ei}(-iu_4\xi_0)\right]\right],$$

$u_4 = (k_{-}h_0 \mp k_{-}h_0\eta)$.

The scattering diagram of the fourth component $P_{-4}^{(2)}(\xi,\eta,\varphi)$ are shown in Fig.7, for $\theta_0 = 30^0$ $(k_{-}h_0 = 5)$. Their configuration is primarily determined by the function $1/\eta(1-\eta_0)(1-\eta)$ of equation (18). As indicated above, this function has a maximum in the backward direction and slightly depends on the angle of incidence. Increasing of the spheroidal scatterer wave size results increases lateral scattering.

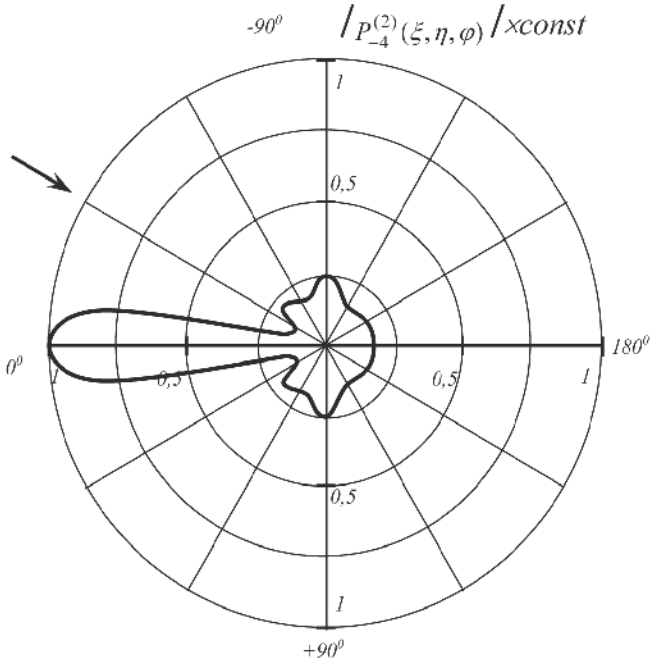


Fig. 7. Scattering diagram of the spatial component $P_{-4}^{(2)}(\xi,\eta,\varphi)$ by a rigid elongated spheroid for: f_2 = 1000 kHz, f_1 =880 kHz, F_{-} =120 kHz, $k_{-}h_0$ =5, $\theta_0 = 30^0$, ξ_0 =1.005, ξ =7.

Fig.8 presents the scattering diagram of the total acoustic pressure in the difference-frequency wave $P_{-}^{(2)}(\xi,\eta,\varphi)$ according to the asymptotic expressions for spatial components. In this case, the angle of incidence is $\theta_0 = 30^0$ $(k_{-}h_0 = 5)$, and the coordinate $\xi = 7$.

Fig.9 shows wave scattering diagrams of difference frequency $P_{-}^{(2)}(\xi,\eta,\varphi)$ on rigid elongated spheroid ξ_0=1,005 with different incidence angle values of inflation incident waves $\theta_0 = 0^0$; 90^0.

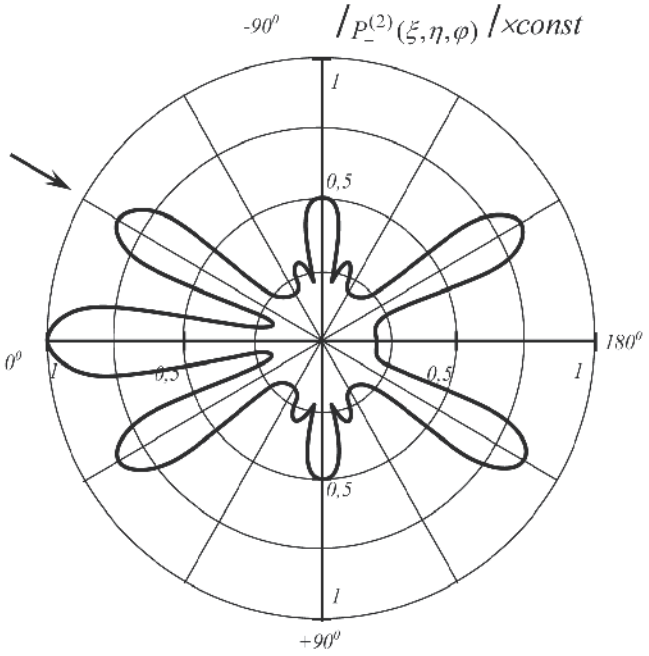


Fig. 8. Scattering diagram of the total acoustic pressure the difference-frequency wave $P_{-}^{(2)}(\xi,\eta,\varphi)$ by a rigid elongated spheroid for: f_2 = 1000 kHz, f_1 =880 kHz, F_{-} =120 kHz, k_h_0 =5, $\theta_0 = 30^0$, ξ_0 =1,005, ξ =7.

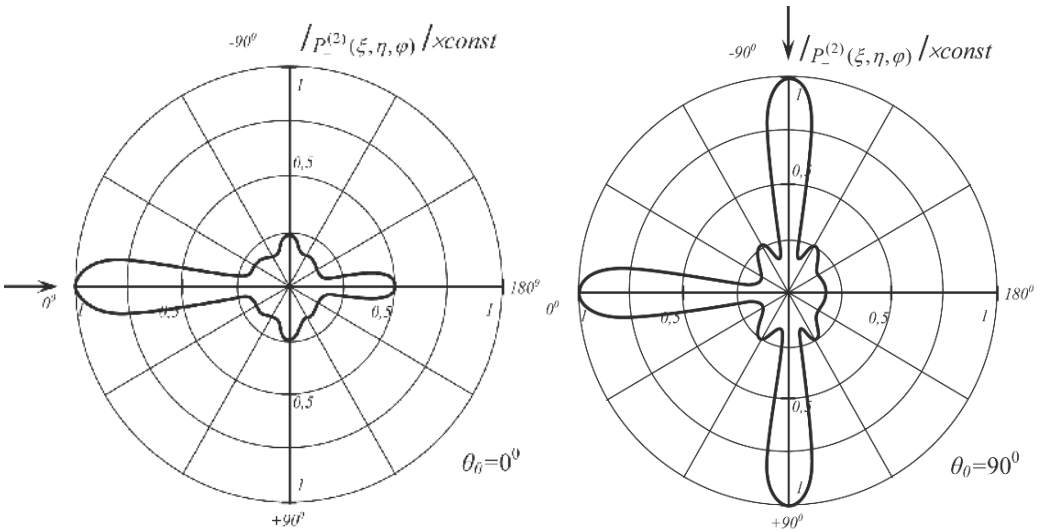


Fig. 9. Scattering diagrams of the total acoustic pressure the difference-frequency wave $P_{-}^{(2)}(\xi,\eta,\varphi)$ by a rigid elongated spheroid for: f_2 = 1000 kHz, f_1 =880 kHz, F_{-} =120 kHz, k_h_0 =5, ξ_0 =1,005, ξ =7, $\theta_0 = 0^0$; 90^0.

With incidence angle $\theta_0 = 0^0$ diagrams have got the basic maximums back, with the increase of spheroid wave dimension, the modest lateral scattering appears. With incidence angle $\theta_0 = 60^0$ diagrams are of the similar form $\theta_0 = 30^0$, with conformable maximums in decrease direction, in mirrorlike, as well as back.

With incidence angle $\theta_0 = 90^0$ diagrams have got the basic maximums back and lateral directions. With the wave dimension growth, modest intermediate levels can be observed. It follows from Fig.9 that angle value change θ_0 leads generally to the change of maximums position in the line of incidence and reflex angle.

It is emphasized that the figures illustrate the dependence of acoustic pressure $P_{-}^{(2)}(\xi,\eta,\varphi)$ on the polar angle $\theta = \arccos\eta$ but not on the angle of asymptote of the hyperbola η. This presentation is conventionally employed for the scattering diagrams in spheroidal coordinates [Cpence & Ganger, 1951], [Kleshchyov & Sheiba, 1970].

The diagrams are presented in the *xoz* plane (Fig.4). Polar angle θ varies in the range 0^0 to 360^0; the value of the angle $\theta = 0^0$ corresponds to the position of x axis, and the value $\theta = 90^0$ corresponds to *z* axis. The arrow here shows the direction of the initial plane wave incidence. The axisymmetry of the diagrams with respect to x axis has been taken into account and two diagrams with positive and negative directions of the angle $\theta = \pm 180^0$ have been combined.

Fig.10 shows a spatial simulation of the scattering diagram of the total acoustic pressure $P_{-}^{(2)}(\xi,\eta,\varphi)$ for $\theta_0 = 30^0$ ($k\_h_0 = 5$, $\xi = 7$, an arrow indicates the direction of the initial wave incidence). It is a surface of revolution, and the rotation axis is the larger axis of the elongated spheroid, that is the *x*- axis.

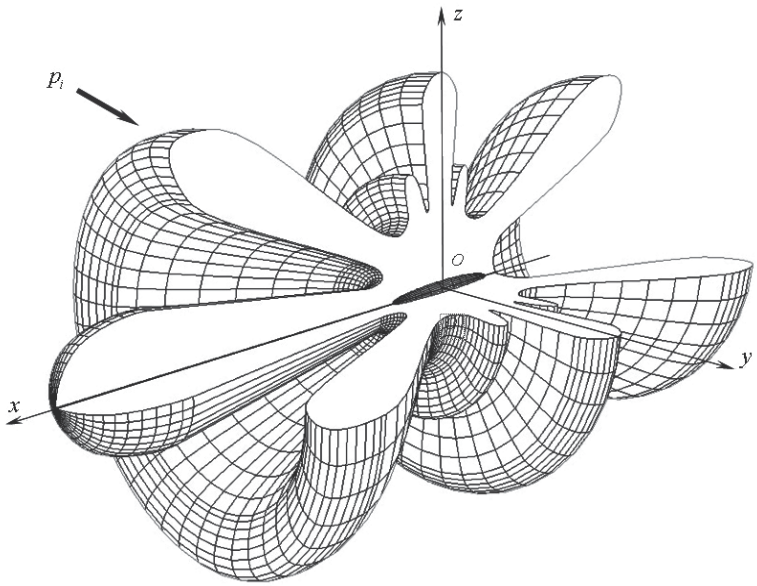


Fig. 10. Spatial model of scattering diagram of the total acoustic pressure the difference-frequency wave $P_{-}^{(2)}(\xi,\eta,\varphi)$ by a rigid elongated spheroid for: f_1 =880 kHz, $F\_$ =120 kHz, k_h_0 =5, $\theta_0 = 30^0$, ξ =7.

5. Discussion

Although investigation of the linear scattering of acoustic waves by the elongated spheroid has been considered previously, results of the scattering of the nonlinearly interacting acoustic wave were not reported. In most previous publications, the problem is investigated when the angles of incidence of acoustic waves are $\theta = 0^0$ and 90^0 [Kleshchyov & Sheiba, 1970], [Tetyuchin & Fedoryuk, 1989].

In article [Kleshchyov & Sheiba, 1970] the calculated diagrams of plane acoustic wave scattering by a similar size spheroid ($\xi_0 = 1{,}005$, $kh_0 = 10$) at angle of incidence $\theta = 30^0$ are presented. Also in this work the scattering diagram has maximums symmetrical to the angle of incidence (mirror lobes) with respect to z axis [Burke, 1966], [Boiko, 1983]. At angle of incidence $\theta = 0^0$ forward scattering dominates. The basic maximum is aligned with 140^0. When the angle of incidence is $\theta = 90^0$ (lateral incidence), there are only two maximums – forward and backward.

An analysis of the acoustic pressure distribution of the difference-frequency wave scattered field shows that the scattering diagrams have maximums in a backward direction. In direction to the angle of incidence, in lateral and transverse directions, plane waves have maximums. Incident high-frequency plane waves form the scattering field in backward and forward directions, and scattered spheroidal waves form the scattering field in transverse direction. An increase in the wave size of the spheroidal scatterer changes maximum levels, and an increase in the size of the interacting area around the elongated spheroidal scatterer leads to narrowing of these maximums.

It is important to note that in this work we considered the case when the scattered field is generated by the secondary wave sources located in the volume around the spheroid. In the case of the linear scattering, these sources are located on the surface of the spheroid. The mirror maximums 30^0 and 150^0 appear as a result of the asymptotics of the first spatial sum $P_{-1}^{(2)}(\xi, \eta, \varphi)$ as confirmed in [2]. Therefore, the plotted scattering diagrams are in conformity with the results of 90^0 [Burke, 1966], [Kleshchyov & Sheiba, 1970], [Boiko, 1983], [Tetyuchin & Fedoryuk, 1989].

As for the numerical evaluation of the acoustic pressure, it is necessary to note the following. In view of the complexity of mathematical calculations, the obtained asymptotics allow for qualitative evaluation of the spatial distribution of the acoustic pressure in the scattered field. It would be more adequate to compare the results with experimental data. Unfortunately, experiments in non-linear conditions have not been carried out. For the sake of better understanding of contribution of the separated sums into the cumulative acoustic field, results were presented for two values of the wave dimension and the angle of incidence.

It should be noted, that description of wave processes in spheroidal coordinates have several peculiarities. For example, comparing the acoustic pressure distribution at the distance from the scatterer, the results given in [Abbasov & Zagrai, 1994], [Abbasov & Zagrai, 1998], [Abbasov, 2007] can be taken. Spheroidal coordinates in a far field transform into spherical ones $(h_0 \to 0)$ and $P_{-}^{(2)}(\xi, \eta, \varphi) \to P_{-}^{(2)}(r, \theta, \varphi)$. The results of this research are in agreement with results of prior studies of the scattering process described in spherical coordinates.

6. Conclusion

Summing up the secondary field studies on the difference frequency wave with interacting acoustic wave scattering on elongated spheroid, it should be noted that:

- the statement of the problem has been formulated and problem peculiarities of scattering in elongated spheroidal coordinates has been described, the solution of non-homogeneous wave equation in the second approximation and Helmholtz nonhomogeneous equation on the difference frequency wave has been obtained;
- high-frequency asymptotic expressions of general acoustic pressure of difference frequency wave have been obtained; they consist of spacing terms, characterizing nonlinear interaction between incident plane and scattered spheroidal waves;
- the assumption diagrams of difference frequency wave scattering on different distances from spheroidal scatterer, for different incident angles and different wave dimensions: $k\_h_0 = 0{,}5 \div 1$, incident angles $\theta_0 = 0^0$, 30^0, 60^0, 90^0, radial distances ξ =3; 7; 15, have been obtained;
- the obtained diagrams of difference frequency wave scattering have basic maximums in back, lateral directions and in the incidence and reflex line (reflection lobe) of inflation waves, three-dimensional diagram models of difference frequency wave scattering on elongated spheroid have been featured.

The method of successive approximations has been used for the description of wave processes with weak non-linearity. The diagrams are presented that illustrate the distribution of acoustic pressure of the scattered field. In view of the obtained theoretical results, the method of successive approximations is an adequate tool for solving the problem of the scattering of non-linearly interacting waves by an elongated spheroid.

7. References

Abbasov, I.B. (2007). *Scattering nonlinear interacting acoustic waves: sphere, cylinder and a spheroid*. Fizmatlit, Moscow, 160p.

Abbasov, I.B., Zagrai, N.P. (1994). Scattering of interacting plane waves by a sphere. *Acoust. Phys.* Vol. 40, No. 4, P. 473-479.

Abbasov, I.B., Zagrai, N.P. (1998). The investigation of the second field of the summarized frequency originated from scattering of nonlinearly interacting sound waves at a rigid sphere. *Journal of Sound and Vibration.* Vol. 216, No. 1, P. 194-197.

Abramovitz, M., Stegun, I. (1971). *Handbook of special functions with formulas, graphs, and mathematical tables*. Dover, New York, 830p.

Andebura, V. A. (1969). Acoustic properties of spheroidal radiators. *Akust. Zh.* Vol. 15, No.44, P. 513-522.

Andebura, V. A. (1976). About akustic - mechanical characteristics spheroidal a radiator and scatterer. *Akust. Zh.* Vol. 22, No. 4, P. 481-486.

Babailov, E.P., Kanevsky V.A. (1988). Sound scattering the gas-filled spheroidal bubble of fishes. *Akust. Zh.* Vol. 34, No. 1, P. 19-23.

Belkovich, V. M, Grigoriev, V. A, Katsnelson, B.G., Petnikov, V.G. (2002). About possibilities of use of acoustic diffraction in monitoring problems cetaceans. *Akust. Zh.* Vol. 48, No. 2, P. 162-166.

Boiko, A.I. (1983). The scattering of plane sound wave from thin revolve body. *Akust. Zh.* Vol. 29, No. 3, P. 321-325.

Burke, J.E. (1966). Long-wavelength scattering by hard spheroids. *Journ. Acoust. Soc. Amer*. Vol. 39, No. 5, P. 826-831.

Chertock, G. (1961). Sound radiation from circular pistons of elliptical profile. *Journ. Acoust. Soc. Amer.* Vol. 33, No. 7, P. 871-8876.

Corn, H., Corn, T. (1961). *Mathematical Handbook*. cMgraw-Hill Book Company, New York. 720p.

Cpence, R., Ganger, S. (1951). The scattering of sound from a prolate spheroid. *Journ. Acoust. Soc. Amer*. Vol. 23, No. 6, P. 701-706.

Fedoryuk, M.V. (1981). The scattering of sound wave from thin acoustically rigid revolve body. *Akust. Zh.* Vol. 27, No. 4, P. 605-609.

Guyer, R.A., Johnson, P.A. (1999). Nonlinear mesoscopic elasticity evidence for a new class of materials. *Physics Today*. Vol. 52, No. 4, P. 30-36.

Haslett, R. (1962). Determination of the acoustic scatter patterns and cross sections of fish models and ellipsoids. *Brit. Journ. Appl. Phys.* Vol. 13, No. 12, P. 611-620.

Kleshchyov, A.A, Sheiba, L.S. (1970). The scattering of sound wave from an ideal elongated spheroids. *Acoust. Phys.(Akust. Zh.)* Vol. 16, No.2, P. 264-268.

Kleshchyov, A.A. (1992). *Hydroacoustic scatterers*. Sudostroenie. St. Peterburg. 248p.

Kleshchyov, A.A. (2004). Physical model of sound scattering by jamb of fishes who is at border of section of. *Akust. Zh.* Vol. 50, No. 4, P. 512-515.

Kleshchyov, A.A., Clyukin, I.I. (1987). *The foundation of hydroacoustic*. Sudostroenie, Leningrad. 224p.

Kleshchyov, A.A., Rostovtsev, D.M. (1986). Sound scattering elastic and liquid ellipsoidal rotation shells. *Akust. Zh.* Vol. 32, No. 5, P. 691-694.

Kuzkin, V.M. (2003). Scattering of sound waves on a body in плоскослоистом a wave guide. *Akust. Zh.* Vol. 49, No. 1, P. 77-84.

Lebedev, A.V., Ostrovsky,i L.A., Sutin A.M. (2005). Nonlinear acoustic spectroscopy of local defects in geomaterials. *Akust. Zh.* Vol. 51, No. add., P. 103-117.

Lebedev, A.V., Salin, B.M. (1997). An experimental method of definition of dispersion section of the elongated bodies. *Akust. Zh.* Vol. 43, No. 3, P. 376-385.

Lyamshev, L.M., Sakov, P.V. (1992). Nonlinear scattering of sound from an pulsted sphere. *Soviet Physics Acoustics,* Vol. 38, No. 1, P. 51-57.

Novikov, B.K., Rudenko, O.V., Timoshenko, V.I. (1987). *Nonlinear underwater acoustic*. Acoustical Society of America, New York, 264 p.

Prudnikov, A.P., Brychkov, Yu. A., Marichhev, O.I. (1983). *Integrals and rows*. Nauka. Moscow. 752p.

Skudrzyk, E. (1971). *The foundations of acoustics*. Springer, New York, 542p.

Stanton, T.K. (1989). Simple approximate formulas for backscattering of sound by spherical and elongated objects. *Journ. Acoust. Soc. Amer*. Vol. 86, No. 4, P. 1499-1510.

Tetyuchin, M.Yu., Fedoryuk, M.V. (1989). The diffraction of plane sound wave from a elongated rigid revolved body in the liquid. *Akust. Zh.* Vol. 35, No. 1, P. 126-130.

Tikhonov, A.N., Samarskyi, A.A. (1966). *The equations of mathematical physics*. Nauka, Moscow. 724p.

Weksler, N.D., Dubious, B., Lave, A. (1999). The scattering of acoustic wave from an ellipsoidal shell. *Akust. Zh.* Vol. 45, No. 1, P. 53-58.

Werby, M.F., Green, L.H. (1987). Correspondence between acoustical scattering from spherical and end-on incidence spherical shells. *Journ. Acoust. Soc. Amer.* Vol. 81, No. 2, P. 783-787.

6

Acoustic Waves in Phononic Crystal Plates

Xin-Ye Zou, Xue-Feng Zhu, Bin Liang and Jian-Chun Cheng
Nanjing University,
People's Republic of China

1. Introduction

Recently, the study on elastic waves in phononic crystal plates is becoming a research hotspot due to its potential applications, especially in wireless communication, transducer and sensor system [1-10]. The phononic crystal plates commonly consist of two materials with large contrast in elastic properties and densities, arranging in a periodic (or quasiperiodic) array. The absolute band gaps in composite plates can forbid the propagation of all elastic wave modes in all directions. Comparing with the bulk wave and surface acoustic wave devices, phononic crystal plates have better performance in elastic wave propagation since the phase speed of most Lamb wave modes (except for A_0 mode) is faster than surface wave mode, and also the wave energy in plates is totally confined between the upper and nether free-stress boundaries regardless of the air damp and self-dissipation, which provides a special potentiality in micro-electronics in wireless communication.

The propagation of Lamb waves is much more complicated than bulk wave and surface acoustic wave in terms of the free-stress boundaries which can couple the longitudinal and transversal strain components. The first attempt to describe the propagation of Lamb waves with wavelength comparable with the lattice is due to Auld and co-workers [11-12], who studied 2D composites within the couple-mode approximation. Alippi et al. [13] have presented an experimental study on the stopband phenomenon of lowest-order Lamb waves in piezoelectric periodic composite plates and interpreted their results in terms of a theoretical model, which provides approximate dispersion curves of the lowest Lamb waves in the frequency range below the first thickness mode by assuming no coupling between different Lamb modes. The transmissivity of the finite structure to Lamb wave modes was also calculated by taking into account the effective plate velocities of the two constituent materials [14]. Based on a rigorous theory of elastic wave, Chen et al.[1] have employed plane wave expansion (PWE) method and transient response analysis (TRA) to demonstrate the existence of stop bands for lower-order Lamb wave modes in 1D plate. Gao et al.[8] have developed a virtual plane wave expansion (V-PWE) method to study the substrate effect on the band gaps of lower-order Lamb waves propagating in thin plate with 1D phononic crystal coated on uniform substrate. They also studied the quasiperiodic (Fibonacci system) 1D system and find out the existence of split in phonon band gap [2]. In order to reduce the computational complexity without losing the accuracy, Zhu et al.[9] have promoted an efficient method named harmony response analysis (HRA) and supercell plane wave expansion (SC PWE) to study the behavior of Lamb wave in silicon-based 1D composite plates. Zou et al.[10] have employed V-PWE method to study the band gaps of plate-mode waves in 1D piezoelectric composite plates with substrates.

The chapter is structured as follows: we firstly introduce the theory and modeling used in this chapter in Section 2. In Section 3, we focus on the band gaps of lower-order Lamb waves in 1D composite thin plates without/with substrate. In Section 4, we study the lamb waves in 1D quasiperiodic composite thin plates. In Section 5, we focus on acoustic wave behavior in silicon-based 1D phononic crystal plates for different structures, and finally in Section 6, we study the band gaps of plate-mode waves in 1D piezoelectric composite plates without/with substrates.

2. Theory and modeling of phononic crystal plates

In this section, we give the theory and modeling of phononic crystal plates with different structures: the periodic structure without/with substrate, and the quasiperiodic structure.

2.1 Periodic structure without substrate by PWE method

As shown in Fig. 1, the periodic composite plate consists of material A with width d_{A}, material B with d_{B}, lattice spacing $D = d_{\mathrm{A}} + d_{\mathrm{B}}$, and filling rate defined by $f = d_{\mathrm{A}} / D$. The wave propagates along the x direction of a plate bounded by planes $z = 0$ and $z = L$.

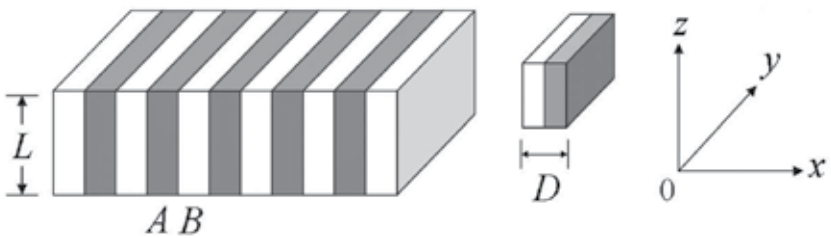


Fig. 1. 1D periodic composite plate consisting of alternate A and B strips.

In the periodic structure, all field components are assumed to be independent of the y direction. In an inhomogeneous linear elastic medium with no body force, the equation of motion for displacement vector $\boldsymbol{u}(x,z,t)$ can be written as

$$\rho(x)\ddot{\boldsymbol{u}}_p = \partial_q[c_{pqmn}(x)\partial_n \boldsymbol{u}_m], \quad (p = 1,2,3), \tag{1}$$

where $\rho(x)$ and $c_{pqmn}(x)$ are the x-dependent mass density and elastic stiffness tensor, respectively. Due to the spatial periodicity in the x direction, the material constants, $\rho(x)$ and $c_{pqmn}(x)$ can be expanded in the Fourier series with respect to the 1D reciprocal lattice vectors (RLVs), as follows

$$\rho(x) = \sum_G e^{jGx}\rho_{\mathrm{G}}, \tag{2}$$

$$c_{pqmn}(x) = \sum_G e^{jGx}c_{pqmn}^G, \tag{3}$$

where ρ_G and c_{pqmn}^G are expansion coefficients of the mass density and elastic stiffness tensor, respectively. From the Bloch theorem and by expanding the displacement vector $\boldsymbol{u}(x,z,t)$ into Fourier series, one obtains

$$\boldsymbol{u}(x,z,t)=\sum_{G} e^{jk_x x-j\omega t}(e^{jGx}\boldsymbol{A}_G e^{jk_z z}), \tag{4}$$

where k_x is a Bloch wave vector and ω is the circular frequency, $\boldsymbol{A}_G=(A_G^1,A_G^2,A_G^3)$ is the amplitude vector of the partial waves, and k_z is the wave number of the partial waves along the z direction. Substituting Eqs. (2)-(4) into Eq. (1), one obtains homogenous linear equations to determine both (A_G^1,A_G^2,A_G^3) and k_z.

$$\begin{pmatrix} c_{11}(k_x+G)(k_x+G')+c_{44}k_z^2-\rho\omega^2 & 0 & c_{12}(k_x+G)+c_{44}(k_x+G')k_z \\ 0 & c_{44}(k_x+G)(k_x+G')+c_{44}k_z^2-\rho\omega^2 & 0 \\ c_{12}(k_x+G')+c_{44}(k_x+G)k_z & 0 & c_{44}(k_x+G)(k_x+G')+c_{11}k_z^2-\rho\omega^2 \end{pmatrix}\begin{pmatrix} A_{G'}^1 \\ A_{G'}^2 \\ A_{G'}^3 \end{pmatrix}=0, \tag{5}$$

Supposing that the materials *A* and *B* are cubic materials, it is obvious that the wave motion polarized in the *y*-direction, namely SH wave, decouples to the wave motions polarized in the *x*- and *z*-directions, namely, P and SV waves. It is relatively simple to discuss the SH wave so that we focus our attentions to P and SV waves, and the equation of motion for Lamb waves becomes

$$\begin{pmatrix} c_{11}(k_x+G)(k_x+G')+c_{44}k_z^2-\rho\omega^2 & c_{12}(k_x+G)+c_{44}(k_x+G')k_z \\ c_{12}(k_x+G')+c_{44}(k_x+G)k_z & c_{44}(k_x+G)(k_x+G')+c_{11}k_z^2-\rho\omega^2 \end{pmatrix}\begin{pmatrix} A_{G'}^1 \\ A_{G'}^3 \end{pmatrix}=0, \tag{6}$$

If one truncates the expansions of Eqs. (2) and (3) by choosing n RLVs, one will obtain $4n$ eigenvalues $k_z^{(l)}$, $(l=1-4n)$. For the Lamb waves, all of the $4n$ eigenvalues $k_z^{(l)}$ must be included. Accordingly, displacement vector of the Lamb waves can be taken of the form

$$\boldsymbol{u}(x,z,t)=\sum_{G}{}' e^{i(k_x+G)x-i\omega t}\left(\sum_{l=1}^{4n}\boldsymbol{A}_G e^{iK_z^{(l)}z}\right)=\sum_{G}{}' e^{i(k_x+G)x-i\omega t}\left(\sum_{l=1}^{4n}X_l\varepsilon_G^{(l)}e^{ik_z^{(l)}z}\right), \tag{7}$$

where $\varepsilon_G^{(l)}$ is the associated eigenvector for the eigenvalue $k_z^{(l)}$, X_l is the weighting coefficient to be determined, and the prime of the summation expresses that the sum over G is truncated up to n.

The boundary conditions are the stress-free on the upper (z = 0) and rear (z = L) surfaces

$$\boldsymbol{T}_{p3}\big|_{z=0,L}=c_{p3mn}\partial_n u_m\big|_{z=0,L}=0 \qquad (p=1,3). \tag{8}$$

which $\boldsymbol{T}_{p3}$ is the stress tensor and L is the plate thickness. Eq. (8) leads to $4n$ homogeneous linear equations for X_l l = (1- $4n$), as follows

$$\begin{bmatrix} H_{1,G}^{(1)} & H_{1,G}^{(2)} & \cdots & H_{1,G}^{(4n)} \\ H_{2,G}^{(1)} & H_{2,G}^{(2)} & \cdots & H_{2,G}^{(4n)} \\ H_{3,G}^{(1)} & H_{3,G}^{(2)} & \cdots & H_{3,G}^{(4n)} \\ H_{4,G}^{(1)} & H_{4,G}^{(2)} & \cdots & H_{4,G}^{(4n)} \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{4n} \end{bmatrix}=\tilde{H}X=0, \tag{9}$$

where $\tilde{H}$ is a $4n \times 4n$ matrix with components

$$H_{1,G}^{(l)} = C_{G-G'}^{44}[(k_x + G')\varepsilon_{G'}^{3(l)} + k_z^{(l)}\varepsilon_{G'}^{1(l)}], \tag{10a}$$

$$H_{2,G}^{(l)} = C_{G-G'}^{11}k_z^{(l)}\varepsilon_{G'}^{3(l)} + C_{G-G'}^{12}(k_x + G')\varepsilon_{G'}^{1(l)}, \tag{10b}$$

$$H_{3,G}^{(l)} = C_{G-G'}^{44}[(k_x + G')\varepsilon_{G'}^{3(l)} + k_z^{(l)}\varepsilon_{G'}^{1(l)}] \times \exp(jk_z^{(l)}L), \tag{10c}$$

$$H_{4,G}^{(l)} = [C_{G-G'}^{11}k_z^{(l)}\varepsilon_{G'}^{3(l)} + C_{G-G'}^{12}(k_x + G')\varepsilon_{G'}^{1(l)}] \times \exp(jk_z^{(l)}L). \tag{10d}$$

From Eq. (9) one notes that to obtain nontrivial solution for the X_l, the determinant of the boundary condition matrix should be equal to zero. The ω of the Lamb wave modes are thus found by searching for the values of ω that simultaneously make the Eq. (6) and $\det(\tilde{H})$ equal to zero. In practice, an iterative search procedure is usually required to find these ω [15-16].

2.2 Periodic structure with substrate by V-PWE method

As shown in Fig.2, the composite plate with substrate consists of the 1D phononic crystal (PC) layer coated on *C* substrate. The PC layer consists of the material *A* with the width d_A and the material *B* with the width d_B.

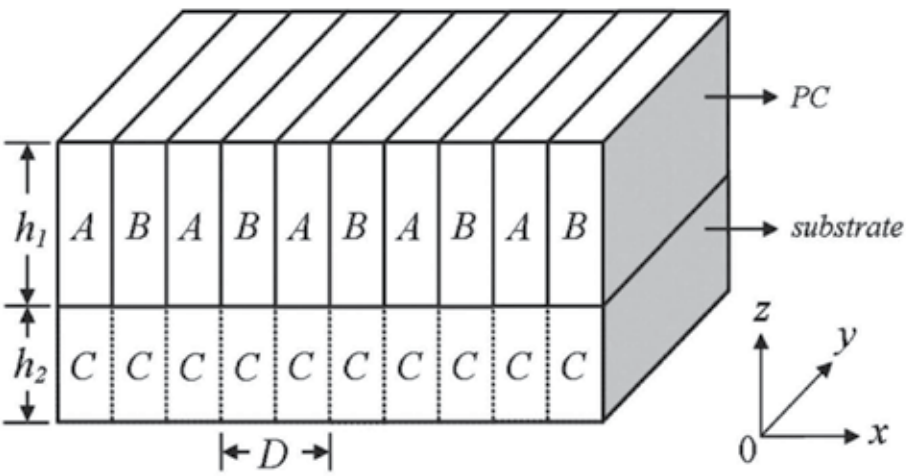


Fig. 2. The 1D periodic composite plate consisting of alternate *A* and *B* strips with a substrate *C*.

We develop a V-PWE method to calculate the dispersion curves of Lamb wave modes propagating along the *x* direction in the presence of the uniform substrate. Here, we give the equations of V-PWE method for the piezoelectric periodic structure with substrate. One can have the equations for non-piezoelectric situation by omitting the piezoelectricity components and the electrical boundary conditions.

In the situation of piezoelectric composite plate with substrate, the equations governing the motion of lattice displacement $\boldsymbol{u}^m(x,z,t)$ and electrical displacement $\boldsymbol{D}^m(x,z,t)$ in this inhomogeneous system are given by

$$\rho^m(x)\ddot{u}_j^m = \partial_i T_{ij}^m, \tag{11}$$

$$\partial_i D_i^m = 0, \tag{12}$$

$$T_{ij}^m = c_{ijkl}^m(x)\partial_l u_k^m + e_{lij}^m(x)\partial_l \phi^m, \tag{13}$$

$$D_i^m = e_{ikl}^m(x)\partial_l u_k^m - \varepsilon_{il}^m(x)\partial_l \phi^m, \tag{14}$$

where $i,j,k,l = x,z$; $m = 1,2$ (1 represents phononic layer; 2 represents the substrate, respectively). $\boldsymbol{T}_{ij}^m(x,z,t)$, $\boldsymbol{D}^m(x,z,t)$, $\boldsymbol{u}^m(x,z,t)$, $\varphi^m(x,z,t)$, $\rho^m(x)$, $c_{ijkl}^m(x)$, $e_{lij}^m(x)$, and $\varepsilon_{il}^m(x)$ are the stress vector, electrical displacement vector, displacement vector, electric potential, x-dependent mass density, elastic stiffness, piezoelectric, and dielectric constant tensors, respectively. It comes into notice that in fact the material constants depend on the z-direction due to the existence of the substrate, as follows

$$\alpha(x,z) = \begin{cases} \alpha^1(x), & (0 < z < h_1) \\ \alpha^2, & (-h_2 < z < 0) \end{cases} \tag{15}$$

where $\alpha = (\rho, c_{ijkl}, e_{lij}, \varepsilon_{il})$, $(\rho^2, c_{ijkl}^2, e_{lij}^2, \varepsilon_{il}^2)$ are the material constants for the substrate.

Due to the spatial periodicity, the Bloch theorem can be applied to the PC layer, but it cannot be simply applied to the substrate layer. However, one notice that the triangle basic function set in the Fourier series is an orthogonal and complete set, each components in the Fourier series must satisfied the boundary conditions at the interface between the PC layer and the substrate at z = 0, namely the continuities of the normal stress, normal displacement, normal electrical displacement and electric potential.

$$T_{iz}^1\Big|_{z=0} = T_{iz}^2\Big|_{z=0},\ u_{iz}^1\Big|_{z=0} = u_{iz}^2\Big|_{z=0},\ D_z^1\Big|_{z=0} = D_z^2\Big|_{z=0},\ \phi^1\Big|_{z=0} = \phi^2\Big|_{z=0},\ (i = x,z). \tag{16}$$

Therefore, the displacement and electric potential fields in the substrate layer also must be expanded to the Fourier series with the period that is same as the PC layer in order to satisfy the boundary conditions. Then the substrate layer can be treated as a virtual periodic structure that has the same filling fraction and period as the PC layer. Thereupon, the Bloch theorem can be applied to both the PC and the substrate layers.

Due to the spatial periodicity in the x direction, the material constants can be expanded in Fourier series with respect to the 1-D reciprocal-lattice vector (RLV) G, as follows:

$$\alpha(x) = \sum_G e^{jGx}\alpha_G, \tag{17}$$

where α_G is the corresponding Fourier coefficient. Utilizing the Bloch theorem and expanding the displacement vector and electric potential into Fourier series in the PC and the substrate layers, one obtains

$$\boldsymbol{u}^m(x,z,t)=\sum_G e^{j(k_x x-\omega t)}(e^{jGx}\mathrm{A}_G^m e^{jk_z^m z}), \tag{18}$$

$$\boldsymbol{\phi}^m(x,z,t)=\sum_G e^{j(k_x x-\omega t)}(e^{jGx}A_{G3}^m e^{jk_z^m z}), \tag{19}$$

where k_x is a Bloch wave vector, ω is the circular frequency, and k_z^m is the wave number along the z-direction, $\boldsymbol{A}_G^m=(A_{G1}^m,A_{G2}^m)$ and A_{G3}^m are the amplitude vectors of the partial waves and electric potential, respectively. Substituting Eqs. (17)-(19) into Eqs. (11)-(14), one can obtain the eigenvalue problem with respect to k_z^m:

$$(\boldsymbol{A}^m k_z^{m2}+\boldsymbol{B}^m k_z^m+\boldsymbol{C}^m)\cdot \mathrm{U}^m=0, \tag{20}$$

where $\boldsymbol{U}^m=\{A_{G1}^m,A_{G2}^m,A_{G3}^m\}^T$ is called the generalized displacement vector, the $3n\times 3n$ matrices $\boldsymbol{A}^m$, $\boldsymbol{B}^m$, and $\boldsymbol{C}^m$ are functions of k_x, G, ω, ρ_G^m, c_G^{ijklm}, e_G^{lijm}, ε_G^{ilm}, and n is the number of RLV.

Here, we consider the stress-free boundary conditions and two kinds of the electrical boundary conditions. For the 1-D problem, we have

the stress free boundary conditions:

$$T_{iz}^1\Big|_{z=h_1}=0,\quad T_{iz}^2\Big|_{z=-h_2}=0,\quad (i=x,z), \tag{21}$$

the OC boundary conditions:

$$\begin{aligned} &D_z^1\Big|_{z=h_1}=D_z^{air}\Big|_{z=h_1},\quad \phi^1\Big|_{z=h_1}=\phi^{air}\Big|_{z=h_1},\\ &D_z^2\Big|_{z=-h_2}=D_z^{air}\Big|_{z=-h_2},\ \phi^2\Big|_{z=-h_2}=\phi^{air}\Big|_{z=-h_2},\\ &(D_z^{air}=-\varepsilon_0\frac{\partial\phi^{air}}{\partial z},\quad \varepsilon_0=1\times10^{-11}\mathrm{F/m},) \end{aligned} \tag{22}$$

the SC boundary conditions:

$$\phi^1\Big|_{z=h_1}=0,\qquad \phi^2\Big|_{z=-h_2}=0. \tag{23}$$

Putting $A_G^{j(l)m}=X_l^m\beta_G^{j(l)m}$ $(j=1-3,\ l=1-6n,\ m=1,2)$, where $\beta_G^{j(l)m}$ is the associated eigenvector of the eigenvalue $k_z^{(l)m}$, and X_l^m is the weighting coefficient that can be determined from the boundary conditions for different layers, one obtains: $\mathbf{H}\cdot\mathbf{X}=0$ from the equations (16), (21) and (22) [or (23)], where $\mathbf{H}$ is a $12n\times12n$ matrix. The existence of a nontrivial solution of X_l^m needs the determinant of matrix $\mathbf{H}$ to be zero

$$\det(\mathbf{H})=0. \tag{24}$$

Then one can obtain the dispersion relations of the Lamb waves propagating in a 1-D PC layer coated on a substrate.

2.3 Periodic structure without/with substrate by FE method

In order to study the elastic wave in the phononic crystal plates, transient response analysis (TRA) and the harmony response analysis (HRA) are presented here by finite element (FE) method.

First, the TRA is employed to calculate the transmitted power spectra (TPS) for the finite periodic structure. The FE solution involves the discretization of the domain into a number of elements, approximating the displacement values interior to the elements in term of its nodal value through the shape functions of the chosen element and the determination of nodal values [17].

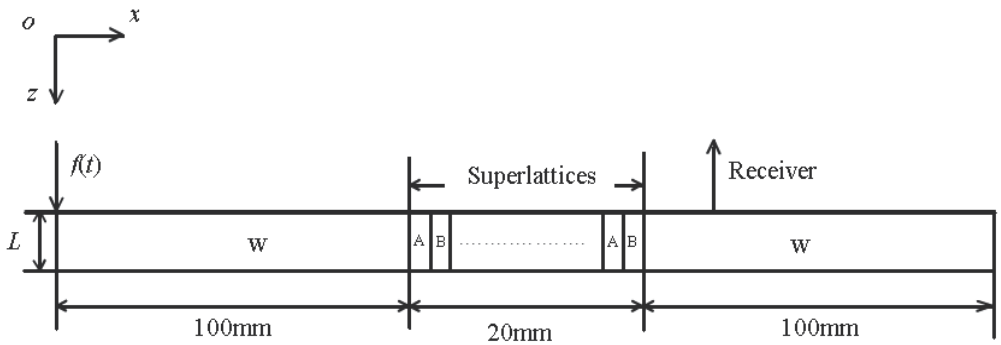


Fig. 3. Modified plate geometry in the Finite Element calculations

Fig. 3 shows the configuration of the modified composite plate in the Finite Element calculations, in which the superlattices with ten periods (the length is 20 mm) is bounded by two pure tungsten plates (the length is 100 mm) at two sides. Lamb waves are excited by the force function $f(t)$ that is a triangle wave at $x = 0$, and are received at $x = 140$ mm [18]. The generation source is far from the periodic structure in order to obtain approximate plane waves when the wavefronts reach it. The step sizes of temporal and spatial discretization in the finite model are fine enough for the convergence of the numerical results (increasing the number of elements of the finite element mesh is equivalent to increasing the number of harmonics in PWE method). The vertical displacement of a node at upper surface of the plate behind the superlattices array at $x = 140$ mm is collected as function of time. For a sufficiently large number of these vertical displacement data on the time axis, the displacement fields are Fourier transformed into the frequency domain to yield the TPS.

We also promote an efficient method named HRA to study the propagation and transmission of acoustic waves in 1D phononic crystal plates. Comparing with TRA [1,2,8], HRA is more time-saving due to its direct calculation in frequency domain and more powerful for the acquirability of displacement field under certain frequency load, which can be further employed to the designation of various phononic crystal functionalities such as filters, resonators and waveguides. With this method, we can study the cases both without and with substrates. Taking the gradient of the displacement fields, we can further study the strain distribution in the plates, and it is really a very direct way to understand how the band gaps form in phononic crystal plates by comparing displacement fields under different frequency loads (inside/outside band gaps).

Any continual periodic loads can produce continual periodic response (harmony response) in phononic crystal plates. HRA is a method used to define the stabilized response of linear structures under time-harmonic loads. By calculating the responses (usually displacement fields) under different frequency loads, we can obtain the transmitted power spectra in the detected region. HRA is a linear analysis regardless of any nonlinear characteristics. For multi-element structure, the Newton's second law can be expressed as follows:

$$[\boldsymbol{M}]\{\ddot{\boldsymbol{D}}\}+[\boldsymbol{C}]\{\dot{\boldsymbol{D}}\}+\{\boldsymbol{R}^{int}\}=\{\boldsymbol{R}^{ext}\} \tag{25a}$$

$$\{\boldsymbol{R}^{int}\}=[\boldsymbol{K}]\{\boldsymbol{D}\} \tag{25b}$$

$$\{\boldsymbol{R}^{ext}\}=\{\boldsymbol{F}\}\exp(i\Omega t) \tag{25c}$$

where $[\boldsymbol{M}]$, $[\boldsymbol{C}]$ and $[\boldsymbol{K}]$ are general mass matrix, damping matrix and stiffness matrix, respectively; $\{\boldsymbol{D}\}$ and $\{\boldsymbol{F}\}\exp(i\Omega t)$ are nodal degree of freedom vector and nodal external load vector, respectively. Eq (25a) describes a dynamic balance among inertial force, damping force, inner force $\{\boldsymbol{R}^{int}\}$ and external load force $\{\boldsymbol{R}^{ext}\}$. The forced vibration of the structure will finally come to a stabilized status in which every node moves in harmonic motion with the same frequency (Ω). Further, we can express $\{\boldsymbol{D}\}$ into:

$$\{\boldsymbol{D}\}=\{\bar{\boldsymbol{D}}\}\exp(i\Omega t) \tag{26}$$

where $\{\bar{\boldsymbol{D}}\}$ is the complex nodal degree of freedom vector. By substituting Eq (26) into Eqs (25a)-(25c), we can obtain:

$$\left([\boldsymbol{K}]+i\Omega[\boldsymbol{C}]-\Omega^2[\boldsymbol{M}]\right)\{\bar{\boldsymbol{D}}\}=\{\boldsymbol{F}\} \tag{27}$$

where $\{\bar{\boldsymbol{D}}\}$ can be obtained using Frontal solver. We choose imaginary component of $\{\bar{\boldsymbol{D}}\}$ to build up the stabilized displacement field under different frequency loads. It is necessary to mention that TRA requires much more substeps to obtain the nodal degree of freedom vector at certain detected time for the reason that the time step Δt should obey the following criterion for numerical convergence in Newmark method: [19]

$$\Delta t \le \frac{\Omega_{crit}}{2\pi f_{max}} \tag{28}$$

where f_{max} is the maximum frequency of interest. Ω_{crit} is defined to be:

$$\Omega_{crit}=\left(\frac{\gamma}{2}-\beta\right)^2 \tag{29}$$

where β is chosen to be $(\gamma+1/2)^2/4$ with $\gamma \ge 1/2$ to achieve as large high frequency dissipation as possible. We choose $\beta=0.2756$, $\gamma=0.55$ in the numerical calculation of TRA. In each substep, a very complex iteration is employed, which takes the form:

$$\left[\boldsymbol{K}^{eff}\right]\{\boldsymbol{D}\}_{n+1}=\left\{\boldsymbol{R}^{ext}\right\}_{n+1}+[\boldsymbol{M}]\left(\frac{\{\boldsymbol{D}\}_{n+1}}{\beta\Delta t^2}+\frac{\left\{\dot{\boldsymbol{D}}\right\}_n}{\beta\Delta t}+\frac{\left\{\ddot{\boldsymbol{D}}\right\}_n(1-2\beta)}{2\beta}\right)$$
$$+[\boldsymbol{C}]\left(\frac{\{\boldsymbol{D}\}_n\gamma}{\beta\Delta t}+\frac{\left\{\dot{\boldsymbol{D}}\right\}_n(\gamma-\beta)}{\beta}+\frac{\left\{\ddot{\boldsymbol{D}}\right\}_n\Delta t(\gamma-2\beta)}{2\beta}\right) \tag{30a}$$

$$\left\{\dot{\boldsymbol{D}}\right\}_{n+1}=\frac{\left(\{\boldsymbol{D}\}_{n+1}-\{\boldsymbol{D}\}_n\right)\gamma}{\beta\Delta t}-\frac{\left\{\dot{\boldsymbol{D}}\right\}_n(\gamma-\beta)}{\beta}-\frac{\left\{\ddot{\boldsymbol{D}}\right\}_n\Delta t(\gamma-2\beta)}{2\beta} \tag{30b}$$

$$\left\{\ddot{\boldsymbol{D}}\right\}_{n+1}=\frac{\left(\{\boldsymbol{D}\}_{n+1}-\{\boldsymbol{D}\}_n-\Delta t\left\{\dot{\boldsymbol{D}}\right\}_n\right)}{\beta\Delta t^2}-\frac{\left\{\ddot{\boldsymbol{D}}\right\}_n(1-2\beta)}{2\beta} \tag{30c}$$

where $\left[\boldsymbol{K}^{eff}\right]=[\boldsymbol{M}]/\left(\beta\Delta t^2\right)+[\boldsymbol{C}]\gamma/(\beta\Delta t)+[\boldsymbol{K}]$. The initial condition for Eqs (30a)-(30c) is shown as follows:

$$\left\{\ddot{\boldsymbol{D}}\right\}_0=[\boldsymbol{M}]^{-1}\left(\left\{\boldsymbol{R}^{ext}\right\}_0-[\boldsymbol{K}]\{\boldsymbol{D}\}_0-[\boldsymbol{C}]\left\{\dot{\boldsymbol{D}}\right\}_0\right) \tag{31}$$

With Eqs (30a)-(30c) and (31), we can obtain $\{\boldsymbol{D}\}_1$, $\{\boldsymbol{D}\}_2$, $\{\boldsymbol{D}\}_3$, and so forth. From the above-mentioned details, it is obvious that the numerical calculation of TRA is more complicated than that of HRA and therefore requires more computation resources when the model being larger.

In TRA or HRA, we need to suppress reflections from the hard boundary to get rid of the unwanted resonance peaks. Based on the wave equation in spherical coordinate, artificial boundary can be equivalent to many continuous distribution parallel viscous-spring systems. The coefficients of stiffness and damping are given as follows:

$$K_T=\frac{\alpha_T G}{LN};\ K_N=\frac{\alpha_N G}{LN} \tag{32a}$$

$$C_T=\frac{\sqrt{G\rho}}{N};\ C_N=\frac{\sqrt{E\rho}}{N} \tag{32b}$$

where K_T and K_N are tangential and normal stiffness coefficients of springs, respectively; ρ is the material density of matrix silicon; C_T and C_N are tangential and normal damping coefficients, respectively; G and E are shear modulus and Young's modulus of matrix silicon, respectively; L and N are the distance from exciting source to artificial boundary and number of viscous-spring systems attached to the boundary, respectively; α_T and α_N are the tangential and normal modified coefficients for artificial boundary, respectively. α_T and α_N are assigned with 0.67 and 1.33, respectively [20].

2.4 Periodic structure without/with Substrate by SC-PWE method

The super-cell plane wave expansion (SC-PWE) method is another efficient way to calculate the plate-mode waves of the phononic crystal plates. As shown in Fig. 4, we establish a 3D

model in Cartesian coordination to calculate the elastic band structures of 1D phononic crystal plates, where the periodic composite plate consists of alternate *A* and *B* strips, *C* is the LIM layer, and *D* is the substrate, respectively [21].

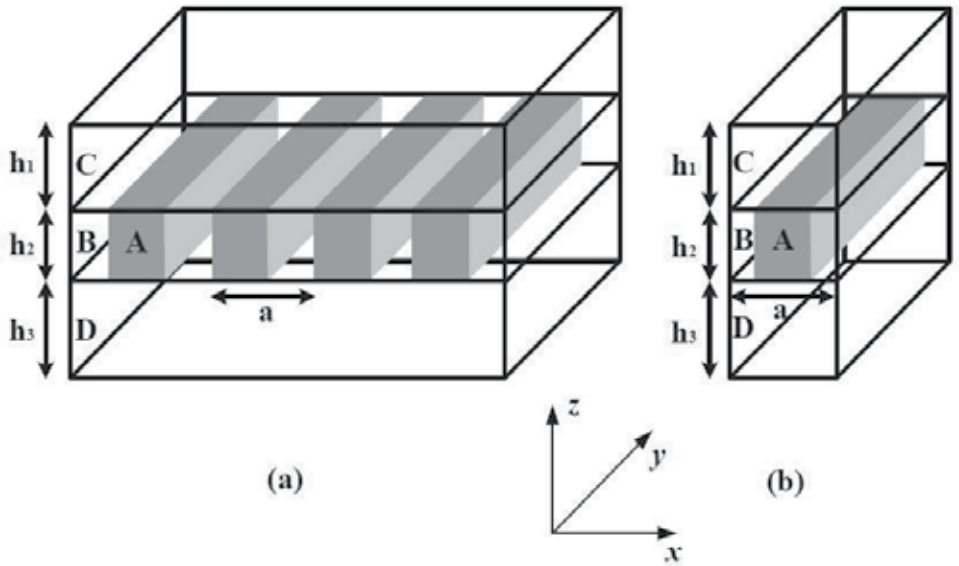


Fig. 4. (a) 1D Lamb wave phononic crystal plate sandwiched between two layers of homogeneous materials, and (b) 3D super-cell used in the computation.

The LIM is an imaginary material with relatively low elastic moduli for approximately meeting the requirement of free boundary condition and an extremely low mass density, which leads the sound speed in the LIM to be much larger than that in usual solid material. In this chapter, the LIM is assumed as an isotropic material with $C_{11} = 2 \times 10^6$ N/m², $C_{44} = 1 \times 10^6$ N/m², $C_{12} = 0$ N/m² and $\rho = 1 \times 10^{-4}$ kg/m³. The choice of such unphysical high sound speeds for the LIM is in good agreement with the numerical condition derived by Tanaka et al [22]. With these values, both good numerical convergence and computing accuracy can be achieved. The thickness of plate h is assumed to be 2 mm and $h = h_2 + h_3$ where $h_3 = 0$ mm for the case without substrate. The thickness of the LIM layer h_1 is defined to be $5h_2$ to reduce unexpected wave coupling between two nearest phononic layers in z direction [23]. In the absence of body force and strain in y direction, the SH mode wave in 1D plate can be decoupled. Regardless of the wave propagating in y direction, the elastic wave equations of phononic crystal are given by:

$$\rho(\boldsymbol{r})\ddot{u}_p = \partial_q[C_{pqmn}(\boldsymbol{r})\partial_n u_m] \quad (p = 1,2,3) \tag{33}$$

where $\boldsymbol{r} = (x, z)$. This equation can be solved by a standard Fourier expansion to $\rho(\boldsymbol{r})$, $C_{pqmn}(\boldsymbol{r})$ and $u(\boldsymbol{r}, t)$, which are all position-dependent values. For convenience, we put $\alpha = (\rho, C_{pqmn})$ and then we can obtain the following equations:

$$\alpha(\boldsymbol{r}) = \sum_{\mathbf{G}_x} \sum_{G_z} \alpha_{\mathbf{G}} \exp[i(G_x x + G_z z)] \tag{34}$$

$$u(\boldsymbol{r},t)=\sum_{G_x}\sum_{G_z}u_{\boldsymbol{G}}\exp[i(G_x x+G_z z)]e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} \tag{35}$$

where $\boldsymbol{k}=(k_x,k_z)$ is the Bloch wave vector and the 2D reciprocal-lattice vector $\boldsymbol{G}=(G_x,G_z)$, respectively. Substituting equations (34) and (35) into wave equation (33), we can obtain:

$$\omega^2\begin{pmatrix}\rho_{G-G'} & & \\ & \rho_{G-G'} & \\ & & \rho_{G-G'}\end{pmatrix}u_{G'}=\begin{pmatrix}M^{11}_{G,G'} & M^{12}_{G,G'} & M^{13}_{G,G'}\\ M^{21}_{G,G'} & M^{22}_{G,G'} & M^{23}_{G,G'}\\ M^{31}_{G,G'} & M^{32}_{G,G'} & M^{33}_{G,G'}\end{pmatrix}u_{G'} \tag{36a}$$

The explicit expressions of the matrix elements $M^{lm}_{G,G'}$, ($l=m=1-3$) are:

$$\begin{aligned}
&M^{11}_{\mathbf{G},\mathbf{G}'}=C^{11}_{\mathbf{G}-\mathbf{G}'}(k_x+G'_x)(k_x+G_x)+C^{44}_{\mathbf{G}-\mathbf{G}}G'_zG_z \qquad && M^{12}_{G,G'}=0\\
&M^{13}_{G,G'}=C^{12}_{G-G}G'_z(k_x+G_x)+C^{44}_{G-G'}(k_x+G'_x)G_z \qquad && M^{21}_{G,G'}=0\\
&M^{22}_{G,G'}=C^{44}_{G-G'}(k_x+G'_x)(k_x+G_x)+C^{44}_{G-G}G'_zG_z \qquad && M^{23}_{G,G'}=0\\
&M^{31}_{G,G'}=C^{44}_{G-G}G'_z(k_x+G_x)+C^{12}_{G-G'}(k_x+G'_x)G_z \qquad && M^{32}_{G,G'}=0\\
&M^{33}_{G,G'}=C^{44}_{G-G'}(k_x+G'_x)(k_x+G_x)+C^{11}_{G-G}G'_zG_z
\end{aligned} \tag{36b}$$

where the Fourier coefficients $C^{pq}_{G-G'}$ are related to $C_{pqmn}(\boldsymbol{r})$ in a conventional manner. As shown in equation (36a), characteristic frequency ω is exactly the squared generalized eigenvalue of density matrix and elastic constant matrix. The coefficients $C^{pq}_{G-G'}$ and $\rho_{G-G'}$ takes the form:

$$\alpha_{G-G'}=\frac{1}{V_c}\iiint_{(Supercell)}\alpha(\boldsymbol{r})\exp[-i(\boldsymbol{G}-\boldsymbol{G}')\cdot\boldsymbol{r}]d^3\boldsymbol{r} \tag{37}$$

where V_c is the volume of super-cell. With the above-mentioned equations, we can easily obtain the band structure of 1D phononic crystal plate.

2.5 Quasiperiodic structure by FE method

As shown in Fig.5, the quasiperiodic composite plate consists of material *A* of width d_A and material *B* of width d_B. The lattice spacing is $D=d_A+d_B$. When the distribution of materials *A* and *B* is arranged according to the Fibonacci sequence, one obtains a quasiperiodic system [24]. We create the Fibonacci sequence *B*, *BA*, *BAB*, *BABBA*, *BABBABAB*, *BABBABABBABBA*, … according to the production rule $S_j=S_{j-1}\,|\,S_{j-2}$ for $j\geq 3$ with $S_1=B$ and $S_2=BA$. When *A* and *B* are put along the chain alternately, a periodic model is obtained. We introduce parameter $\Phi=d_A/d_B$ to describe the ratio of the two components. Φ is fixed at 1.0 and the number of layers *N* is 13 throughout the section unless otherwise stated. The wave propagates along the *x* direction of the plate bounded by planes $z=0$ and $z=L$. We consider a 2D problem, in which all field components are assumed to be *y* independent.

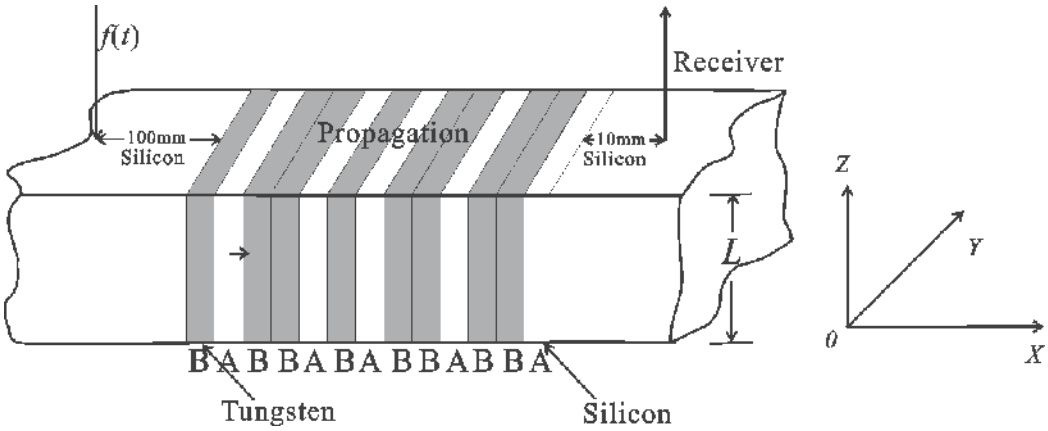


Fig. 5. The configuration of 1D quasiperiodic composite plate consisting of Tungsten and Silicon strips arranged following the Fibonacci sequence.

To demonstrate the structures of the band gaps for Lamb waves in the 1D quasiperiodic systems and the difference from that of periodic systems, we calculate TPS of the transient Lamb waves by using the TRA. We suppose that a Lamb wave is excited by a line laser pulse with a spatial Gaussian distribution (Gaussian radius = 0.2 mm). The laser pulse, which is normally incident to the surface of the studied plates, generates the Lamb wave propagating along the x direction. The laser-generated force source $f(t)$ is simulated as a delta function, which is perpendicular to the surface of the plate [25].

The elastic properties of the materials in the numerical calculations are the same as mentioned in above sections; and the thickness of the plates (L) of 1.0 mm. The step sizes of temporal and spatial discretization in the FE calculations are fine enough for the convergence of the numerical results. Lamb waves are excited by the force function $f(t)$ at x=0, and are received at the point 10 mm away from the superlattices array. The generation source is far from the Fibonacci superlattices in order to obtain approximately plane waves when the wave fronts reach the plate. The received vertical displacement in time domain is Fourier-transformed into the frequency domain to yield the TPS.

We also adopt the HRA to study three quasiperiodic systems. Two Generalized Fibonacci Systems (Type A and Type B) [26] are obtained inductively through the following transformations:

$$A \to AAB,\ B \to A \quad \text{for Type A Fibonacci System} \tag{38a}$$

$$A \to ABB,\ B \to A \quad \text{for Type B Fibonacci System} \tag{38b}$$

We can generate the two quasiperiodic systems, as shown as follows:

$$AABAABAAABAABAAABAAB\ldots \text{ for Type A Fibonacci System} \tag{39a}$$

$$ABBAAABBABBABBAAABBA\ldots \text{ for Type B Fibonacci System} \tag{39b}$$

It is interesting to find out that Generalized Fibonacci Systems are very flexible in forms and by changing the transformations ($A \to ABA$, $B \to A$ …) we can obtain many other quasiperiodic systems.

Then, we can introduce the third quasiperiodic system (Double-period System) into this model. The recursion relation for Double-period System is $A \to AB$, $B \to AA$ [27]. With the recursion relation, we can obtain the sequence of the Double-period System:

$$ABAAABABABAAABAAABAA\ldots \text{for Double-period System} \tag{40}$$

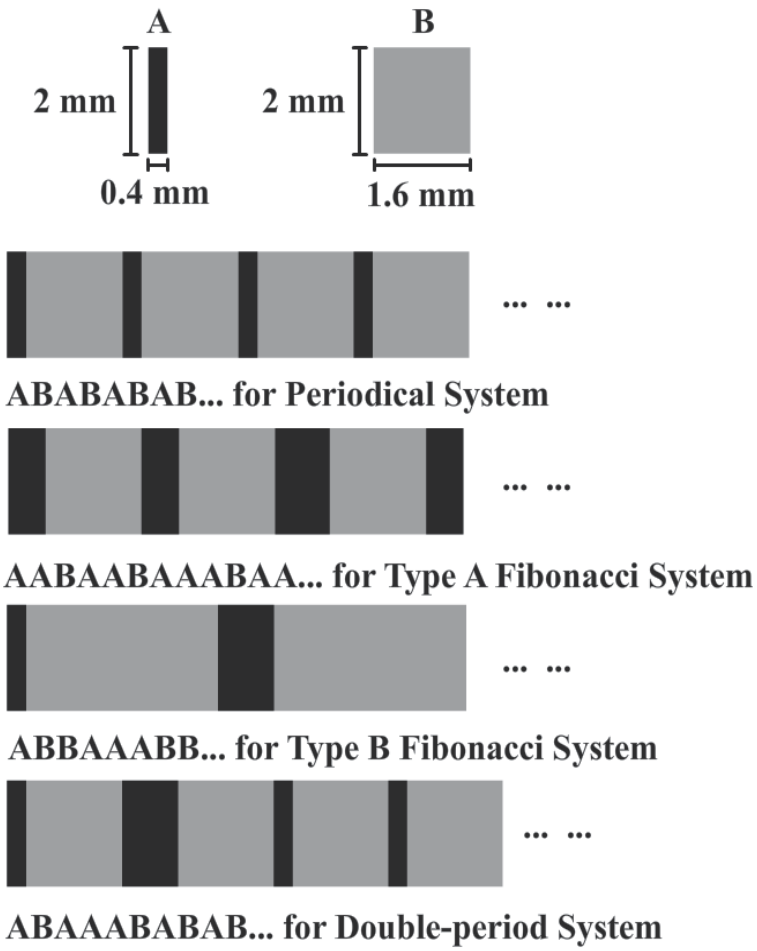


Fig. 6. The schematic diagram of the four systems: Periodic System, Type A Fibonacci System, Type B Fibonacci System and Double-period System, respectively.

Fig. 6 shows the scheme of the four different systems, namely, Periodic System, Type A Fibonacci System, Type B Fibonacci System and Double-period System, respectively. In numerical simulations, the homogeneous media A and B are gold and silicon, with the thicknesses of media A and B are 2 mm, and the widths of media A and B are 0.4 and 1.6 mm, respectively.

3. Lower-order lamb waves in 1D composite thin plates without/with substrate

In order to demonstrate the existence of band gaps for low-order Lamb wave modes in the 1D periodic structure as shown in Fig.1, we have calculated the dispersion curves for a cubic medium (silicon) of a 1 mm thick plate by considering only the fundamental term in the Fourier and Floquet series [16], as shown in Fig. 7(a). Fig. 7(b) displays the dispersion

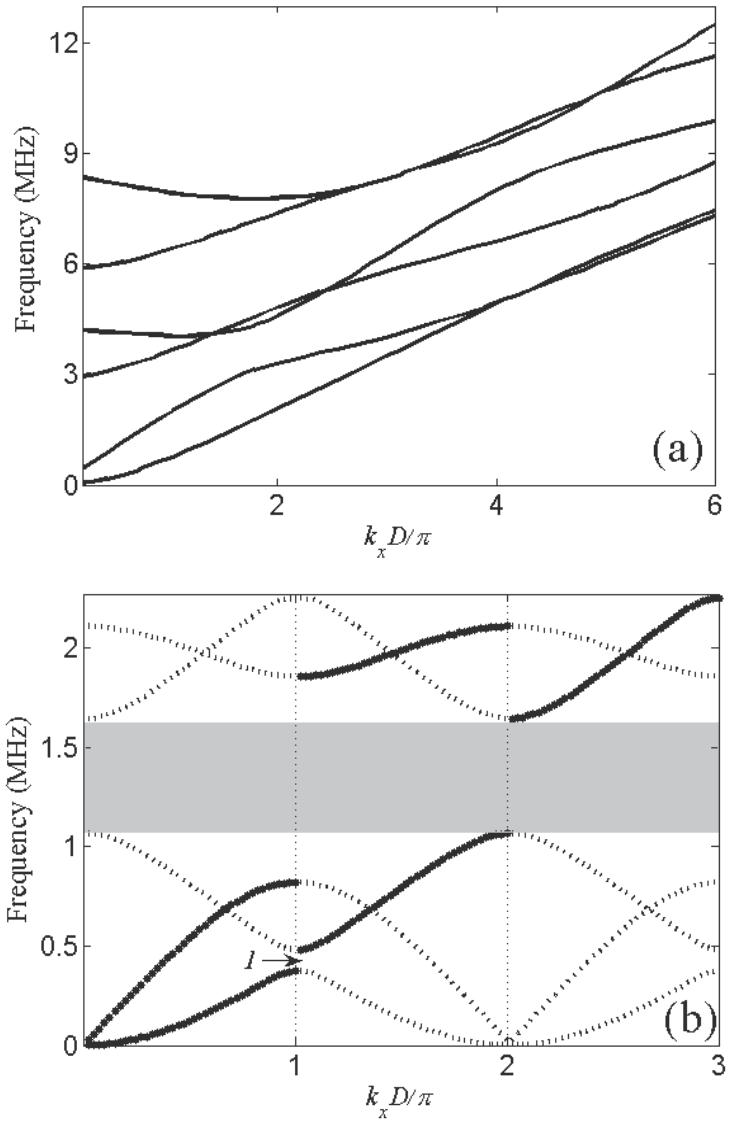


Fig. 7. Schematic representation of the Lamb wave dispersion curve for (a) a homogeneous Si plate with $L = 1.0$ mm (b) composite thin plate (W / Si) with $f = 0.5$, $L = 1.0$ mm, and $D = 2.0$ mm.

curves of four lower-order modes along the boundary of the mini-Brillouin zone with filling ratio $f = 0.5$, $L = 1.0$ mm, and $D = 2.0$ mm. One can obviously observe the modifications produced by resonant reflections in the strip lattice. The dashed vertical line identifies the frequency zone where all the Lamb wave modes are resonantly reflected by the periodic lattice of strips. The proposed approach allows one to identify that the forward propagating Lamb wave modes are not coupled with the backward propagating modes.

By comparing Fig. 7(a) with Fig. 7(b), one can easily find that there exists a band gap from 1065 to 1642 kHz for the lower-order Lamb wave modes propagating in the 1D periodic structure. The gap width ($\Delta\Omega$) is 577 kHz and the corresponding gap/midgap ratio ($\Delta\Omega / \Omega_m$, Ω_m is the midgap frequency) is approximately 0.426. In order to analyze the influence of the ratio L / D for the band gap width, we also calculate the dispersion curves of the lower-order modes with $f = 0.5$, $L = 2.0$ mm, and $D = 2.0$ mm, as shown in Fig. 8. It is apparent that there are two band gaps (from 806 to 1167 kHz and from 1438 to 1863 kHz, respectively) for the ratio $L / D = 1$. The gap widths are 361 and 425 kHz, and the corresponding gap/midgap ratios are about 0.366 and 0.255, respectively.

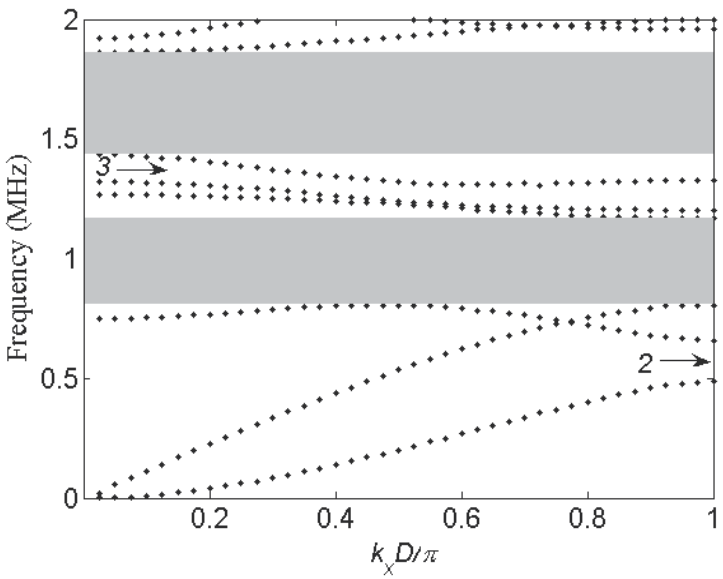


Fig. 8. Dispersion curves of Lamb wave modes for 1D finite thickness composite plate with filling ratio $f = 0.5$, $L = 2.0$ mm, and $D = 2.0$ mm.

Basically, there are three parameters that influence the formation of band gaps, i.e., L / D, f, and the contrast between the physical parameters of the constituents. It is rather intuitive that L / D is very crucial for the formation of a band gap. If it is either too small or too large, there should be no band gaps for lower-order modes. Fig. 9 depicts the gap width of the lowest band gap as a function of L / D with $f = 0.5$ and $D = 2$ mm for tungsten/silicon supperlattices. It is noteworthy to point out that the lowest band gap opens up over a domain of the ratio of L / D defined by $0.15 \le L / D \le 1.64$. The maximum value of gap width appears at $L / D \approx 0.53$ for the lowest band gap and reaches 610 kHz as shown in Fig. 9.

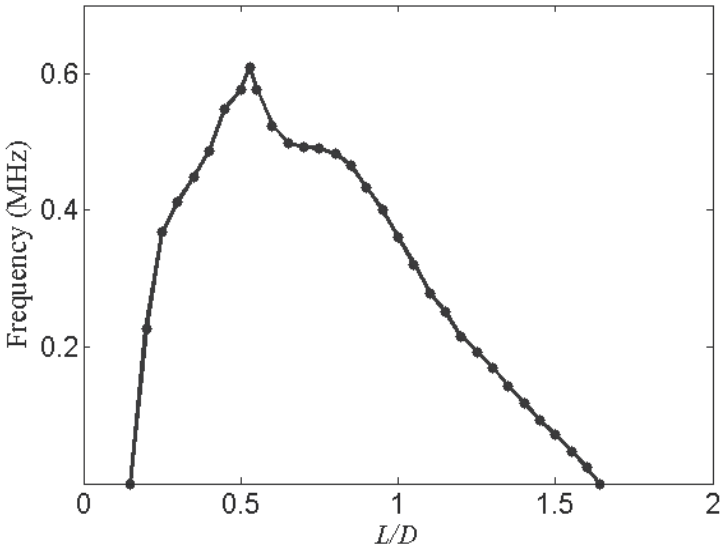


Fig. 9. The width of the lowest band gap at the filling fraction f = 0.5 versus the value of L/D.

It is noted that the value of the normalized gap width of the lowest band gap in the systems increases progressively with the increase of the value of the ratio of L/D until a critical value and then decreases. In fact, a plate can support a number of Lamb wave modes depending on the value of the ratio L/λ, where λ is the acoustic wavelength. When the periodicity of these Lamb waves matches the lattice spacing, stop bands appear in the Lamb wave dispersion curves [28]. There is a high interaction when the wavelength of Lamb wave is close the lattice constant, which induces mode conversion and reflections. When the wavelength of Lamb wave is different from the periodicity of the lattice constant, the interaction is weak. On the another hand, the midgap frequency of forbidden gap is inversely proportional to the lattice constant D [29], therefore, the value of the ratio of L/D is important for the width of the band gap for the Lamb waves in the periodic composite systems.

In order to demonstrate further the existence of the band gaps for the lower-order modes in the 1D periodic structure, the finite element method (FEM) is employed to calculate the transmitted power spectra (TPS) for the finite periodic structure as shown in Fig.3.

Fig. 10 shows the TPS for the 1D composite structure plate with $f = 0.5$, $L = 1.0$ mm, and $D = 2.0$ mm. There is a broad region from 1060 to 1630 kHz that is less than –30dB. The result shows good agreement with that by PWE method. The TPS is also depicted in Fig. 10 from a pure Tungsten plate with the same dimensions, and no sharp attenuation in any frequency domain is observed.

For the second sample, f, D, and the configuration are the same with the first one, and only the thickness of the plate is different (L=2 mm). Fig. 11 depicts the TPS for 1D plate with periodic structure and without periodic structure. The frequency range of the gaps of Lamb waves by PWE is almost the same with those of large attenuation in the calculated TPS .The first gap extends from the frequency of 804 up to 1176 kHz and the second from 1436 to 1869 kHz, which are less than –45dB.

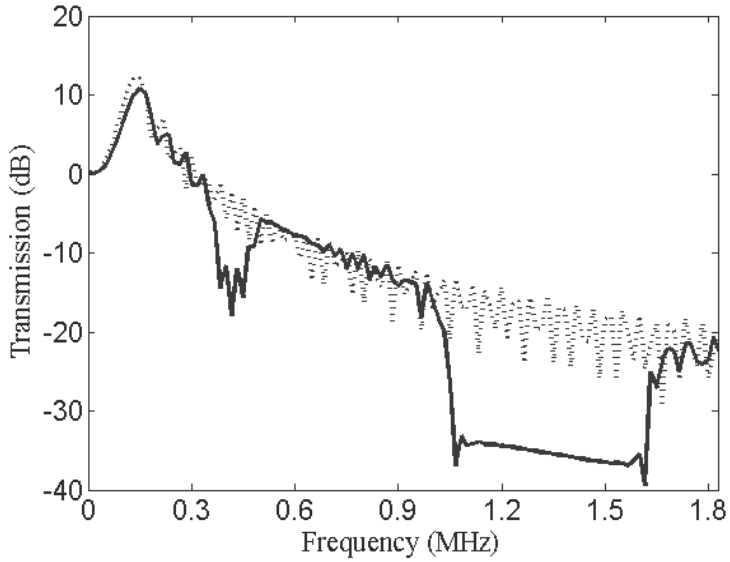


Fig. 10. The TPS computed by the FE method with $f = 0.5$, $L = 1.0$ mm, and $D = 2.0$ mm through the composite pate (solid line) and a pure Tungsten plate (dashed line).

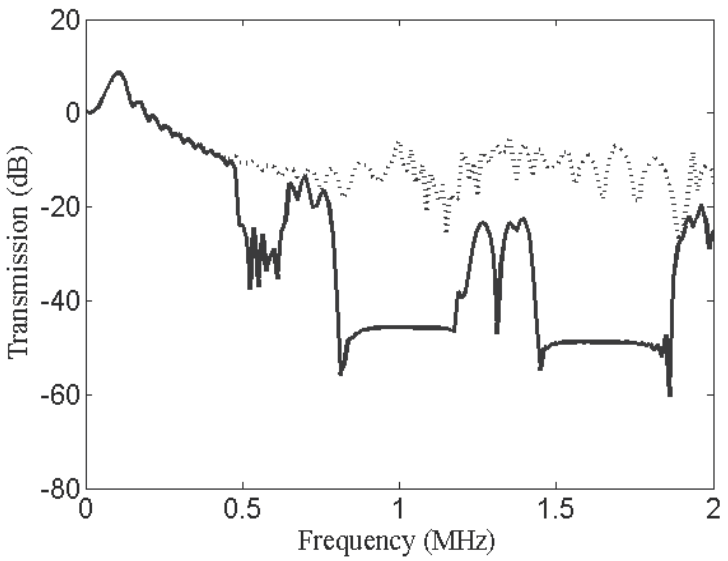


Fig. 11. The TPS computed by the FE method with $f = 0.5$, $L = 2.0$ mm, and $D = 2.0$ mm: through the composite plate (solid line) and through a pure Tungsten plate (dashed line).

It is interesting to notice that there are some slight dips centered at about 0.4MHz in Fig. 6, or 0.5MHz and 1.3MHz in Fig. 11. These dips attribute to the band gaps of antisymmetric modes, but not absolute band gap of both symmetric and antisymmetric modes, which can be observed in Figs. 7(b) and 8, indicated by the arrows 1, 2 and 3. On the other hand, the

lower frequency components of the modes is mainly excited for the thin plate [30], so that the TPS energy displays a monotonic decrease as the frequency increases as shown in Fig. 10. When the plate is thicker, the higher frequency components increase, which makes the attenuation of the TPS energy less, as shown in Fig. 11.

Since various modified photonic crystal structures, such as regular arrangements of individual metal nanoparticles on dielectric substrate, have been the subject of extensive research in recent years, it is meaningful to study the elastic modes in the system with PC layer coated on uniform substrate [31]. In fact, from the application point of view, when the thickness of the periodic thin plate is of the order of magnitude of several hundred micrometers, one needs to use substrate to support such a plate. Therefore, the effect of the substrate is important and cannot be neglected.

We study the influence of substrate on the band structure of the Lamb wave in PC layer by FEM. We also employ V-PWE to calculate the dispersion curves of Lamb wave. As we shall demonstrate that the locations and widths of band gaps on the dispersion curves from the V-PWE method are in good agreement with the results from the TPS by FEM.

In order to demonstrate the influences of different substrates on band gaps in the PC layer with substrate, we study three types substrate: hard material (Tungsten), soft material (Rubber) and medium hardness material (Silicon). The corresponding TPS (in green, blue, and red, respectively) are shown in Fig. 12 (a-b) for different h_2. For comparison, we also show the case without substrate (black).

It can be easily seen that there exist two band gaps for the Lamb modes propagating in the 1D periodic model. The first gap extends from the frequency of 820 kHz up to 1160 kHz and the second one from 3050 kHz to 3360 kHz, which is less than –38 dB, as shown in Fig. 12 (black line).

Comparing the TPS of the 1D PC layer without substrate (black line) with that of coated Tungsten substrate (green line) [Fig. 12(a)], we can see that the width of the first band gap decreases to some slight dips centered at about 1.0MHz (green line). The second band gap has a little decay in frequency domain (from 2850 to 3250 kHz). This demonstrates that the influence of hard substrate on band gaps is strong even when the substrate is very thin. The band gaps disappear rapidly when the substrate becomes thicker as shown in Fig. 12(b). It is because when the substrate becomes thicker, more energy will go via the substrate instead of via the PC, so that the interference from different periodic layers becomes less important.

The TPS for the 1D PC layer coated on Rubber substrate with different thickness h_2 are also shown in Fig. 12(a-b) (blue line). It is seen that there is no obvious change in the band gaps when the substrate is thin, especially, for the first band gap at low frequency region. However, as the thickness of substrate increases, interesting things happen, as shown in Fig. 12(b) (h_2=0.5mm). The band gap in the PC layer coated on Rubber substrate does not disappear but becomes smoother in the band gaps domain. Three band gaps appear, namely, from 750 kHz up to 1200 kHz, from 1720 kHz up to 1950 kHz and from 2920 kHz up to 3640 kHz, respectively. Compared with the band gap of the PC plate without substrate, we can see a broad band appears in the range of 1720 kHz and 1950 kHz. From Fig. 12(b), it is clearly seen that more band gaps appear, such as some band gaps in low frequency domain, which is opposite to the hard substrate. The appearance of more bands is due to the more interference from the boundary as Rubber is softer than Tungsten and Silicon, therefore there is more reflection at the interface between the PC and the substrate.

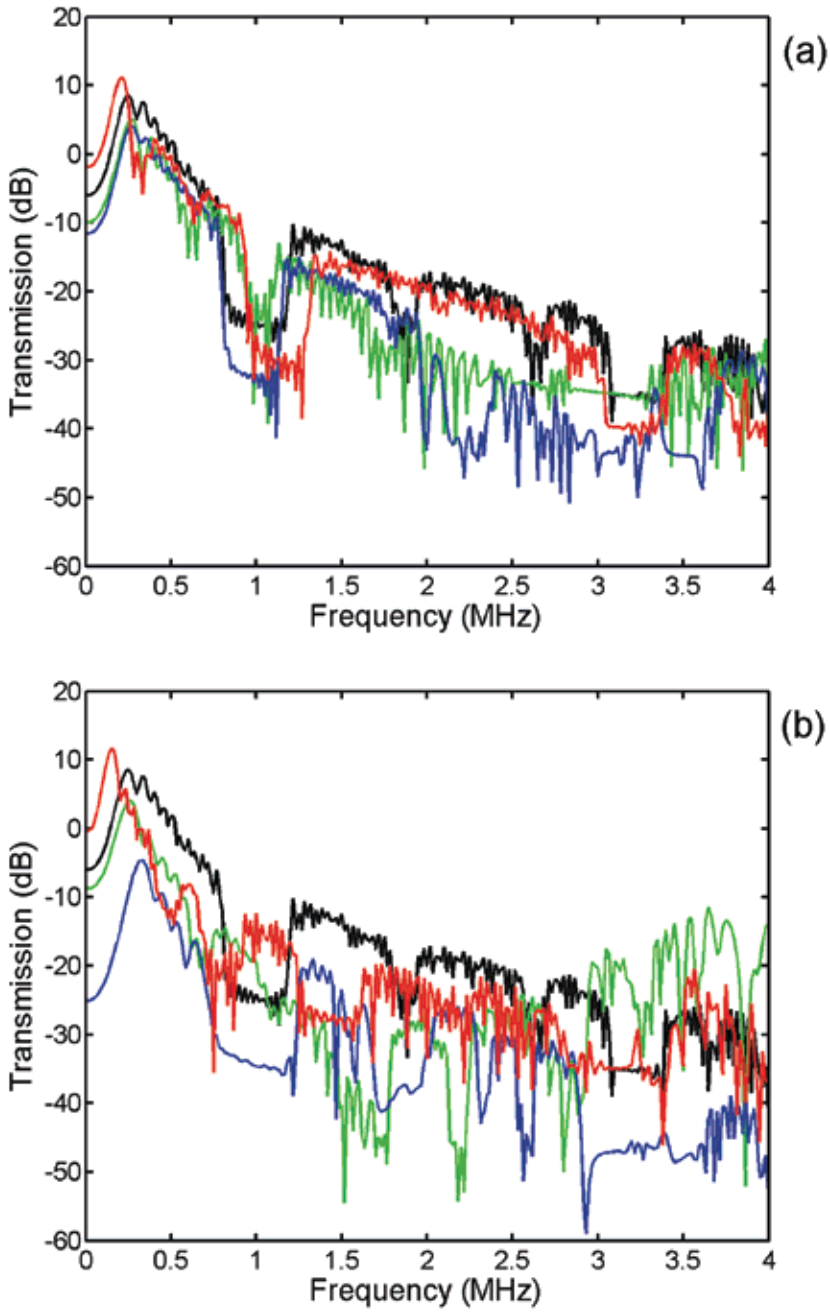


Fig. 12. (color online) The TPS for the 1D PC layer without substrate (black line), and for the 1D PC layer coated on Tungsten substrate (green line), Rubber substrate (blue line), Silicon substrate (red line), respectively, with different h_2 : (a) $h_2 = 0.125$ mm; (b) h_2=0. 50 mm

We also show the TPS for the 1D PC layer coated on Silicon substrate with different substrate thickness h_2. From Fig. 12(a) (red line), it can be easily found that there exist two band gaps for the Lamb modes propagating in the 1D PC layer coated on Silicon substrate. The first gap extends from 920 kHz up to 1280 kHz and the second from 3050 kHz to 3400 kHz, which are less than –40dB. Compared with Fig. 12(a) (black line and red line), we can see there is no obvious change between the band gaps when the substrate is thin. Although there are some band gaps appearing like some band gaps in low frequency domain when the thickness of substrate increases, the depth of band gap decreases. For example, one can see that although there is a band gap at about 1.5 MHz, the depth of band gaps for the model of Silicon substrate becomes very small as the thickness of substrate increases. Therefore, the influence of the Silicon substrate is between those of the hard substrate and the soft substrate.

To verify our numerical results, we calculate the dispersion curves of Lamb wave modes propagating along the x direction in the presence of the uniform substrate by V-PWE method.

Fig. 13 displays the dispersion curves of the lower-order modes of the 1D PC layer coated on Silicon substrate with different substrate thickness h_2. It is apparent that there are two band gaps (from 980 to 1285 kHz and from 3020 to 3380 kHz, respectively) for the h_2=0.125mm, as shown in Fig. 13(a). The gap widths are 305 kHz and 360 kHz, respectively, and the corresponding gap/mid-gap ratios are about 0.269 and 0.112, respectively. The results calculated by the V-PWE method show that the locations and widths of band gaps on the dispersion curves are in good agreement with the results on the transmitted power spectra by FEM, as shown in Fig. 12(a) (red line).

Some band gaps appear in low frequency domain with the increase in the thickness of substrate, which is also found by V-PWE method. For example, we can see that there are three band gaps (from 685 to 820 kHz, from 1320 to 1590 kHz and from 3120 to 3250 kHz) for the model of Silicon substrate with the thickness of 0.5mm as shown in Fig. 13(b), which is in good agreement with the results by FEM as shown in Fig. 12(b) (red line).

Here, we give a qualitative physical explanation of above results. When the substrate is Tungsten material, because the ratio of acoustic impedances of Tungsten and Silicon $\rho_S C_S / \rho_T C_T \approx 0.2$ (where ρ_S (C_S) and ρ_T (C_T) are the mass densities (the acoustic velocities of longitudinal wave) of Silicon and Tungsten, respectively), the interface between the PC layer and the substrate is equivalent to a hard boundary condition, at which the phase change of the reflected wave pressure is less than 90°. The superposition of the reflective wave will destroy the formation condition of band gap, as the formation of band gap is due to the destructive interference of the reflective waves. Therefore, the influences on band gaps are significant even when the substrate is very thin. On the other hand, due to the interface is not strictly strong, the Lamb wave can transmit partially to the uniform substrate, and then the band gaps disappear rapidly when the substrate becomes thicker.

In contrast, when the substrate is Rubber material, because the acoustic impedances of Silicon is approximately seven times of that of Rubber, the interface between the PC layer and the substrate can be approximately considered a soft boundary, at which the phase change of the reflected wave is larger than 90°. The superposition of the reflective waves will lead to the band gap. As the substrate is very thin, the influences on band gaps are negligible. On the other hand, as the interface is not strictly a pressure-released boundary, the Lamb wave can transmit partially to the uniform substrate. Because the mass density

and the elastic constants of Silicon are much larger than that of Rubber, the acoustic wave will be localized in the soft Rubber material. Therefore, band gaps become deeper as the thickness of substrate increases. If the substrate is Silicon, which is the same as the matrix material, the acoustic wave does not reflect at $z = 0$, In this case, the influence of the substrate is between those of the hard substrate and the soft substrate.

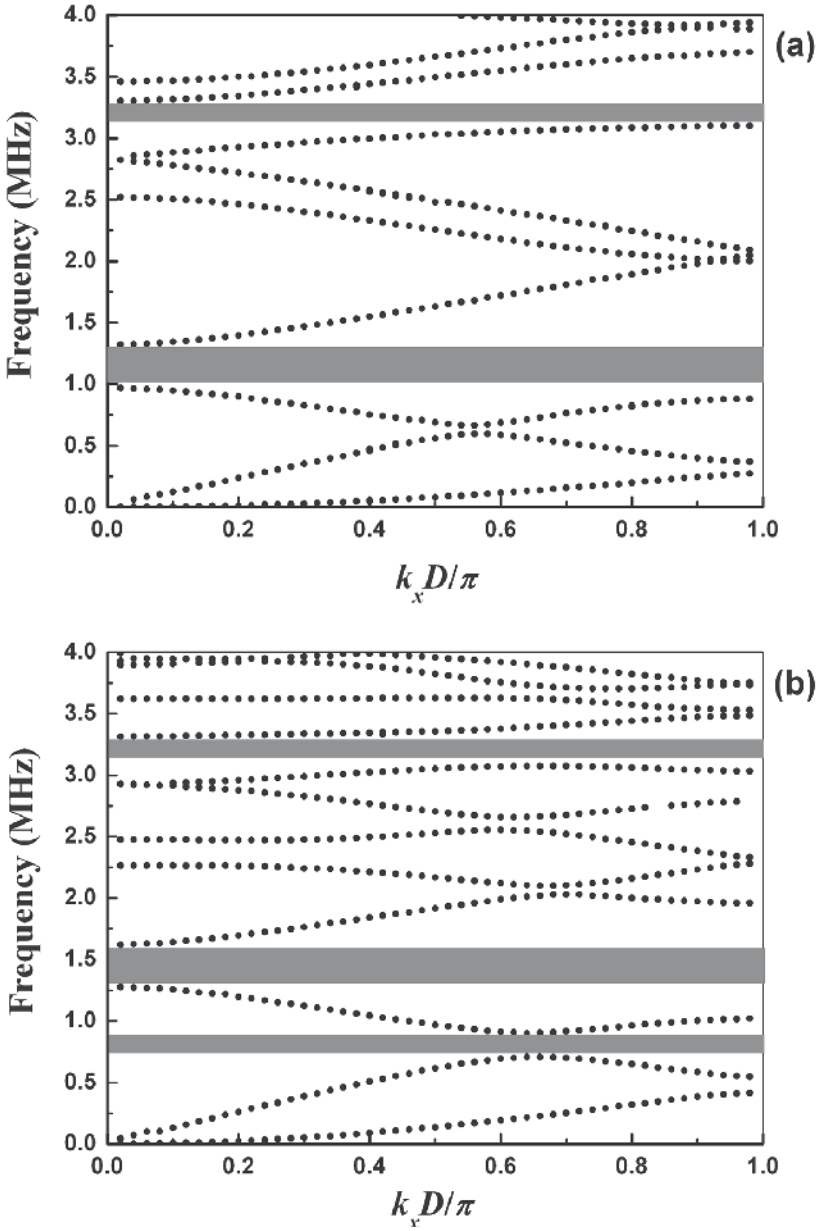


Fig. 13. The dispersion curves of Lamb modes of the 1D PC layer coated on Silicon substrate with different h_2 : (a) h_2=0.125 mm; (b) h_2=0.50 mm

4. Lamb waves in 1D quasiperiodic composite thin plates

In this section, we study numerically the band gaps of Lamb waves in 1D quasiperiodic thin plate. The motivation of the study lies in the factor that a lot of real-world materials are quasiperiodic [32-33]. In particular, since Merlin et al.[34] reported the realization of Fibonacci superlattices, a lot of interesting physical phenomena have been observed in *x*-ray scattering spectra, Raman scattering spectra, and propagating modes of acoustic waves on corrugated surfaces [35-37].

First, we show the dependence of TPS on L/D. From Fig. 14(a-e), the TPS are shown for the periodic and quasiperiodic composite plates with L/D= 0.3, 0.5, 0.54, 0.6, and 0.68, respectively. For comparison, the TPS for a pure Silicon plate of 1 mm thickness is also shown in order to demonstrate the band gaps. Fig. 14(a) shows that for such a pure silicon plate there is no band gap at all. However, two band gaps are clearly seen in the periodic system. The first band extends from frequency of 570 up to 760 kHz and the second one from the 1550 up to 1960 kHz. With the same parameters, the two bands are not so obvious in a quasiperiodic plate.

When L/D is increased to 0.5 [see Fig. 14(b)], interesting things happen. It is evident that for the periodic model there exists a band gap from 1050 up to 1615 kHz. However, for the quasiperiodic plate, a clear band split is seen from 1085 up to 1286 kHz and from 1460 up to 1710 kHz, and a new band appears in the range of 2010-2275 kHz.

As L/D is increased to 0.54 [Fig. 14(c)] and 0.6 [Fig. 14(d)], the only band gap in the periodic system does not change too much; it just shifts a little toward the high frequency. However, the situation changes in the quasiperiodic system. In the case of $L/D = 0.54$, the band gap is split into two subbands, namely, from 1210 up to 1380 kHz and from 1505 up to 1780 kHz. Two more new bands appear from 2050 up to 2420 kHz and from 2750 up to 2950 kHz. In the case of $L/D = 0.6$, only two bands appear, namely, from 1360 up to 1949 kHz and from 2205 up to 2685 kHz.

From the results shown in Figs. 14(a)-(d), we can say that the band structures of a quasiperiodic system depend strongly (or sensitively) on the parameter L/D, whereas that in a periodic system does not. A quasiperiodic system has more forbidden gaps than that a periodic system has. This can be explained from the following. The 1D Fibonacci sequence is the project of the 2D square periodic lattice; it implicitly includes the periodicity of a multidimensional space. In fact, a quasiperiodic structure may be considered as a system made up of many periodic structures [38].

Moreover, the change of the ratio L/D also leads to the changes of the number of splitting band gaps. Physically, as the ratio L/D changes to an appropriate value, due to reflections at the plate boundaries, the interaction between longitudinal and transversal strain components becomes strong. For the Lamb modes, the restriction of boundary conditions leads to intermode Bragg-like reflections in the quasiperiodic superlattices [39]. As a result, much more physical phenomena are present compared with the bulk wave propagation in the Fibonacci chains.

In general, there are three parameters that influence the formation of band gaps, namely, L/D, Φ, and λ (the acoustic wavelength). The number of Lamb wave modes in a plate depends on the value of L/λ. The midgap frequency of forbidden gap is inversely proportional to the lattice spacing D [29]. Therefore, it is rather intuitive that L/D is very crucial for the formation of band gaps for Lamb waves. In fact, it is also found that the difference between the forbidden gaps in quasiperiodic and periodic systems disappears

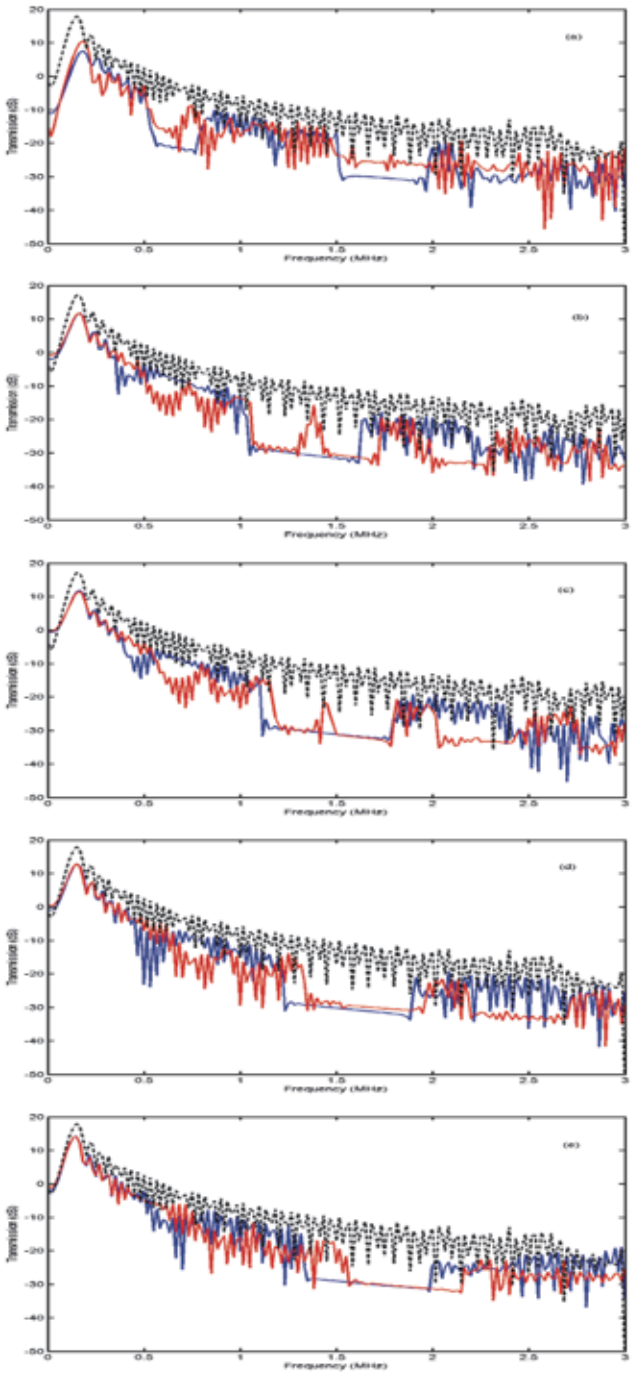


Fig. 14. (color online) The TPS for the periodic plate (blue), the quasiperiodic plate (red), and a pure Silicon plate (dashed black), respectively. (a) L/D=0.3, (b) L/D=0.5, (c) L/D=0.54, (d) L/D=0.6, (e) L/D=0.68.

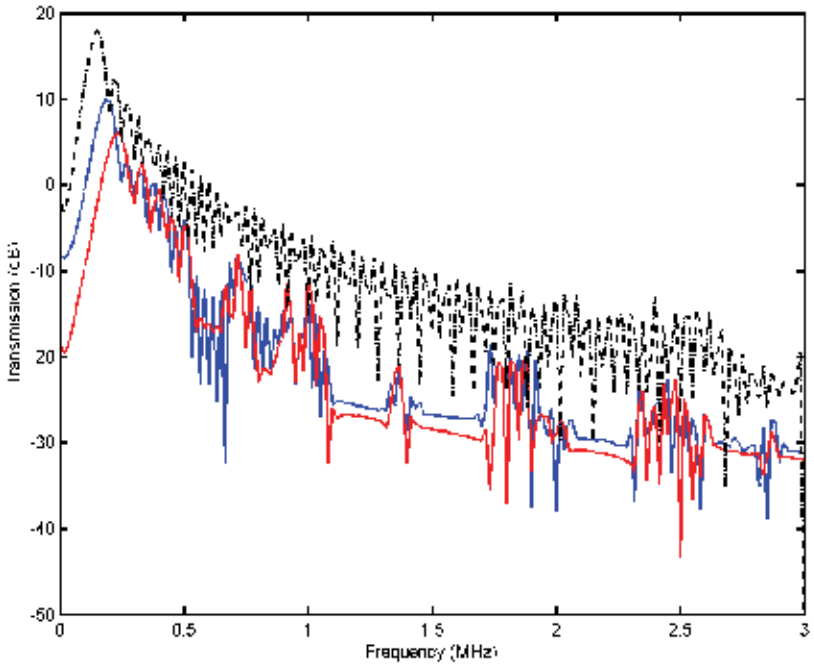


Fig. 15. (color online) The TPS of the quasiperiodic plate with N=21 (blue) and N=34 (red), and a pure Silicon plate (dashed black); L/D =0.5.

when the ratio L/D is larger than 0.68, as shown in Fig. 14(e). In this figure, one can see that there is only one forbidden gap in both the periodic and quasiperiodic systems. The gap extends from 1350 (1570) up to 1970 (2136) kHz for the periodic (quasiperiodic) system, respectively. It means that the difference of band gaps between quasiperiodic and periodic systems basically disappears as the lattice spacing decreases.

Furthermore, in order to investigate the finite size effect on band gaps, we calculate the TPS for $N = 21$ and 34 for $L / D = 0.5$. The results are shown in Fig. 15, which tells us that the number of splitting band gaps in quasiperiodic superlattices does not increase with the addition of the layer number of Fibonacci sequences. The result is quite different from those in the quasiperiodic photonic and phononic crystals of the bulk waves [40-41].

Lastly we study the influence of the thickness of sublattices on the band gap. We calculate the TPS for the cases of $d_A / D = 0.618$ and $d_B / D = 0.618$. The results are shown in Fig. 16. There is only one band gap in the structure of $d_A / D = 0.618$ ($d_B / D = 0.382$). The gap extends from the frequency of 1565 up to 1790 kHz. However, four band gaps are observed in the systems with $d_A / D = 0.382$ ($d_B / D = 0.618$). The four bands are from 950 up to 1130 kHz, from 1310 up to 1550 kHz, from 1780 up to 2030 kHz, and from 2250 up to 2530 kHz, respectively. One can easily find that the material (Tungsten) with larger values of the elastic constant and mass density influences the band gap more than the material (Silicon) with smaller values of the elastic constant and mass density.

In conclusion, we have examined the band gap structures of Lamb waves in the 1D quasiperiodic composite thin plates by calculating the TPS from the FEM. The band gap structures of the Lamb waves are quite different from those of bulk waves. Specifically, the

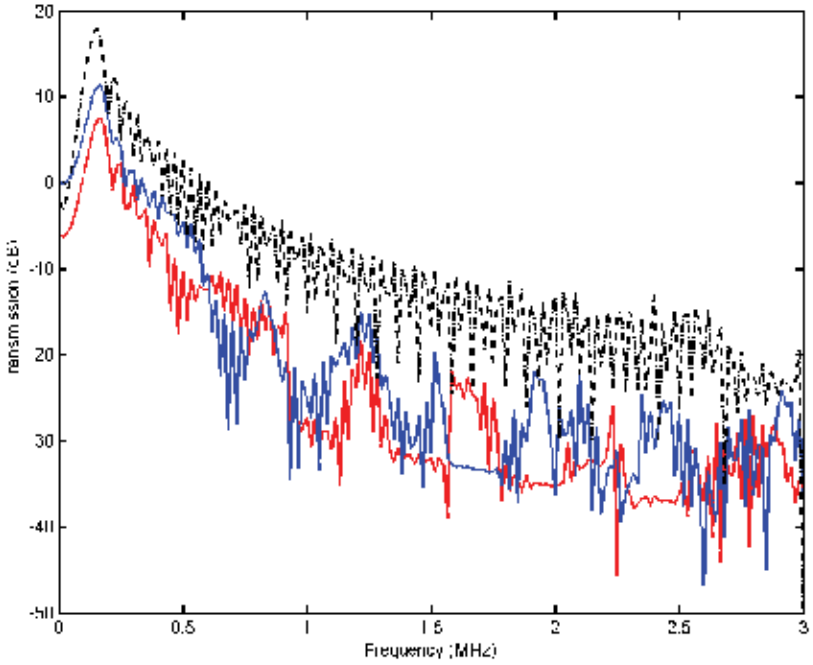


Fig. 16. (color online) The TPS of the quasiperiodic plate $d_A / D = 0.618$ (blue) and $d_B / D = 0.618$ (red), and a pure Silicon plate (dashed black); L/D =0.5.

number of splitting band gaps depends strongly on the values of L/D owing to resonance of the coupling of the longitudinal and transversal strain components at the plate boundaries. However, the split of band gaps is independent of the layer number of Fibonacci sequences. Moreover, we have found that the structure of the band gaps depends very sensitively on the thickness ratio of the sublattices A and B in the quasiperiodic structures which might find applications in nondestructive diagnosis.

5. Acoustic wave behavior in silicon-based 1D phononic crystal plates

In this section, we employ HRA to study the propagation and transmission of acoustic waves in silicon-based 1D phononic crystal plates without/with substrate. We also employ HRA to study quasiperiodic systems such as Generalized Fibonacci Systems and Double-period System, and the results show that some new phononic band gaps form in quasiperiodic systems, which hold the potential in the application of acoustic filters and couplers.

In Fig. 17, the parameters of finite element models for both TRA and HRA are set to be: the plate thickness $H = 2$ mm, the distance from exciting source to the left edge of plate (also the distance from the receiver to the right edge of plate) $L_1 = 15$ cm, the length of superlattice $S = 20$ cm, the number of finite elements per meter $N = 10000$ m^{-1}, the distance between exciting source and receiver $L_2 = 30$ cm, the width of the exciting source region (source function is Guassian function) $\delta = 4$ mm. In fact, the theoretical models for TRA and HRA are analogous to laser-generated Lamb wave system and piezoelectricity-generated Lamb wave system, respectively.

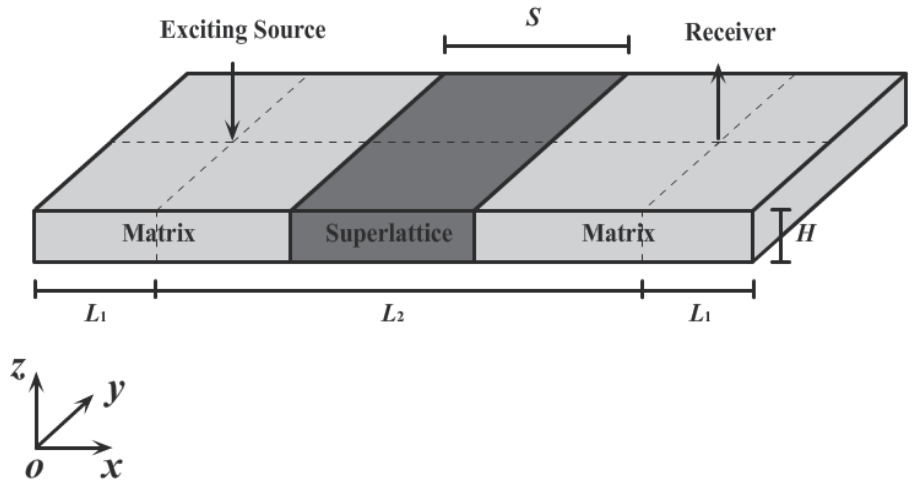


Fig. 17. The plate geometry in the finite element models for both TRA and HRA method; the upper surface is located at $z = H$.

We choose two cases (without/with substrate and different quasiperiodic systems) to investigate the acoustic wave behavior in phononic crystal plates.

For the plate without substrate, we set filling factor $f = 0.2$, lattice constant $a = 2$ mm, plate thickness $H = 2$ mm, without substrate. The number of inclusions is 100 and all the inclusions are embedded periodically in the middle of plate.

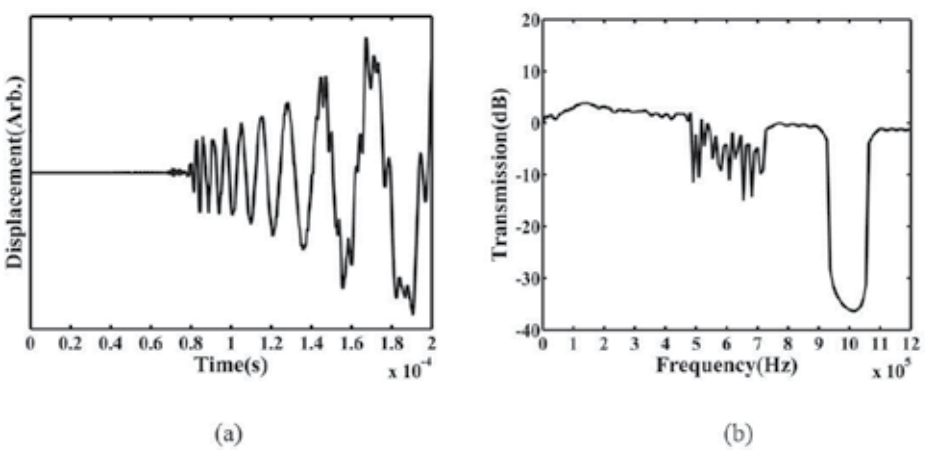


Fig. 18. (a) The transient vertical displacement at the upper surface of phononic crystal plate without substrate, calculated by TRA method; (b) Normalized transmitted power spectrum for phononic crystal plate without substrate.

In TRA, as seen in Fig. 18(a), the transient vertical displacement at the upper surface of phononic crystal plate is shown when the time ranges from 0 to 200 μs. Transforming the vertical displacement from time domain to frequency domain and normalizing by the transmitted power spectrum of homogeneous plate, we can obtain the normalized transmitted power spectrum of phononic crystal plate with periodic superlattice, as shown

in Fig. 18(b), and an obvious band gap is observed in the range from 0.9512 to 1.047 MHz, which means the elastic wave located in this gap is extremely attenuated. Applying the Super-cell PWE or HRA, we recalculate the band structure and normalized transmitted power spectrum, respectively for comparison and the data are shown in Fig. 19.

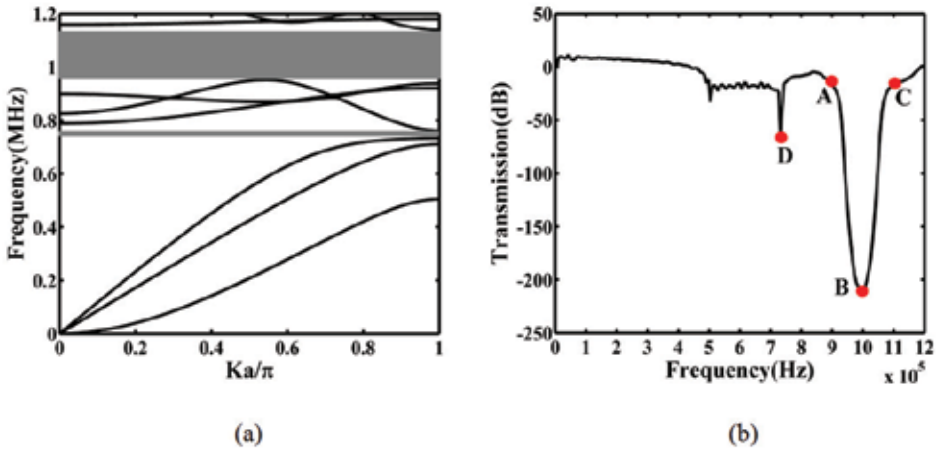


Fig. 19. (a) Dispersion curves of Lamb wave modes for phononic crystal plate without substrate, calculated by Super-cell PWE; (b) Normalized transmitted power spectrum for phononic crystal plate without substrate, calculated by HRA method.

From both Fig. 19(a) and 19(b), we can see a main band gap located around 1 MHz (0.9511~1.1300 MHz in Fig. 19(a); 0.9510~1.0560 MHz in Fig. 19(b)), which accords with the Fig. 18(b). Note that there exists a very narrow band gap in low frequency zone as shown in Fig. 19(a) (0.7332 MHz~0.762 MHz), or the D point (0.7335 MHz) in Fig. 19(b). Therefore, the result of HRA is more consistent with Super-cell PWE than of TRA, and importantly the HRA method is more efficient in calculations of not only normalized transmitted power spectrum but also space distribution of elastic wave field for the reason mentioned above. Hereon we choose three points (A: 0.9 MHz, B: 1 MHz, C: 1.1 MHz) in Fig. 19(b) for the study of propagation of Lamb waves under different frequency loads (inside/outside the band gap).

As seen from Fig. 20, the displacement fields under different frequency loads are quite different. In Fig. 20(b), the load frequency locates inside the band gap and the displacement field seems like being blocked by the superlattice, in which the periodic structure forbids the propagation of elastic waves along the plate. However, when the load frequency locates outside the band gap in Fig. 20(a) and 20(c), the elastic waves propagate without any obvious attenuation.

Then, we add an extra substrate to the established model. The thickness of substrate is set to be 0.2 mm. Applying the Super-cell PWE and HRA, we can obtain the dispersion curves of Lamb wave modes and normalized transmitted power spectrum, respectively, as shown in Fig. 21, in which the first band gap exists in low frequency zone (0.7413~0.7767 MHz in Fig. 21(a); 0.7520~0.7730 MHz in Fig. 21(b)) and the main band gap (second band gap) locates at high frequency zone (0.9852~1.1240 MHz in Fig. 21(a); 0.9853~1.0580 MHz in Fig. 21(b)). Comparing Fig. 21 with Fig. 19, one can observe that the first band gap width in the plate with substrate is larger than that of the plate without substrate and main band gap (the

second band gap) width is narrowed and shifted towards high frequency zone, which accord with previous works [4,23,42].

In addition to the periodic systems, we adopt the HRA to study the quasiperiodic systems. The normalized transmitted power spectra are calculated for phononic crystal plates with the above three quasiperiodic systems, as shown in Fig. 22(a)-(c), in which the normalized transmitted power spectrum of periodic system is also plotted for comparison.

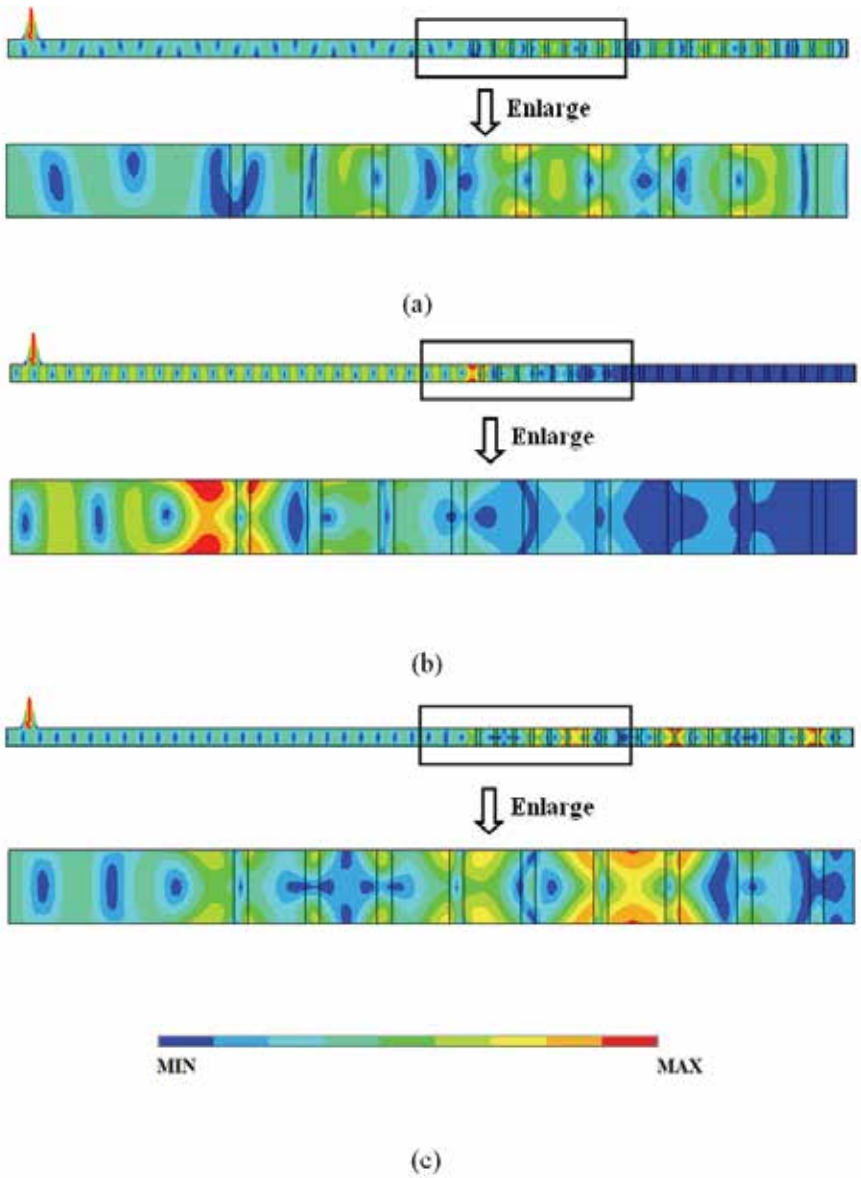


Fig. 20. The displacement fields at the frequency loads of 0.9 MHz (A point in Fig. 19(b)) (a), 1 MHz (B point in Fig. 19(b)) (b) and 1.1 MHz (C point in Fig. 19(b)) (c), respectively. Corresponding plot in each figure is enlarged.

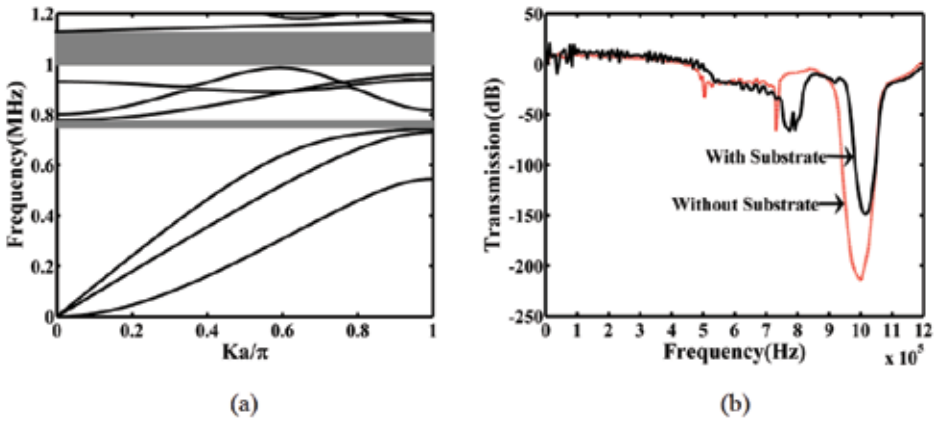


Fig. 21. (a) Dispersion curves of Lamb wave modes for phononic crystal plate with substrate, calculated by Super-cell PWE; (b) Normalized transmitted power spectra for phononic crystal plates both with and without substrate (substrate thickness: 0.2 mm), calculated by HRA method.

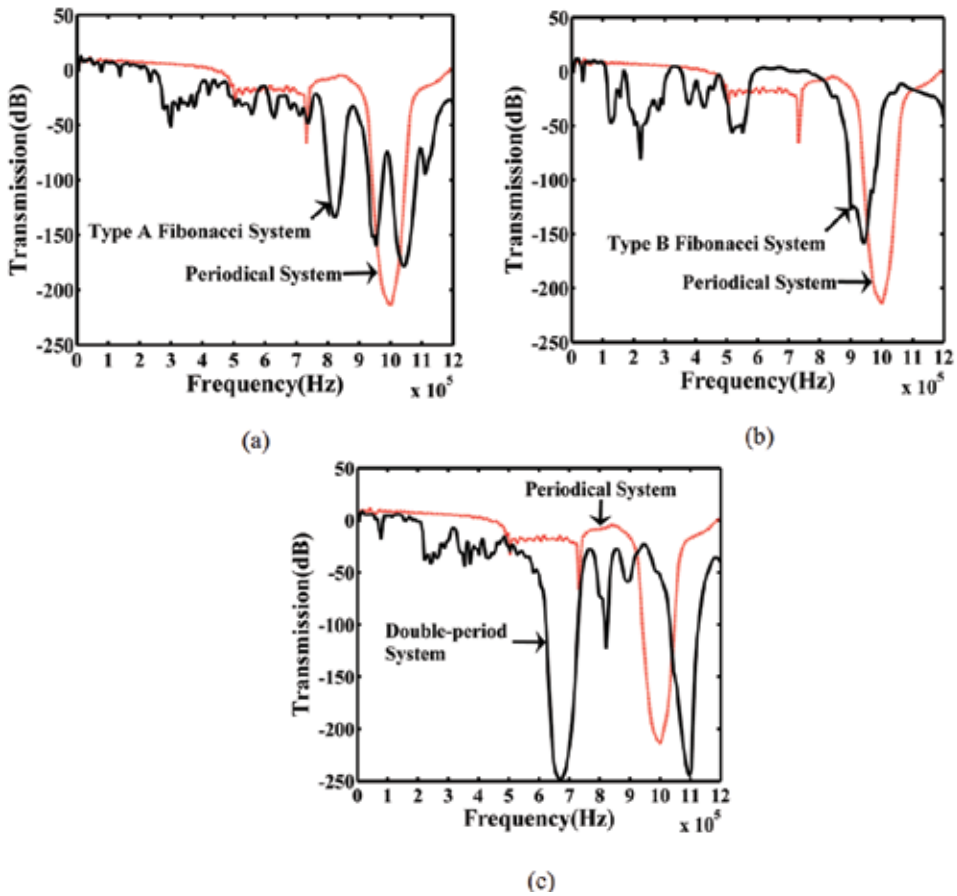


Fig. 22. Normalized transmitted power spectra for Type A Fibonacci System (a), Type B Fibonacci System (b) and Double-period System (c), respectively.

As shown in Fig. 22(a), for Type A Fibonacci System, a new band gap is opened in low frequency zone (0.7925~0.8622 MHz) and the main band gap (corresponding to the one in the periodic system) splits into two separated sub-band gaps (0.9282~0.9667 MHz and 0.9942~1.0730 MHz). In addition, an obvious attenuation is observed in lower frequency zone (0.276~0.384 MHz).

For Type B Fibonacci System, the main band gap shifts to the low frequency zone (0.8915~0.9978 MHz) with its gap width almost unchanged, as shown in Fig. 22(b), and three new band gaps form in lower frequency zone (0.114~0.162 MHz, 0.192~0.288 MHz and 0.498~0.57 MHz).

For Double-period System, the main band gap shifts to high frequency zone (1.0180~1.1340 MHz) and four new band gaps are opened in the low frequency zone (0.216~0.294 MHz, 0.336~0.468 MHz, 0.6238~0.7265 MHz and 0.792~0.84 MHz), as shown in Fig. 22(c).

From the above-mentioned information, we convincingly demonstrate the band gap distribution of quasiperiodic systems is more complicated and meaningful than of periodic systems and the reason is supposed to be that quasiperiodicity unlike periodicity can provide more than one reciprocal lattices.

6. Band gaps of plate-mode waves in 1D piezoelectric composite plates without/with substrates

As well known, the ceramic material will have the piezoelectricity only after it is polarized. In convenience, we define the non-polarized PZT-5H ceramic as the non-piezoelectric material, which has the same elastic constants as the polarized PZT-5H. Fig. 23 provides five schematic representations of the plate-mode waves for non-polarization, *x*-polarization with OC, *x*-polarization with SC, *z*-polarization with OC, and *z*-polarization with SC, respectively. The first band gaps (FBG) widths shown by the gray area in Fig. 23(a)-(e) are 2.088, 2.072, 2.368, 2.368, and 2.6 MHz, respectively. On the whole, the FBG are always broadened by polarizing piezoelectric ceramics at the same values of f and h/D. Comparing the Fig. 23(b) and (c) (or Fig. 23(d) and (e)), the FBG width with SC is larger than that with OC for the same polarized direction, whereas the FBG width of *z*-polarization with SC is the largest. In our example, the FBG width of *z*-polarization with OC is equal to the FBG width of *x*-polarization with SC, which means the *z*-polarized PZT-5H ceramics is easy to produce a larger FBG width.

The V-PWE method is applied to calculate the dispersion curves of Lamb wave propagating in the *x*-direction when the existence of uniform substrate. Since the substrate affects the width and starting frequency of the PC layer, the thickness of the substrate will be an important parameter of the system. Meanwhile, the filling fraction f is another critical parameter that affects the formation, width and starting frequency of the FBG [43,44]. Fig. 24 (a) and (b) display the dependence of the FBG widths and starting frequencies with the filling fraction f and the ratio of h_2 to h_1 at h_1 =0.8mm, D =2mm with OC when the PC layer is coated on an epoxy substrate. As shown in Fig. 24(a), the FBG width increases progressively with the increase of the value of f at a certain value of h_2 / h_1 until a critical value then decreases and the width decreases gradually with the increase of the value of h_2 / h_1 at a certain value of f . The FBG width takes the maximum value when there is no substrate, and decreases with the increase of the substrate's thickness at any values of f . The FBG width takes a larger value when f and h_2 / h_1 take values in the domain 0.45-0.65

and 0-0.8, respectively. This domain is useful in the engineering field. The FBG width decreases slowly when h_2 / h_1 takes values from 0 to 0.4 (Δh_{slow}) and decreases rapidly when h_2 / h_1 takes values from 0.4 to 0.80 (Δh_{rapid}) as f takes values from 0.45 to 0.65(Δf). As shown in Fig. 25(b), the FBG starting frequency decreases gradually with the increase of the value of f at a certain value of h_2 / h_1 until a critical value then increases. The starting frequency increases progressively with the increase of the value of h_2 / h_1 at a certain value of f until a critical value then decreases, but the change of the starting frequency is small. On the whole, the epoxy substrate reduces the FBG width obviously and has little influence on the FBG starting frequency.

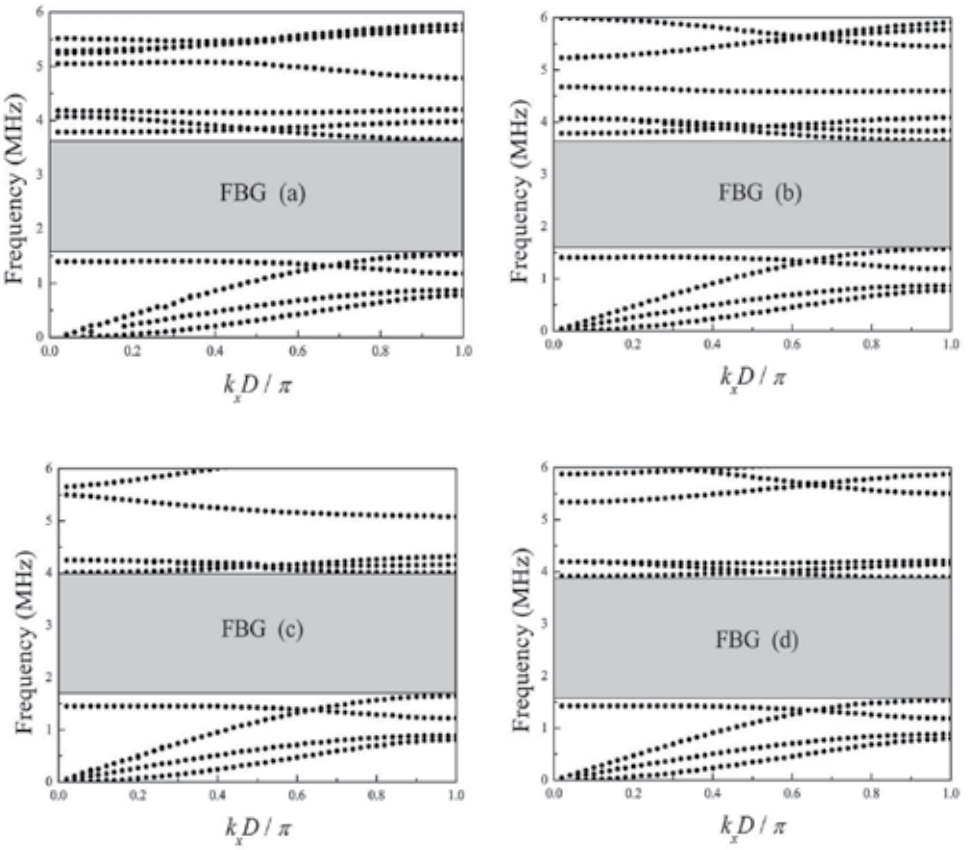


Fig. 23. The 1D plate-mode waves for different polarizations under different boundary conditions with $f = 0.5$ and h / D =0.8 (D=2mm): (a) Non-polarization, (b) x-polarization with OC, (c) x-polarization with SC, (d) z-polarization with OC, and (e) z-polarization with SC.

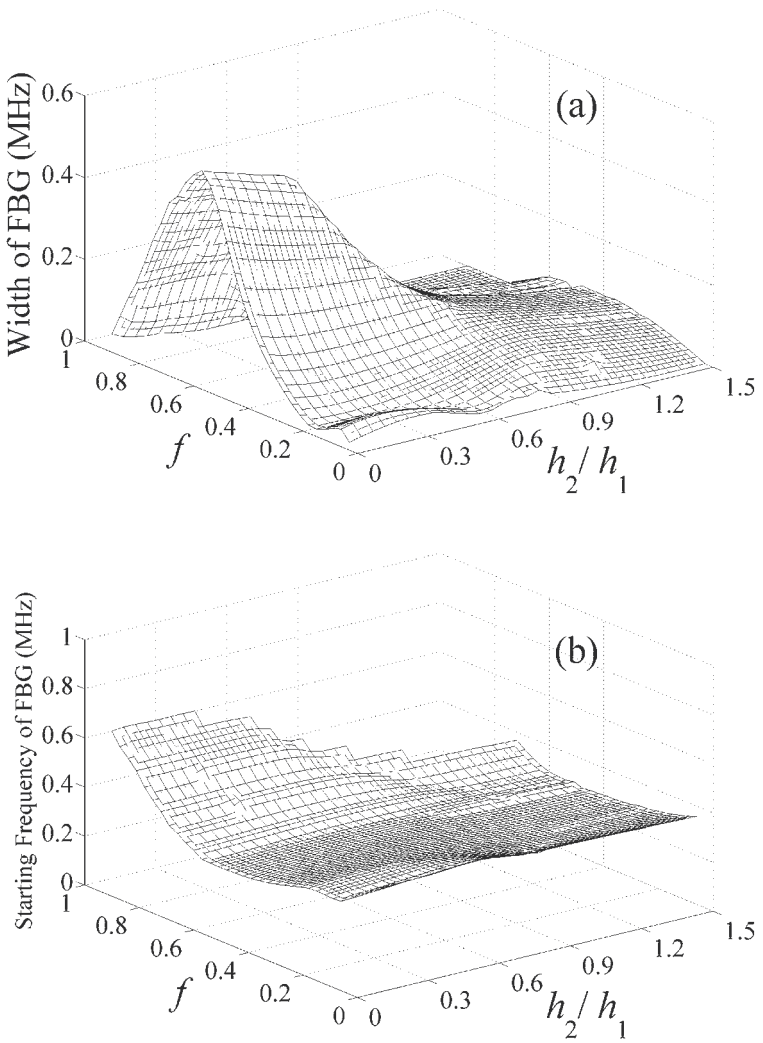


Fig. 24. The FBG widths (a) and starting frequencies (b) versus f and h_2 / h_1 (h_1 =0.8mm, D =2mm) with OC coated on epoxy substrate.

7. Conclusions

In this chapter, we first examine the band structures of lower-order Lamb wave modes propagating in the 1D periodic composite thin plate based on the PWE for infinitely long periodic systems and have calculated the TPS for finite systems by using the FE method. As shown, the TPS through a superlattice with ten periods has prominent dips at frequencies corresponding to the gaps in band structure. A crucial parameter, namely, the ratio of L/D, was discussed, and the value of the ratio of L/D was emerging as critical parameters in determining the existence of band gaps for the Lamb waves in the periodic structures. Thus, we can achieve the needed width of band gap for Lamb wave by varying the thickness of

plate. Then, we study the substrate effect on the band gaps of lower-order Lamb waves propagating in thin plate with 1D phononic crystal layer coated on uniform substrate. The results show that when the substrate is hard, the influences on band gap are significant, and the band gaps disappear rapidly as the substrate becomes thicker. However, when the substrate is soft, the depth of band gaps becomes larger as the thickness of substrate increases. A virtual plane wave expansion method is developed to calculate the dispersion curves of Lamb wave. The locations and widths of band gaps on the dispersion curves are in good agreement with the results from the transmitted power spectra by FEM.

The band gap structure of Lamb waves in the 1D quasiperiodic composite thin plate is also studied by calculating the TPS from the FEM. The band gap structures of the Lamb waves are quite different from those of bulk waves. Specifically, the number of splitting band gaps depends strongly on the values of L/D owing to resonance of the coupling of the longitudinal and transversal strain components at the plate boundaries. However, the split of band gaps is independent of the layer number of Fibonacci sequences. Moreover, we have found that the structure of the band gaps depends very sensitively on the thickness ratio of the sublattices A and B in the quasiperiodic structures which might find applications in nondestructive diagnosis.

We have promoted an efficient HRA method to investigate the acoustic wave behavior in silicon-based 1D phononic crystal plates. The HRA method can not only save much time in the calculation of transmitted power spectrum but also acquire information of the displacement field under different frequency loads at the same time. Applying HRA and supercell PWE, we have studied the periodic structures both without and with substrate. From the displacement field map, we find that the elastic wave is completely blocked by the superlattice when the load frequency is inside the acoustic band gap. After introducing different kinds of quasiperiodic structures, we studied the normalized transmitted power spectra in details and find out that the original main band gap in periodic structure may split or shift to low or high frequency zones in different quasiperiodic structures. Furthermore, new band gaps in low frequency zone may be opened which provide potential application in the field of wave filtering as well as sound isolation.

Finally, we study the band gaps of plate-mode waves in 1D piezoelectric composite plates without/with substrates. We found that the FBG is always broadened by polarizing piezoelectric ceramics, and the FBG widths with SC are always larger than that with OC for the same polarization. The FBG width decreases gradually as the substrate's thickness increases and the FBG starting frequency increases progressively as the thickness increases on the whole. Our researches show that it is possible to control the width and starting frequency of the FBG in the engineering according to need by choosing suitable values of the substrate's thickness, the filling fraction with different electrical boundary conditions.

8. References

[1] J. J. Chen, K. W. Zhang, J. Gao, and J. C. Cheng, Phys. Rev. B 73, 094307 (2006).
[2] J. Gao, J. C. Cheng, and B. W. Li, Appl. Phys. Lett. 90, 111908 (2007).
[3] J. H. Sun and T. T. Wu, Phys. Rev. B 74, 174305 (2006).
[4] J. O. Vasseur, P. A. Deymier, B. Djafari-Rouhani, Y. Pennec, and A. -C. Hladky-Hennion, Phys. Rev. B 77, 085415 (2008).

[5] S. Mohammadi, A. A. Eftekhar, A. Khelif, W. D. Hunt, and A. Adibi, Appl. Phys. Lett. 92, 221905 (2008).
[6] S. Mohammadi, A. A. Eftekhar, W. D. Hunt, and A. Adibi, Appl. Phys. Lett. 94, 051906 (2009).
[7] C. J. Rupp, M. L. Dunn, and K. Maute, Appl. Phys. Lett. 96, 111902 (2010).
[8] J. Gao, X. Y. Zou, and J. C. Cheng, Appl. Phys. Lett. 92, 023510 (2008).
[9] X. F. Zhu, T. Xu, S. C. Liu, and J. C. Cheng, J. Appl. Phys. 106, 104901 (2009).
[10] X.-Y. Zou, B. Liang, Q. Chen, and J.-C. Cheng, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 56, 361 (2009).
[11] B. A. Auld, Y. A. Shui, and Y. Wang, J. Phys. (Paris) 45, 159 (1984).
[12] B. A. Auld and Y. Wang, Proc.-IEEE Ultrason. Symp. 528 (1984).
[13] A. Alippi, F. Craciun, and E. Molinari, Appl. Phys. Lett. 53, 1806 (1988).
[14] A. Alippi, F. Craciun, and E. Molinari, J. Appl. Phys. 66, 2828 (1989).
[15] S. G. Joshi and Y. Jin, J. Appl. Phys. 69, 8018 (1991).
[16] M. Wilm, S. Ballandras, V. Laude, and T. Pastureaud, J. Acoust. Soc. Am. 112, 943 (2002).
[17] S. Zhang and J. C. Cheng, Phys. Rev. B 68, 245101 (2003).
[18] F. Moser, L. J. Jacobs, and J. M. Qu, NDT & E Int. 32, 225 (1999).
[19] K. J. Bathe, Finite Element Procedures, Prentice-Hall, Englewood Cliffs, NJ, 1996.
[20] Y. Gu, J. Liu and Y. Du, Engineering Mechanics 24, 12 (2007) (Chinese).
[21] X. F. Zhu, S. C. Liu, T. Xu, T. H. Wang, and J. C. Cheng, Chin. Phys. B 19, 4 (2010).
[22] Y. Tanaka, Y. Tomoyasu, and S. I. Tamura, Phys. Rev. B 62, 7387 (2000).
[23] Z. L. Hou and B. M. Assouar, Phys. Lett. A 372, 2091 (2008).
[24] Y. Zhang and H. Zhao, Phys. Rev. E 66, 026106 (2002).
[25] J. C. Cheng, S. Y. Zhang, and L. Wu, Appl. Phys. A: Mater. Sci. Process. 61, 311 (1995).
[26] Y. J. Cao and Y. Xu, Actra Physica Sinica 57, 3620 (2008).
[27] E. L. Albuquerque and M. G. Cottam, Physics Reports 376, 225 (2003).
[28] W. A. Smith and B. A. Auld, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 38, 40 (1991).
[29] M. S. Kushwaha, Appl. Phys. Lett. 70, 3218 (1997).
[30] J. C. Cheng and S. Y. Zhang, Appl. Phys. Lett. 74, 2087 (1999).
[31] S. Linden, J. Kuhl, and H. Giessen, Phys. Rev. Lett. 86, 4688 (2001).
[32] S. Ostlund, R. Pandit, D. Rand, H. J. Schellnhuber, and E. D. Siggia, Phys. Rev. Lett. 50, 1873 (1983).
[33] D. Shechtman, I. Blech, D. Gratias, and J. W. Cahn, Phys. Rev. Lett. 53, 1951 (1984).
[34] R. Merlin, K. Bajema, R. Clarke, F. -Y. Juang, and P. K. Bhattacharya, Phys. Rev. Lett. 55, 1768 (1985).
[35] J. Todd, R. Merlin, R. Clarke, K. M. Mohanty, and J. D. Axe, Phys. Rev. Lett. 57, 1157 (1986).
[36] C. Wang and R. A. Barrio, Phys. Rev. Lett. 61, 191 (1988).
[37] J. -P. Desideri, L. Macon, and D. Sornette, Phys. Rev. Lett. 63, 390 (1989).
[38] M. W. C. Dharma-Wardana, A. H. MacDonald, D. J. Lockwood, J. -M. Baribeau, and D. C. Houghton, Phys. Rev. Lett. 58, 17 (1987).
[39] S. Tamura and J. P. Wolfe, Phys. Rev. B 36, 6 (1987).
[40] Y. E. Hassouani, H. Aynaou, E. H. E. Boudouti, B. Djafari-Rouhani, A. Akjouj, and V. R. Velasco, Phys. Rev. B 74, 035314 (1990).
[41] J. Q. You, Q. B. Yang, and J. R. Yan, Phys. Rev. B 41, 11 (1990).
[42] Z. Hou and B. M. Assouar, J. Phys. D: Appl. Phys. 42, 0850103 (2009).
[43] X.-Y. Zou, Q. Chen, B. Liang, and J.-C. Cheng, Smart Mater. Struct. 17, 015008 (2008)
[44] X.-Y. Zou, Q. Chen, and J.-C. Cheng, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 54, 1430 (2007).

7

Frequency-Domain Numerical Modelling of Visco-Acoustic Waves with Finite-Difference and Finite-Element Discontinuous Galerkin Methods

Romain Brossier[1,2], Vincent Etienne[1], Stéphane Operto[1] and Jean Virieux[2]
[1]*Geoazur - CNRS - UNSA - IRD – OCA*
[2]*LGIT - University Josepth Fourier*
France

1. Introduction

Seismic exploration is one of the main geophysical methods to extract quantitative inferences about the Earth's interior at different scales from the recording of seismic waves near the surface. Main applications are civil engineering for cavity detection and landslide characterization, site effect modelling for seismic hazard, CO_2 sequestration and nuclear-waste storage, oil and gas exploration, and fundamental understanding of geodynamical processes. Acoustic or elastic waves are emitted either by controlled sources or natural sources (i.e., earthquakes). Interactions of seismic waves with the heterogeneities of the subsurface provide indirect measurements of the physical properties of the subsurface which govern the propagation of elastic waves (compressional and shear wave speeds, density, attenuation, anisotropy). Quantitative inference of the physical properties of the subsurface from the recordings of seismic waves at receiver positions is the so-called seismic inverse problem that can be recast in the framework of local numerical optimization. The most complete seismic inversion method, the so-called full waveform inversion (Virieux & Operto (2009) for a review), aims to exploit the full information content of seismic data by minimization of the misfit between the full seismic wavefield and the modelled one. The theoretical resolution of full waveform inversion is half the propagated wavelength. In full waveform inversion, the full seismic wavefield is generally modelled with volumetric methods that rely on the discretization of the wave equation (finite difference, finite element, finite volume methods).

In the regime of small deformations associated with seismic wave propagation, the subsurface can be represented by a linear elastic solid parameterized by twenty-one elastic constants and the density in the framework of the constitutive Hooke's law. If the subsurface is assumed isotropic, the elastic constants reduce to two independent parameters, the Lamé parameters, which depend on the compressional (P) and the shear (S) wave speeds. In marine environment, the P wave speed has most of the time a dominant footprint in the seismic wavefield, in particular, on the hydrophone component which records the pressure wavefield. The dominant footprint of the P wave speed on the seismic

wavefield has prompted many authors to develop and apply seismic modelling and inversion under the acoustic approximation, either in the time domain or in the frequency domain.

This study focuses on frequency-domain modelling of acoustic waves as a tool to perform seismic imaging in the acoustic approximation. In the frequency-domain, wave modelling reduces to the resolution of a complex-valued large and sparse system of linear equations for each frequency, the solution of which is the monochromatic wavefield and the right-hand side (r.h.s) is the source. Two key issues in frequency-domain wave modelling concern the linear algebra technique used to solve the linear system and the numerical method used for the discretization of the wave equation. The linear system can be solved with Gauss elimination techniques based on sparse direct solver (e.g., Duff et al.; 1986), Krylov-subspace iterative methods (e.g., Saad; 2003) or hybrid direct/iterative method and domain decomposition techniques (e.g., Smith et al.; 1996). In the framework of seismic imaging applications which involve a large number of seismic sources (i.e., r.h.s), one motivation behind the frequency-domain formulation of acoustic wave modelling has been to develop efficient approaches for multi-r.h.s modelling based on sparse direct solvers (Marfurt; 1984). A sparse direct solver performs first a LU decomposition of the matrix which is independent of the source followed by forward and backward substitutions for each source to get the solution (Duff et al.; 1986). This strategy has been shown to be efficient for 2D applications of acoustic full waveform inversion on realistic synthetic and real data case studies (Virieux & Operto; 2009). Two drawbacks of the direct-solver approach are the memory requirement of the LU decomposition resulting from the fill-in of the matrix during the LU decomposition (namely, the additional non zero coefficients introduced during the elimination process) and the limited scalability of the LU decomposition on large-scale distributed-memory platforms. It has been shown however that large-scale 2D acoustic problems involving several millions of unknowns can be efficiently tackled thanks to recent development of high-performance parallel solvers (e.g., MUMPS team; 2009), while 3D acoustic case studies remain limited to computational domains involving few millions of unknowns (Operto et al.; 2007). An alternative approach to solve the time-harmonic wave equation is based on Krylov-subspace iterative solvers (Riyanti et al.; 2006; Plessix; 2007; Riyanti et al.; 2007). Iterative solvers are significantly less memory demanding than direct solvers but the computational time linearly increases with the number of r.h.s. Moreover, the impedance matrix, which results from the discretization of the wave equation, is indefinite (the real eigenvalues change sign), and therefore ill-conditioned. Designing efficient pre-conditioner for the Helmholtz equation is currently an active field of research (Erlangga & Nabben; 2008). Efficient preconditioners based on one cycle of multigrid applied to the damped wave equation have been developed and leads to a linear increase of the number of iterations with frequency when the grid interval is adapted to the frequency (Erlangga et al.; 2006). This makes the time complexity of the iterative approaches to be $\mathcal{O}(N^4)$, where N denotes the dimension of the 3D N^3 cubic grid. Intermediate approaches between the direct and iterative approaches are based on domain decomposition methods and hybrid direct/iterative solvers. In the hybrid approach, the iterative solver is used to solve a reduced system for interface unknowns shared by adjacent subdomains while the sparse direct solver is used to factorize local impedance matrices assembled on each subdomains during a preprocessing step (Haidar; 2008; Sourbier et al.; 2008). A short review of the time and memory complexities of the direct, iterative and hybrid approaches is provided in Virieux et al. (2009).

The second issue concerns the numerical scheme used to discretize the wave equation. Most of the methods that have been developed for seismic acoustic wave modelling in the frequency domain rely on the finite difference (FD) method. This can be justified by the fact that, in many geological environments such as offshore sedimentary basins, the subsurface of the earth can be viewed as a weakly-contrasted medium at the scale of the seismic wavelengths, for which FD methods on uniform grid provide the best compromise between accuracy and computational efficiency. In the FD time-domain method, high-order accurate stencils are generally designed to achieve the best trade-off between accuracy and computational efficiency (Dablain; 1986). However, direct-solver approaches in frequency-domain modelling prevent the use of such high-order accurate stencils because their large spatial support will lead to a prohibitive fill-in of the matrix during the LU decomposition (Stekl & Pratt; 1998; Hustedt et al.; 2004). Another discretization strategy, referred to as the mixed-grid approach, has been therefore developed to perform frequency-domain modelling with direct solver: it consists of the linear combination of second-order accurate stencils built on different rotated coordinate systems combined with an anti-lumped mass strategy, where the mass term is spatially distributed over the different nodes of the stencil (Jo et al.; 1996). The combination of these two tricks allows one to design both compact and accurate stencils in terms of numerical anisotropy and dispersion.

Sharp boundaries of arbitrary geometry such as the air-solid interface at the free surface are often discretized along staircase boundaries of the FD grid, although embedded boundary representation has been proposed (Lombard & Piraux; 2004; Lombard et al.; 2008; Mattsson et al.; 2009), and require dense grid meshing for accurate representation of the medium. The lack of flexibility to adapt the grid interval to local wavelengths, although some attempts have been performed in this direction (e.g., Pitarka; 1999; Taflove & Hagness; 2000), is another drawback of FD methods. These two limitations have prompted some authors to develop finite-element methods in the time domain for seismic wave modelling on unstructured meshes. The most popular one is the high-order spectral element method (Seriani & Priolo; 1994; Priolo et al.; 1994; Faccioli et al.; 1997) that has been popularized in the field of global scale seismology by Komatitsch and Vilotte (1998); Chaljub et al. (2007). A key feature of the spectral element method is the combined use of Lagrange interpolants and Gauss-Lobatto-Legendre quadrature that makes the mass matrix diagonal and, therefore, the numerical scheme explicit in time-marching algorithms, and allows for spectral convergence with high approximation orders (Komatitsch & Vilotte; 1998). The selected quadrature formulation leads to quadrangle (2D) and hexahedral (3D) meshes, which strongly limit the geometrical flexibility of the discretization. Alternatively, discontinuous form of the finite-element method, the so-called discontinuous Galerkin (DG) method (Hesthaven & Warburton; 2008), popularized in the field of seismology by Kaser, Dumbser and co-workers (e.g., Dumbser & Käser; 2006) has been developed. In the DG method, the numerical scheme is strictly kept local by duplicating variables located at nodes shared by neighboring cells. Consistency between the multiply defined variables is ensured by consistent estimation of numerical fluxes at the interface between two elements. Numerical fluxes at the interface are introduced in the weak form of the wave equation by means of integration by part followed by application of the Gauss's theorem. Key advantages of the DG method compared to the spectral element method is its capacity of considering triangular (2D) and tetrahedral (3D) non-conform meshes. Moreover, the uncoupling of the elements provides a higher level of flexibility to locally adapt the size of

the elements (*h* adaptivity) and the interpolation orders within each element (*p* adaptivity) because neighboring cells exchange information across interfaces only. Moreover, the DG method provides a suitable framework to implement any kind of physical boundary conditions involving possible discontinuity at the interface between elements. One example of application which takes fully advantage of the discontinuous nature of the DG method is the modelling of the rupture dynamics (BenJemaa et al.; 2007, 2009; de la Puente et al.; 2009). The dramatic increase of the total number of degrees of freedom compared to standard finite-element methods, that results from the uncoupling of the elements, might prevent an efficient use of DG methods. This is especially penalizing for frequency-domain methods based on sparse direct solver where the computational cost scales with the size of the matrix N in $\mathcal{O}(N^6)$ for 3D problems. The increase of the size of the matrix should however be balanced by the fact the DG schemes are more local and sparser than FEM ones (Hesthaven & Warburton; 2008), which makes smaller the numerical bandwidth of the matrix to be factorized.

When a zero interpolation order is used in cells (piecewise constant solution), the DG method reduces to the finite volume method (LeVeque; 2002). The DG method based on high-interpolation orders has been mainly developed in the time domain for the elastodynamic equations (e.g., Dumbser & Käser; 2006). Implementation of the DG method in the frequency domain has been presented by Dolean et al. (2007, 2008) for the time-harmonic Maxwell equations and a domain decomposition method has been used to solve the linear system resulting from the discretization of the Maxwell equations. A parsimonious finite volume method on equilateral triangular mesh has been presented by Brossier et al. (2008) to solve the 2D P-SV elastodynamic equations in the frequency domain. The finite-volume approach of Brossier et al. (2008) has been extended to low-order DG method on unstructured triangular meshes in Brossier (2009).

We propose a review of these two quite different numerical methods, the mixed-grid FD method with simple regular-grid meshing and the DG method with dense unstructured meshing, when solving frequency-domain visco-acoustic wave propagation with sparse direct solver in different fields of application. After a short review of the time-harmonic visco-acoustic wave equation, we first review the mixed-grid FD method for 3D modelling. We first discuss the accuracy of the scheme which strongly relies on the optimization procedure designed to minimize the numerical dispersion and anisotropy. Some key features of the FD method such as the absorbing and free-surface boundary conditions and the source excitation on coarse FD grids are reviewed. Then, we present updated numerical experiments performed with the last release of the massively-parallel sparse direct solver MUMPS (Amestoy et al.; 2006). We first assess heuristically the memory complexity and the scalability of the LU factorization. Second, we present simulations in two realistic synthetic models representative of oil exploration targets. We assess the accuracy of the solutions and the computational efficiency of the mixed-grid FD frequency-domain method against that of a conventional FD time-domain method. In the second part of the study, we review the DG frequency-domain method applied to the first-order acoustic wave equation for pressure and particle velocities. After a review of the spatial discretization, we discuss the impact of the order of the interpolating Lagrange polynomials on the computational cost of the frequency-domain DG method and we present 2D numerical experiments on unstructured triangular meshes to highlight the fields of application where the DG method should outperform the FD method.

Although the numerical methods presented in this study were originally developed for seismic applications, they can provide a useful framework for other fields of application such as computational ocean acoustics (Jensen et al.; 1994) and electrodynamics (Taflove & Hagness; 2000).

2. Frequency-domain acoustic wave equation

Following standard Fourier transformation convention, the 3D acoustic first-order velocity-pressure system can be written in the frequency domain as

$$\begin{aligned}
-i\omega p(x,y,z,\omega) &= \kappa(x,y,z)\left(\frac{\partial v_x(x,y,z,\omega)}{\partial x}+\frac{\partial v_y(x,y,z,\omega)}{\partial y}+\frac{\partial v_z(x,y,z,\omega)}{\partial z}\right)\\
-i\omega v_x(x,y,z,\omega) &= b(x,y,z)\cdot\frac{\partial p(x,y,z,\omega)}{\partial x}+f_x(x,y,z,\omega)\\
-i\omega v_y(x,y,z,\omega) &= b(x,y,z)\cdot\frac{\partial p(x,y,z,\omega)}{\partial y}+f_y(x,y,z,\omega)\\
-i\omega v_z(x,y,z,\omega) &= b(x,y,z)\cdot\frac{\partial p(x,y,z,\omega)}{\partial z}+f_z(x,y,z,\omega),
\end{aligned} \tag{1}$$

where ω is the angular frequency, $\kappa(x, y, z)$ is the bulk modulus, $b(x, y, z)$ is the buoyancy, $p(x, y, z, \omega)$ is the pressure, $v_x(x, y, z, \omega)$, $v_y(x, y, z, \omega)$, $v_z(x, y, z, \omega)$ are the components of the particle velocity vector. $f_x(x, y, z, \omega)$, $f_y(x, y, z, \omega)$, $f_z(x, y, z, \omega)$ are the components of the external forces. The first block row of equation 1 is the time derivative of the Hooke's law, while the three last block rows are the equation of motion in the frequency domain.
The first-order system can be recast as a second-order equation in pressure after elimination of the particle velocities in equation 1, that leads to a generalization of the Helmholtz equation:

$$\frac{\omega^2}{\kappa(\mathbf{x})}p(\mathbf{x},\omega)+\frac{\partial}{\partial x}b(\mathbf{x})\frac{\partial p(\mathbf{x},\omega)}{\partial x}+\frac{\partial}{\partial y}b(\mathbf{x})\frac{\partial p(\mathbf{x},\omega)}{\partial y}+\frac{\partial}{\partial z}b(\mathbf{x})\frac{\partial p(\mathbf{x},\omega)}{\partial z}=s(\mathbf{x},\omega), \tag{2}$$

where $\mathbf{x} = (x,y, z)$ and $s(\mathbf{x},\omega) = \nabla \cdot \mathbf{f}$ denotes the pressure source. In exploration seismology, the source is generally a local point source corresponding to an explosion or a vertical force.
Attenuation effects of arbitrary complexity can be easily implemented in equation 2 using complex-valued wave speeds in the expression of the bulk modulus, thanks to the correspondence theorem transforming time convolution into products in the frequency domain. For example, according to the Kolsky-Futterman model (Kolsky; 1956; Futterman; 1962), the complex wave speed $\overline{c}$ is given by:

$$\overline{c}=c\left[\left(1+\frac{1}{\pi Q}\left|\log(\omega/\omega_r)\right|\right)+i\frac{sgn(\omega)}{2Q}\right]^{-1}, \tag{3}$$

where the P wave speed is denoted by c, the attenuation factor by Q and a reference frequency by ω_r.
Since the relationship between the wavefields and the source terms is linear in the first-order and second-order wave equations, equations 1 and 2 can be recast in matrix form:

$$\left[\mathbf{M}+\mathbf{S}\right]\mathbf{u}=\mathbf{A}\mathbf{u}=\mathbf{b}, \tag{4}$$

where **M** is the mass matrix, **S** is the complex stiffness/damping matrix. The sparse impedance matrix **A** has complex-valued coefficients which depend on medium properties and angular frequency. The wavefield (either the scalar pressure wavefield or the pressure-velocity wavefields) is denoted by the vector **u** and the source by **b** (Marfurt; 1984). The dimension of the square matrix **A** is the number of nodes in the computational domain multiplied by the number of wavefield components. The matrix **A** has a symmetric pattern for the FD method and the DG method discussed in this study but is generally not symmetric because of absorbing boundary conditions along the edges of the computational domain. In this study, we shall solve equation 4 by Gaussian elimination using sparse direct solver. A direct solver performs first a LU decomposition of **A** followed by forward and backward substitutions for the solutions (Duff et al.; 1986).

$$\mathbf{A}\mathbf{u}=\left(\mathbf{L}\mathbf{U}\right)\mathbf{u}=\mathbf{b} \tag{5}$$

$$\mathbf{L}\mathbf{y}=\mathbf{b}; \quad \mathbf{U}\mathbf{u}=\mathbf{y} \tag{6}$$

Exploration seismology requires to perform seismic modelling for a large number of sources, typically, up to few thousands for 3D acquisition. Therefore, our motivation behind the use of direct solver is the efficient computation of the solutions of the equation 4 for multiple sources. The LU decomposition of **A** is a time and memory demanding task but is independent of the source, and, therefore is performed only once, while the substitution phase provides the solution for multiple sources efficiently. One bottleneck of the direct-solver approach is the memory requirement of the LU decomposition resulting from the fill-in, namely, the creation of additional non-zero coefficients during the elimination process. This fill-in can be minimized by designing compact numerical stencils that allow for the minimization of the numerical bandwidth of the impedance matrix. In the following, we shall review a FD method and a finite-element DG method that allow us to fullfill this requirement.

3. Mixed-grid finite-difference method

3.1 Discretization of the differential operators

In FD methods, high-order accurate stencils are generally designed to achieve the best tradeoff between accuracy and computational efficiency (Dablain; 1986). However, direct-solver methods prevent the use of high-order accurate stencils because their large spatial support will lead to a prohibitive fill-in of the matrix during the LU decomposition (Hustedt et al.; 2004). Alternatively, the mixed-grid method was proposed by Jo et al. (1996) to design both accurate and compact FD stencils. The governing idea is to discretize the differential operators of the stiffness matrix with different second-order accurate stencils and to linearly combine the resulting stiffness matrices with appropriate weighting coefficients. The different stencils are built by discretizing the differential operators along different rotated coordinate systems ($\overline{x}$, $\overline{y}$, $\overline{z}$) such that their axes span as many directions as possible in the FD cell to mitigate numerical anisotropy. In practice, this means that the partial derivatives with respect to x, y and z in equations 1 or 2 are replaced by a linear combination of partial derivatives with respect to $\overline{x}$, $\overline{y}$ and $\overline{z}$ using the chain rule followed by the

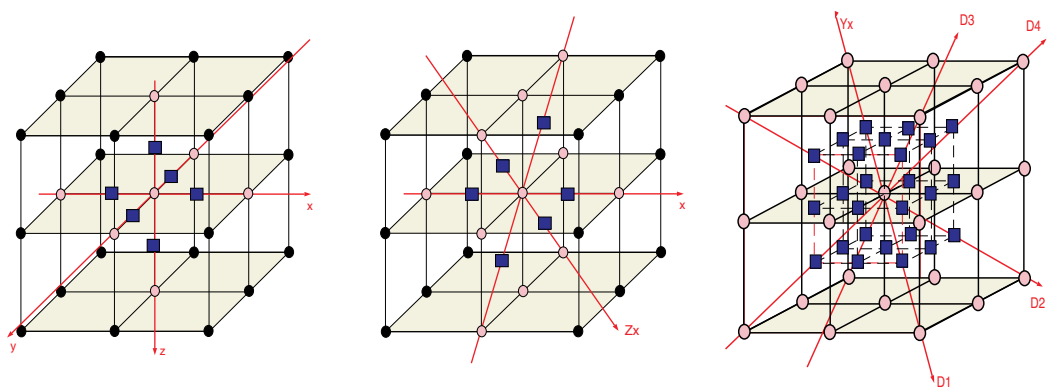
discretization of the differential operators along the axis $\bar{x}$, $\bar{y}$ and $\bar{z}$. In 2D, the coordinate systems are the classic Cartesian one and the 45°-rotated one (Saenger et al.; 2000) which lead to the 9-point stencil (Jo et al.; 1996). In 3D, three coordinate systems have been identified (Operto et al.; 2007) (Figure 1): [1] the Cartesian one which leads to the 7-point stencil, [2] three coordinate systems obtained by rotating the Cartesian system around each Cartesian axis x, y, and z. Averaging of the three elementary stencils leads to a 19-point stencil. [3] four coordinate systems defined by the four main diagonals of the cubic cell. Averaging of the four elementary stencils leads to the 27-point stencil. The stiffness matrix associated with the 7-point stencil, the 19-point stencil and the 27-point stencil will be denoted by $\mathbf{S}_1$, $\mathbf{S}_2$, $\mathbf{S}_3$, respectively.
The mixed-grid stiffness matrix $\mathbf{S}_{mg}$ is a linear combination of the stiffness matrices just-mentioned:

$$\mathbf{S}_{mg} = w_1\mathbf{S}_1 + \frac{w_2}{3}\mathbf{S}_2 + \frac{w_3}{4}\mathbf{S}_3, \tag{7}$$

where we have introduced the weighting coefficients w_1, w_2 and w_3 which satisfy:

$$w_1 + w_2 + w_3 = 1 \tag{8}$$

In the original mixed-grid approach (Jo et al.; 1996), the discretization on the different coordinate systems was directly applied to the second-order wave equation, equation 2, with the second-order accurate stencil of Boore (1972). Alternatively, Hustedt et al. (2004) proposed to discretize first the first-order velocity-pressure system, equation 1, with second-order staggered-grid stencils (Yee; 1966; Virieux; 1986; Saenger et al.; 2000) and, second, to



(a) (b) (c)

Fig. 1. Elementary FD stencils of the 3D mixed-grid stencil. Circles are pressure grid points. Squares are positions where buoyancy needs to be interpolated in virtue of the staggered-grid geometry. Gray circles are pressure grid points involved in the stencil. a) Stencil on the classic Cartesian coordinate system. This stencil incorporates 7 coefficients. b) Stencil on the rotated Cartesian coordinate system. Rotation is applied around x on the figure. This stencil incorporates 11 coefficients. Same strategy can be applied by rotation around y and z. Averaging of the 3 resultant stencils defines a 19-coefficient stencil. c) Stencil obtained from 4 coordinate systems, each of them being associated with 3 main diagonals of a cubic cell. This stencil incorporates 27 coefficients (Operto et al.; 2007).

eliminate the auxiliary wavefields (i.e., the velocity wavefields) following a parsimonious staggered-grid method originally developed in the time domain (Luo & Schuster; 1990). The parsimonious staggered-grid strategy allows us to minimize the number of wavefield components involved in the equation 4, and therefore to minimize the size of the system to be solved while taking advantage of the flexibility of the staggered-grid method to discretize first-order difference operators. The parsimonious mixed-grid approach originally proposed by Hustedt et al. (2004) for the 2D acoustic wave equation was extended to the 3D wave equation by Operto et al. (2007) and to a 2D pseudo-acoustic wave equation for transversely isotropic media with tilted symmetry axis by Operto et al. (2009). The staggered-grid method requires interpolation of the buoyancy in the middle of the FD cell which should be performed by volume harmonic averaging (Moczo et al.; 2002).
The pattern of the impedance matrix inferred from the 3D mixed-grid stencil is shown in Figure 2. The bandwidth of the matrix is of the order of N^2 (N denotes the dimension of a 3D cubic N^3 domain) and was kept minimal thanks to the use of low-order accurate stencils.

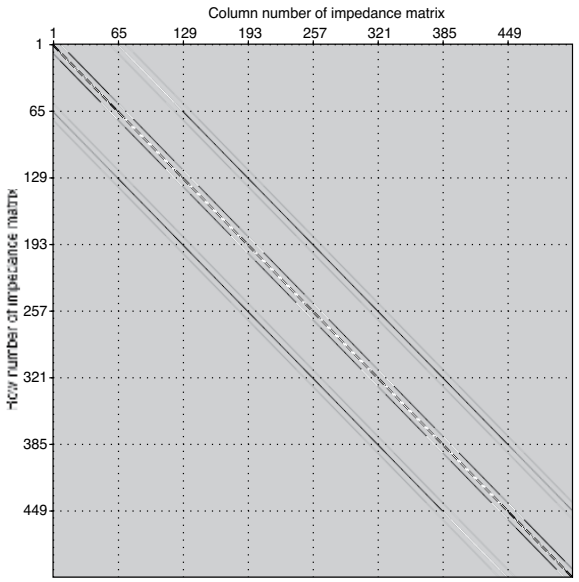


Fig. 2. Pattern of the square impedance matrix discretized with the 27-point mixed-grid stencil (Operto et al.; 2007). The matrix is band diagonal with fringes. The bandwidth is $\mathcal{O}(2N_1N_2)$ where N_1 and N_2 are the two smallest dimensions of the 3D grid. The number of rows/columns in the matrix is $N_1 \times N_2 \times N_3$. In the figure, $N_1 = N_2 = N_3 = 8$

3.2 Anti-lumped mass

The linear combination of the rotated stencils in the mixed-grid approach is complemented by the distribution of the mass term ω^2/κ in equation 2 over the different nodes of the mixed-grid stencil to mitigate the numerical dispersion:

$$\frac{\omega^2}{\kappa_{000}} p_{000} \Rightarrow \omega^2 \left(w_{m1} \left[\frac{p}{\kappa} \right]_0 + w_{m2} \left[\frac{p}{\kappa} \right]_1 + w_{m3} \left[\frac{p}{\kappa} \right]_2 + w_{m4} \left[\frac{p}{\kappa} \right]_3 \right), \tag{9}$$

where

$$w_{m1} + \frac{w_{m2}}{6} + \frac{w_{m3}}{12} + \frac{w_{m4}}{8} = 1. \tag{10}$$

In equation 9, the different nodes of the 27-point stencils are labelled by indices lmn where $l,m,n \in \{-1, 0,1\}$ and 000 denotes the grid point in the middle of the stencil.
We used the notations

$$\left[\frac{p}{\kappa}\right]_0 = \frac{p_{000}}{\kappa_{000}},$$

$$\left[\frac{p}{\kappa}\right]_1 = \frac{p_{100}}{\kappa_{100}} + \frac{p_{010}}{\kappa_{010}} + \frac{p_{001}}{\kappa_{001}} + \frac{p_{-100}}{\kappa_{-100}} + \frac{p_{0-10}}{\kappa_{0-10}} + \frac{p_{00-1}}{\kappa_{00-1}},$$

$$\left[\frac{p}{\kappa}\right]_2 = \frac{p_{110}}{\kappa_{110}} + \frac{p_{011}}{\kappa_{011}} + \frac{p_{101}}{\kappa_{101}} + \frac{p_{-110}}{\kappa_{-110}} + \frac{p_{0-11}}{\kappa_{0-11}} + \frac{p_{-101}}{\kappa_{-101}} + \frac{p_{1-10}}{\kappa_{1-10}} + \frac{p_{01-1}}{\kappa_{01-1}} + \frac{p_{10-1}}{\kappa_{10-1}} + \frac{p_{-1-10}}{\kappa_{-1-10}} + \frac{p_{0-1-1}}{\kappa_{0-1-1}} + \frac{p_{-10-1}}{\kappa_{-10-1}},$$

$$\left[\frac{p}{\kappa}\right]_3 = \frac{p_{111}}{\kappa_{111}} + \frac{p_{-1-1-1}}{\kappa_{-1-1-1}} + \frac{p_{-111}}{\kappa_{-111}} + \frac{p_{1-11}}{\kappa_{1-11}} + \frac{p_{11-1}}{\kappa_{11-1}} + \frac{p_{-1-11}}{\kappa_{-1-11}} + \frac{p_{1-1-1}}{\kappa_{1-1-1}} + \frac{p_{-11-1}}{\kappa_{-11-1}}.$$

This anti-lumped mass strategy is opposite to mass lumping used in finite element methods to make the mass matrix diagonal. The anti-lumped mass approach, combined with the averaging of the rotated stencils, allows us to minimize efficiently the numerical dispersion and to achieve an accuracy representative of 4^{th}-order accurate stencil from a linear combination of 2^{nd}-order accurate stencils. The anti-lumped mass strategy introduces four additional weighting coefficients w_{m1}, w_{m2}, w_{m3} and w_{m4}, equations 9 and 10. The coefficients w_1, w_2, w_3, w_{m1}, w_{m2}, w_{m3} and w_{m4} are determined by minimization of the phase-velocity dispersion in infinite homogeneous medium. Alternatives FD methods for designing optimized FD stencils can be found in Holberg (1987); Takeuchi and Geller (2000).

3.3 Numerical dispersion and anisotropy

The dispersion analysis of the 3D mixed-grid stencil was already developed in details in Operto et al. (2007). We focus here on the sensitivity of the accuracy of the mixed-grid stencil to the choice of the weighting coefficients w_1, w_2, w_3, w_{m1}, w_{m2}, w_{m3}. We aim to design an accurate stencil for a discretization criterion of 4 grid points per minimum propagated wavelength. This criterion is driven by the spatial resolution of full waveform inversion, which is half a wavelength. To properly sample subsurface heterogeneities, the size of which is half a wavelength, four grid points per wavelength should be used according to Shannon's theorem.
Inserting the discrete expression of a plane wave propagating in a 3D infinite homogeneous medium of wave speed c and density equal to 1 in the wave equation discretized with the mixed-grid stencil gives for the normalized phase velocity (Operto et al.; 2007):

$$\tilde{v}_{ph} = \frac{G}{\sqrt{2J\pi}} \sqrt{w_1(3-C) + \frac{w_2}{3}(6-C-B) + \frac{2w_3}{4}(3-3A+B-C)}, \tag{11}$$

where $J = (w_{m1} + 2w_{m2}C + 4w_{m3}B + 8w_{m4}A)$ with

$$A = \cos a \cos b \cos c,$$
$$B = \cos a \cos b + \cos a \cos c + \cos b \cos c,$$
$$C = \cos a + \cos b + \cos c.$$

and $a = \frac{2\pi}{G}\cos\phi\cos\theta$; $b = \frac{2\pi}{G}\cos\phi\sin\theta$; $c = \frac{2\pi}{G}\sin\phi$. Here, the normalized phase velocity is the ratio between the numerical phase velocity ω/k and the wave speed c. $G = \frac{\lambda}{h} = \frac{2\pi}{kh}$ is the number of grid points per wavelength λ. ϕ and θ are the incidence angles of the plane wave. We look for the 5 independent parameters w_{m1}, w_{m2}, w_{m3}, w_1, w_2 which minimize the least-squares norm of the misfit $(1. - \tilde{v}_{ph})$. The two remaining weighting coefficients w_{m4} and w_3 are inferred from equations 8 and 10, respectively. We estimated these coefficients by a global optimization procedure based on a Very Fast Simulating Annealing algorithm (Sen & Stoffa; 1995). We minimize the cost function for 5 angles ϕ and θ spanning between 0 and 45°and for different values of G.

In the following, the number of grid points for which phase velocity dispersion is minimized will be denoted by G_m. The values of the weighting coefficients as a function of G_m are given in Table 1. For high values of G_m, the Cartesian stencil has a dominant contribution (highlighted by the value of w_1), while the first rotated stencil has the dominant contribution for low values of G_m as shown by the value of w_2. The dominant contribution of the Cartesian stencil for large values of G_m is consistent with the fact that it has a smaller spatial support (i.e., $2 \times h$) than the rotated stencils and a good accuracy for G greater than 10 (Virieux; 1986). The error on the phase velocity is plotted in polar coordinates for four values of G (4, 6, 8, 10) and for G_m=4 in Figure 3a. We first show that the phase velocity dispersion is negligible for G=4, that shows the efficiency of the optimization. However, more significant error (0.4 %) is obtained for intermediate values of G (for example, G=6 in Figure 3a). This highlights the fact that the weighting coefficients were optimally designed to minimize the dispersion for one grid interval in homogeneous media. We show also the good isotropy properties of the stencil, shown by the rather constant phase-velocity error whatever the direction of propagation. The significant phase-velocity error for values of G greater than G_m prompt us to simultaneously minimize the phase-velocity dispersion for four values of G: G_m= 4,6,8,10 (Figure 3b). We show that the phase-velocity error is now more uniform over the values of G and that the maximum phase-velocity-error was reduced (0.25 % against 0.4 %). However, the nice isotropic property of the mixed-grid stencil was degraded and the phase-velocity dispersion was significantly increased for G=4. We conclude that the range of wavelengths propagated in a given medium should drive the discretization criterion used to infer the weighting coefficients of the mixed grid stencil and that a suitable trade-off should be found between the need to manage the heterogeneity of the medium and the need to minimize the error for a particular wavelength. Of note, an optimal strategy might consist of adapting locally the values of the weighting coefficients to the local wave speed during the assembling of the impedance matrix. This strategy was not investigated yet.

G_m	4,6,8,10	4	8	10	20	40
wm_1	0.4966390	0.5915900	0.5750648	0.7489436	0.7948160	0.6244839
wm_2	7.51233E-02	4.96534E-02	5.76759E-02	1.39044E-02	3.71392E-03	5.06646E-02
wm_3	4.38464E-03	5.10851E-03	5.56914E-03	6.38921E-03	5.54043E-03	1.42369E-03
wm_4	6.76140E-07	6.14837E-03	1.50627E-03	1.13699E-02	1.45519E-02	6.8055E-03
w_1	5.02480E-05	8.8075E-02	0.133953	0.163825	0.546804	0.479173
w_2	0.8900359	0.8266806	0.7772883	0.7665769	0.1784437	0.2779923
w_3	0.1099138	8.524394E-02	8.87589E-02	6.95979E-02	0.2747527	0.2428351

Table 1. Coefficients of the mixed-grid stencil as a function of the discretization criterion G_m for the minimization of the phase velocity dispersion.

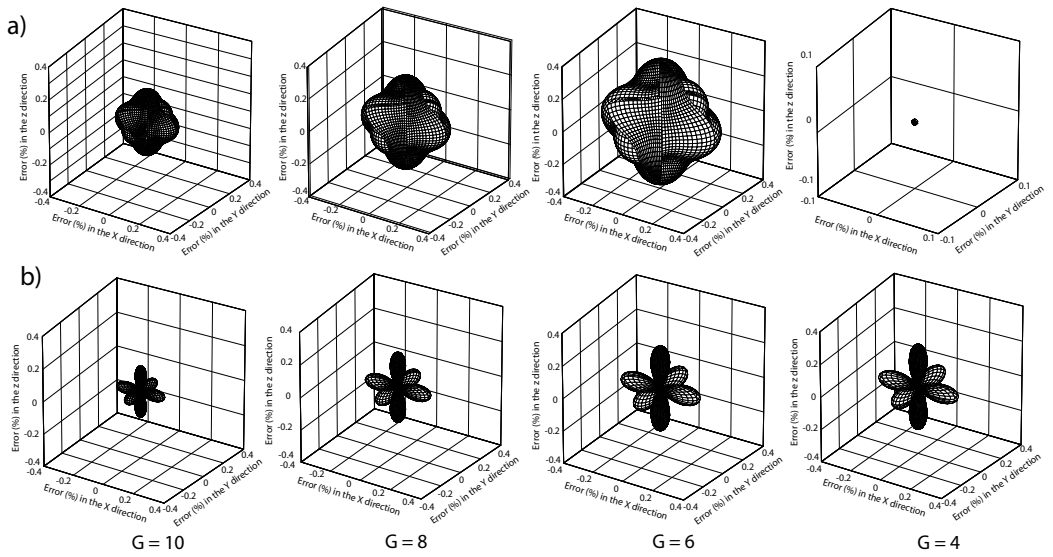


Fig. 3. Phase-velocity dispersion shown in spherical coordinates for four values of G. (a) The phase-velocity dispersion was minimized for G = 4. (b) the phase-velocity dispersion was minimized for 4 values of G: 4, 6, 8 and 10.

Comparison between numerical and analytical pressure monochromatic wavefields computed in a homogeneous medium of wave speed 1.5 km/s and density 1000 kg/m^3 confirms the former theoretical analysis (Figure 4). The frequency is 3.75 Hz corresponding to a propagated wavelength of 400 m. The grid interval for the simulation is 100 m corresponding to G = 4. Simulations were performed when the weighting coefficients of the mixed-grid stencils are computed for G_m = 4 and G_m = {4, 6, 8,10}. The best agreement is obtained for the weighting coefficients associated with G_m = 4 as expected from the dispersion analysis.

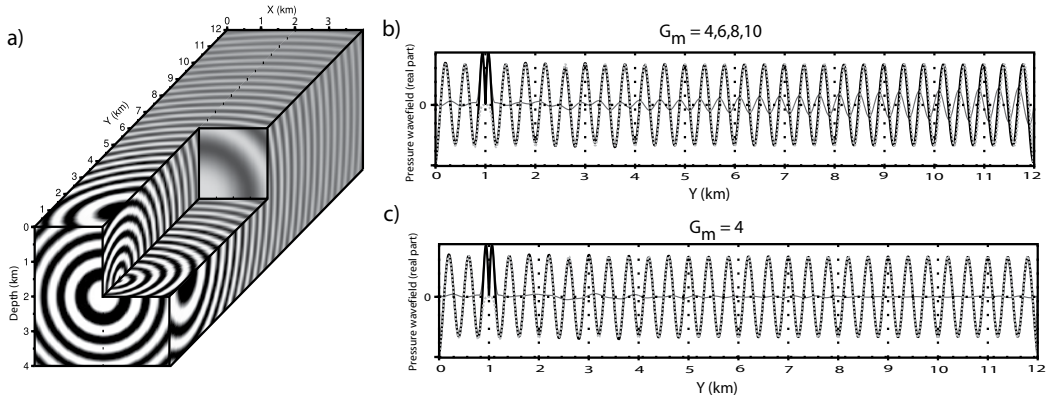


Fig. 4. (a) Real part of a 3.75-Hz monochromatic wavefield computed with the mixed-grid stencil in a 3D infinite homogeneous medium. The explosive point source is at x=2 km, y=1 km, z=2 km. (b-c) Comparison between the analytical (gray) and the numerical solution (black) for a receiver line oriented in the *Y* direction across the source position. The thin black line is the difference. The amplitudes were corrected for 3D geometrical spreading. (b) G_m = 4, 6, 8, 10. (c) G_m = 4.

3.4 Boundary conditions

In seismic exploration, two boundary conditions are implemented for wave modelling: absorbing boundary conditions to mimic an infinite medium and free surface conditions on the top side of the computational domain to represent the air-solid or air-water interfaces.

3.4.1 PML absorbing boundary conditions

We use Perfectly-Matched Layers (PML) absorbing boundary conditions (Berenger; 1994) to mimic an infinite medium. In the frequency domain, implementation of PMLs consists of applying in the wave equation a new system of complex-valued coordinates $\tilde{x}$ defined by (e.g., Chew & Weedon; 1994):

$$\frac{\partial}{\partial \tilde{x}} = \frac{1.}{\xi_x(x)} \frac{\partial}{\partial x} \tag{12}$$

In the PML layers, the damped wave equation writes:

$$\left[\frac{\omega^2}{\kappa(\mathbf{x})} + \frac{1}{\xi_x(x)} \frac{\partial}{\partial x} \frac{b(\mathbf{x})}{\xi_x(x)} \frac{\partial}{\partial x} + \frac{1}{\xi_y(y)} \frac{\partial}{\partial y} \frac{b(\mathbf{x})}{\xi_y(y)} \frac{\partial}{\partial y} + \frac{1}{\xi_z(z)} \frac{\partial}{\partial z} \frac{b(\mathbf{x})}{\xi_z(z)} \frac{\partial}{\partial z} \right] p(\mathbf{x}, \omega) = -s(\mathbf{x}, \omega), \tag{13}$$

where $\xi_x(x) = 1 + i\gamma_x(x)/\omega$ and $\gamma_x(x)$ is a 1D damping function which defines the PML damping behavior in the PML layers. These functions differ from zero only inside the PML layers. In the PML layers, we used $\gamma(x) = c_{pml}\left(1 - cos(\frac{\pi}{2}\frac{L-x}{L})\right)$ where L denotes the width of the PML layer and x is a local coordinate in the PML layer whose origin is located at the outer edges of the model. The scalar c_{pml} is defined by trial and error depending on the width of the PML layer. The procedure to derive the unsplitted second-order wave equation with PML conditions, equation 13, from the first-order damped wave equation is given in Operto et al. (2007).

The absorption of the PML layers at grazing incidence can be improved by using convolutional PML (C-PML) (Kuzuoglu & Mittra; 1996; Roden & Gedney; 2000; Komatitsch & Martin; 2007). In the C-PML layers, the damping function $\xi_x(x)$ becomes:

$$\xi_x(x) = \kappa_x + i\frac{d_x}{\alpha_x + i\omega}, \tag{14}$$

where d_x and α_x are generally quadratic and linear functions, respectively. Suitable expression for κ_x, d_x and α_x are discussed in Kuzuoglu & Mittra (1996); Collino & Monk (1998); Roden & Gedney (2000); Collino & Tsogka (2001); Komatitsch & Martin (2007); Drossaert & Giannopoulos (2007).

3.4.2 Free surface boundary conditions

Planar free surface boundary conditions can be simply implemented in the frequency domain with two approaches. In the first approach, the free surface matches the top side of the FD grid and the pressure is forced to zero on the free surface by using a diagonal impedance matrix for rows associated with collocation grid points located on the top side of the FD grid. Alternatively, the method of image can be used to implement the free surface along a virtual plane located half a grid interval above the topside of the FD grid (Virieux;

1986). The pressure is forced to vanish at the free surface by using a ficticious plane located half a grid interval above the free surface where the pressure is forced to have opposite values to that located just below the free surface.
From a computer implementation point of view, an impedance matrix is typically built row per row. One row of the linear system can be written as:

$$\sum_{i_3=-1,1}\sum_{i_2=-1,1}\sum_{i_1=-1,1} a_{i_1 i_2 i_3}\, p_{i_1 i_2 i_3} = s_{000} \tag{15}$$

where $a_{i_1 i_2 i_3}$ are the coefficients of the 27-point mixed grid stencil and 000 denote the indices of the collocation coefficient located in the middle of the stencil in a local coordinate system. The free surface boundary conditions writes:

$$p_{-1 i_2 i_3} = -p_{0 i_2 i_3}, \tag{16}$$

for $i_2 = \{-1, 0,1\}$ and $i_3 = \{-1, 0,1\}$. The indices i_1=-1 and i_1=0 denotes here the grid points just above and below the free surface, respectively.
For a grid point located on the top side of the computational domain (i.e., half a grid interval below free surface), equation 15 becomes:

$$\sum_{i_3=-1,1}\sum_{i_2=-1,1} a_{1 i_2 i_3}\, p_{1 i_2 i_3} + \sum_{i_3=-1,1}\sum_{i_2=-1,1} \left(a_{0 i_2 i_3} - a_{-1 i_2 i_3}\right) p_{0 i_2 i_3} = s_{000}, \tag{17}$$

where $p_{-1 i_2 i_3}$ has been replaced by the opposite value of $p_{0 i_2 i_3}$ according to equation 16.
Our practical experience is that both implementation of free surface boundary conditions give results of comparable accuracy. Of note, rigid boundary conditions (zero displacement perpendicular to the boundary) or periodic boundary conditions (Ben-Hadj-Ali et al.; 2008) can be easily implemented with the method of image following the same principle than for the free surface condition.

3.5 Source implementation on coarse grids

Seismic imaging by full waveform inversion is initiated at frequency as small as possible to mitigate the non linearity of the inverse problem. The starting frequency for modelling can be as small as 2 Hz which can lead to grid intervals as large as 200 m. In this framework, accurate implementation of point source at arbitrary position in a coarse grid is critical. One method has been proposed by Hicks (2002) where the point source is approximated by a windowed Sinc function. The Sinc function is defined by

$$sinc(x) = \frac{sin(\pi x)}{\pi x}, \tag{18}$$

where $x = (x_g - x_s)$, x_g denotes the position of the grid nodes and x_s denotes the position of the source. The Sinc function is tapered with a Kaiser function to limit its spatial support. For multidimensional simulations, the interpolation function is built by tensor product construction of 1D windowed Sinc functions. If the source positions matches the position of one grid node, the Sinc function reduces to a Dirac function at the source position and no approximation is used for the source positioning. If the spatial support of the Sinc function

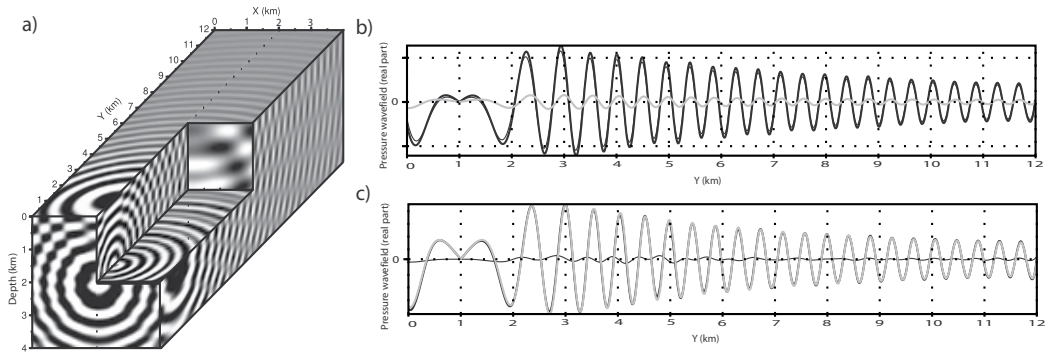


Fig. 5. a) Real part of a 3.75-Hz monochromatic wavefield in a homogeneous half space. (b) Comparison between numerical (black) and analytical (gray) solutions at receiver positions. The Sinc interpolation with 4 coefficients was used for both the source implementation and the extraction of the solution at the receiver positions on a coarse FD grid.

intersects a free surface, part of the Sinc function located above the free surface is mirrored into the computational domain with a reverse sign following the method of image. Vertical force can be implemented in a straightforward way by replacing the Sinc function by its vertical derivative. The same interpolation function can be used for the extraction of the pressure wavefield at arbitrary receiver positions. The accuracy of the method of Hicks (2002) is illustrated in Figure 5 which shows a 3.5-Hz monochromatic wavefield computed in a homogeneous half space. The wave speed is 1.5 km/s and the density is 1000 kg/m^3. The grid interval is 100 m. The free surface is half a grid interval above the top of the FD grid and the method of image is used to implement the free surface boundary condition. The source is in the middle of the FD cell at 2 km depth. The receiver line is oriented in the *Y* direction. Receivers are in the middle of the FD cell in the horizontal plane and at a depth of 6 m just below the free surface. This setting is representative of a ocean bottom survey where the receiver is on the sea floor and the source is just below the sea surface (in virtue of the spatial reciprocity of the Green functions, sources are processed here as receivers and *vice versa*). Comparison between the numerical and the analytical solutions at the receiver positions are first shown when the source is positioned at the closest grid point and the numerical solutions are extracted at the closest grid point (Figure 5b). The amplitude of the numerical solution is strongly overestimated because the numerical solution is extracted at a depth of 50 m below free surface (where the pressure vanishes) instead of 6 m. Second, a significant phase shift between numerical and analytical solutions results from the approximate positioning of the sources and receivers. In contrast, a good agreement between the numerical and analytical solutions both in terms of amplitude and phase is shown in Figure 5c where the source and receiver positioning were implemented with the windowed Sinc interpolation.

3.6 Resolution with the sparse direct solver MUMPS

To solve the sparse system of linear equations, equation 4, we used the massively parallel direct MUMPS solver designed for distributed memory platforms. The reader is referred to Guermouche et al. (2003); Amestoy et al. (2006); MUMPS team (2009) for an extensive description of the method and their underlying algorithmic aspects. The MUMPS solver is based on a multifrontal method (Duff et al.; 1986; Duff and Reid; 1983; Liu; 1992), where the

resolution of the linear system is subdivided into 3 main tasks. The first one is an analysis phase or symbolic factorization. Reordering of the matrix coefficients is first performed in order to minimize fill-in. We used the METIS algorithm which is based on a hybrid multilevel nested-dissection and multiple minimum degree algorithm (Karypis & Kumar; 1999). Then, the dependency graph which describes the order in which the matrix can be factored is estimated as well as the memory required to perform the subsequent numerical factorization. The second task is the numerical factorization. The third task is the solution phase performed by forward and backward substitutions. During the solution phase, multiple-shot solutions can be computed simultaneously from the LU factors taking advantage of threaded BLAS3 (Basic Linear Algebra Subprograms) library and are either assembled on the host or kept distributed on the processors for subsequent parallel computations.

We performed the factorization and the solutions phases in complex arithmetic single precision. To reduce the condition number of the matrix, a row and column scaling is applied in MUMPS before factorization. The sparsity of the matrix and suitable equilibration have made single precision factorization accurate enough so far for the 2D and 3D problems we tackled. If single precision factorization would be considered not accurate enough for very large problems, an alternative approach to double precision factorization may be the postprocessing of the solution by a simple and fast iterative refinement performed in double precision (Demmel (1997), pages 60-61 and Langou et al. (2006); Kurzak & Dongarra (2006)).

The main two bottlenecks of sparse direct solver is the time and memory complexity and the limited scalability of the LU decomposition. By complexity is meant the increase of the computational cost (either in terms of elapsed time or memory) of an algorithm with the size of the problem, while scalability describes the ability of a given algorithm to use an increasing number of processors. The theoretical memory and time complexity of the LU decomposition for a sparse matrix, the pattern of which is shown in Figure 2, is $\mathcal{O}(N^4)$ and $\mathcal{O}(N^6)$, respectively, where N is the dimension of a 3D cubic N^3 grid.

We estimated the observed memory complexity and scalability of the LU factorization by means of numerical experience. The simulations were performed on the SGI ALTIX ICE supercomputer of the computer center CINES (France). Nodes are composed of two quad-core INTEL processors E5472. Each node has 30 Gbytes of useful memory. We used two MPI process per node and four threads per MPI process. In order to estimate the memory complexity, we performed simulations on cubic models of increasing dimension with PML absorbing boundary conditions along the 6 sides of the model. The medium is homogeneous and the source is on the middle of the grid. Figure 6a shows the memory required to store the complex-valued LU factors as a function of N. Normalization of this curve by the real memory complexity will lead to a horizontal line. We found an observed memory complexity of $\mathcal{O}(Log_2(N)N^{3.9})$ (Figure 6b) which is consistent with the theoretical one. In order to assess the scalability of the LU factorization, we consider a computational FD grid of dimensions 177 x 107 x 62 corresponding to 1.17 millions of unknowns. The size of the grid corresponds to a real subsurface target for oil exploration at low frequency (3.5 Hz). We computed a series of LU factorization using an increasing number of processors N_p, starting with $N_{p_{ref}} = 2$. The elapsed time of the LU factorization (T_{LU}) and the parallelism efficiency ($T_{LU}(N_{p_{ref}}) \times N_{p_{ref}} / T_{LU}(N_p) \times N_p$) are shown in Figure 6(c-d). The efficiency drops rapidly as the number of processors increased, down to a value of 0.5 for $N_P = 32$ (Figure 6d). This clearly indicates that the most suitable platform for sparse direct solver should be composed

of a limited number of nodes with a large amount of shared memory. The efficiency of the multi-r.h.s solution phase is significantly improved by using multithreaded BLAS3 library.

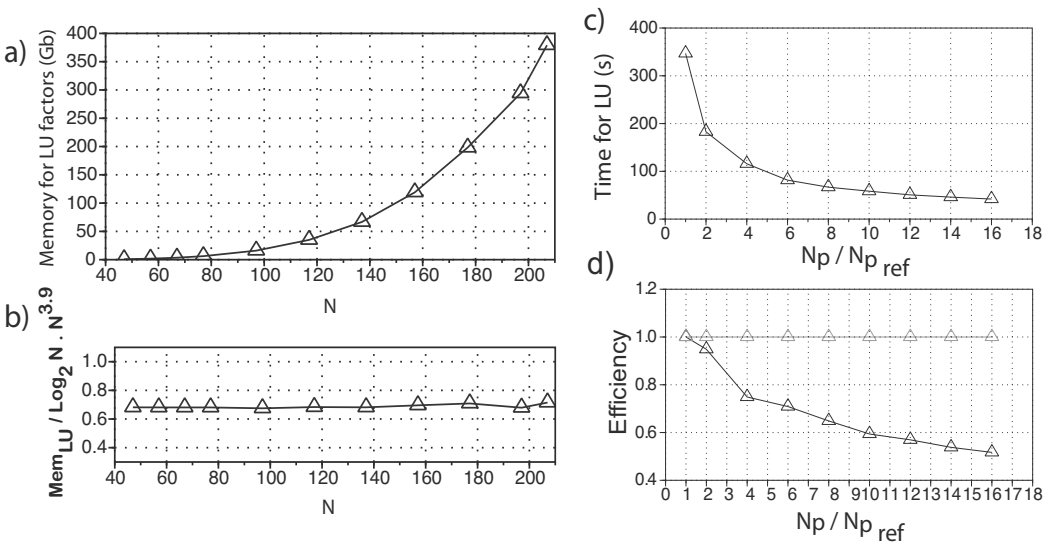


Fig. 6. (a-b) Memory complexity of LU factorization. (a) Memory in Gbytes required for storage of LU factors. (b) Memory required for storage of LU factors divided by $Log_2 N.N^{3.9}$. N denotes the dimension of a 3D N^3 grid. The largest simulation for N = 207 corresponds to 8.87 millions of unknowns. (c-d) Scalability analysis of LU factorization. (c) Elapsed time for LU factorization versus the number of MPI processes. (d) Efficiency.

3.7 Numerical examples

We present acoustic wave modelling in two realistic 3D synthetic velocity models, the SEG/EAGE overthrust and salt models, developed by the oil exploration community to assess seismic modelling and imaging methods (Aminzadeh et al.; 1997). The simulation was performed on the SGI ALTIX ICE supercomputer just described.

3.7.1 3D EAGE/SEG overthrust model

The 3D SEG/EAGE Overthrust model is a constant density onshore acoustic model covering an area of 20 km × 20 km × 4.65 km (Aminzadeh et al.; 1997)(Figure 7a). From a geological viewpoint, it represents a complex thrusted sedimentary succession constructed on top of a structurally decoupled extensional and rift basement block. The overthrust model is discretized with 25 m cubic cells, representing an uniform mesh of 801 × 801 × 187 nodes. The minimum and maximum velocities in the Overthrust model are 2.2 and 6.0 km/s respectively. We present the results of a simulation performed with the mixed-grid FD method (referred to as FDFD in the following) for a frequency of 7 Hz and for a source located at x=2.4 km, y=2.4 km and z=0.15 km. The model was resampled with a grid interval of 75 m that corresponds to four grid points per minimum wavelength. The size of the resampled FD grid is 266 x 266 x 62. PML layers of 8 grid points were added along the 6 sides of the 3D FD grid. This leads to 6.2 millions of pressure unknowns. For the simulation,

we used the weights of the mixed-grid stencil obtained for G_m = 4, 6, 8, 10. These weights provided slightly more accurate results than the weights obtained for G_m = 4, in particular for waves recorded at long source-receiver offsets. The 7-Hz monochromatic wavefield computed with the FDFD method is compared with that computed with a classic $\mathcal{O}(\Delta t^2,\Delta x^4)$ staggered-grid FD time-domain (FDTD) method where the monochromatic wavefield is integrated by discrete Fourier transform within the loop over time steps (Sirgue et al.; 2008) (Figure 7).

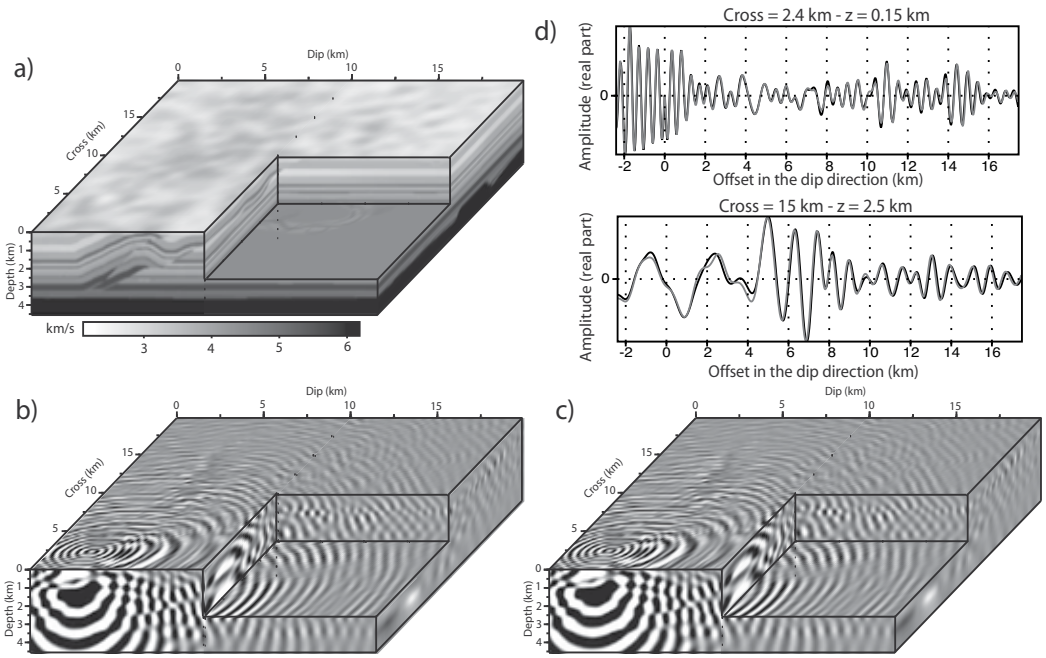


Fig. 7. (a) Overthrust velocity model. (b-c) 7-Hz monochromatic wavefield (real part) computed with the FDFD (b) and FDTD (c) methods.(d) Direct comparison between FDFD (gray) and FDTD (black) solutions. The receiver line in the dip direction is: (top) at 0.15-km depth and at 2.4 km in the cross direction. The amplitudes were corrected for 3D geometrical spreading; (bottom) at 2.5-km depth and at 15 km in the cross direction.

We used the same spatial FD grid for the FDTD and FDFD simulations. The simulation length was 15 s in the FDTD modelling. We obtain a good agreement between the two solutions (Figure 7d). The statistics of the FDFD and FDTD simulations are outlined in Table 2. The FDFD simulation was performed on 32 MPI processes with 2 threads and 15 Gbytes of memory per MPI process. The total memory required by the LU decomposition of the impedance matrix was 260 Gbytes. The elapsed time for LU decomposition was 1822 s and the elapsed time for one r.h.s was 0.97 s. Of note, we processed efficiently groups of 16 sources in parallel during the solution step by taking advantage of the multi-rhs functionality of MUMPS and the threaded BLAS3 library. The elapsed time for the FDTD simulation was 352 s on 4 processors. Of note, C-PML absorbing boundary conditions were implemented in the full model during FDTD modelling to mimic attenuation effects

Model	F (Hz)	h(m)	n_u (10^6)	$M_{LU}(Gb)$	$T_{LU}(s)$	$T_s(s)$	N_p^{fdfd}	N_P^{fdtd}	$T_{fdtd}(s)$
Over.	7	75	6.2	260	1822	0.97	32	4	352
Salt	7.34	50	8.18	402.5	2863	1.4	48	16	211

Table 2. Statistics of the simulation in the overthrust (top row) and in the salt (bottom row) models. $F(Hz)$: frequency; $h(m)$: FD grid interval; n_u: number of unknowns; M_{LU}: memory used for LU factorization in Gbytes; T_{LU}: elapsed time for factorization; T_s: elapsed time for one solution phase; N_p^{fdfd} : number of MPI processors used for FDFD; N_p^{fdtd} : number of MPI processors used for FDTD; T_{fdtd}: elapsed time for one FDTD simulation.

Model	Method	Pre. (hr)	Sol. (hr)	Total (hr)	Pre. (hr)	Sol. (hr)	Total (hr)
Over.	FDTD	0	21.7	21.7	0	0.96	0.96
Over.	FDFD	0.5	0.54	1.04	0.5	0.0134	0.51
Salt	FDTD	0	39	39	0	0.94	0.94
Salt	FDFD	0.8	0.78	1.58	0.80	0.016	0.816

Table 3. Comparison between FDTD and FDFD modelling for 32 (left) and 2000 (right) processors. The number of sources is 2000. *Pre*. denotes the elapsed time for the source-independent task during seismic modelling (i.e., the LU factorization in the FDFD approach). *Sol*. Denotes the elapsed time for multi-r.h.s solutions during seismic modelling (i.e., the substitutions in the FDFD approach).

implemented with memory variables. To highlight the benefit of the direct-solver approach for multi-r.h.s simulation on a small number of processors, we compare the performances of the FDFD and FDTD simulations for 2000 sources (Table 3). If the number of available processors is 32, the FDFD method is more than one order of magnitude faster than the FDTD one thanks of the efficiency of the solution step of the direct-solver approach. If the number of processors equals to the number of sources, the most efficient parallelization of the FDTD method consists of assigning one source to one processor and performing the FDTD simulation in sequential on each processor. For a large number of processors, the cost of the FDFD method is dominated by the LU decomposition (if the 2000 processors are splitted into groups of 32 processors, each group being assigned to the processing of 2000/32 sources) and the computational cost of the two methods is of the same order of magnitude. This schematic analysis highlights the benefit of the FDFD method based on sparse direct solver to tackle efficiently problems involving few millions of unknowns and few thousands of r.h.s on small distributed-memory platforms composed of nodes with a large amount of shared memory.

3.7.2 3D EAGE/SEG salt model

The salt model is a constant density acoustic model covering an area of 13.5 km × 13.5 km × 4.2 km (Aminzadeh et al.; 1997)(Figure 8). The salt model is representative of a Gulf Coast salt structure which contains salt sill, different faults, sand bodies and lenses. The salt model is discretized with 20 m cubic cells, representing an uniform mesh of 676 x 676 x 210 nodes. The minimum and maximum velocities in the salt model are 1.5 and 4.482 km/s respectively. We performed a simulation for a frequency of 7.34 Hz and for one source located at x=3.6 km, y=3.6 km and z = 0.1 km. The model was resampled with a grid interval of 50 m corresponding to 4 grid points per minimum wavelength. The dimension of the

resampled grid is 270 x 270 x 84 which represents 8.18 millions of unknowns after addition of the PML layers. We used the weights of the mixed-grid stencil inferred from G_m = 4, 6, 8, 10. Results of simulations performed with the FDFD and FDTD methods are compared in Figure 8. The length of the FDTD simulation was 15 s.

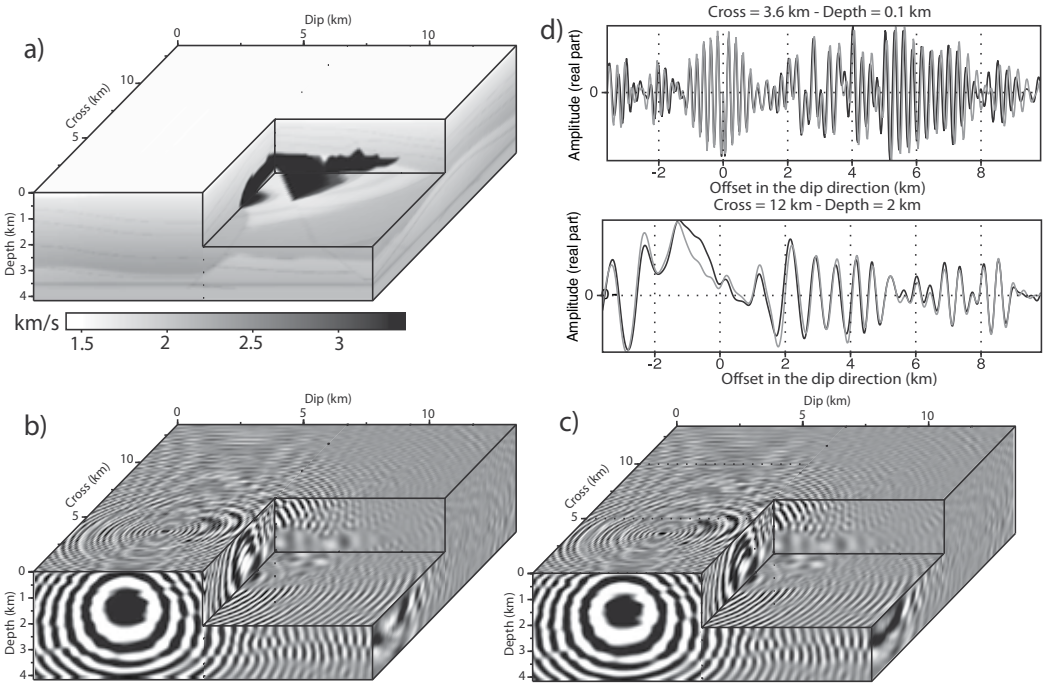


Fig. 8. (a) Salt velocity model. (b-c) 7.33-Hz monochromatic wavefield (real part) computed with the FDFD (b) and the FDTD (c) methods. (d) Direct comparison between FDFD (gray) and FDTD (black) solutions. The receiver line in the dip direction is: (top) at 0.1-km depth and at 3.6 km in the cross direction. The amplitudes were corrected for 3D geometrical spreading. (bottom) at 2.5-km depth and at 15 km in the cross direction.

The statistics of the simulation are outlined in Table 2. We obtain a good agreement between the two solutions (Figure 8d) although we show a small phase shift between the two solutions at offsets greater than 5 km. This phase shift results from the propagation in the high-velocity salt body. This phase shift is higher when the FDFD is performed with weights inferred from G_m = 4. The direct-solver modelling was performed on 48 MPI process using 2 threads and 15 Gbytes of memory per MPI process. The memory and the elapsed time for the LU decomposition were 402 Gbytes and 2863 s, respectively. The elapsed time for the solution step for one r.h.s was 1.4 s when we process 16 rhs at a time during the solution step in MUMPS. The elapsed time for one FDTD simulation on 16 processors was 211 s. As for the overthrust model, the FDFD approach is more than one order of magnitude faster than the FDTD one when a large number of r.h.s (2000) and a small number of processors (48) are used (Table 3). For a number of processors equal to the number of r.h.s, the two approaches have the same cost. Of note, in the latter configuration (N_P=N_{rhs}), the cost of the FDFD modelling and of the FDTD modelling are almost equal in the case of the salt model (0.94 h versus 0.816 h) while the FDFD modelling was almost two times faster in the case of

the smaller overthrust case study (0.96 h versus 0.51 h). This trend simply highlights the higher scalability of the FDTD method.

4. Finite-element Discontinuous Galerkin method in the frequency domain

We just presented applications of the FD frequency-domain method in weakly-contrasted media with flat topography where the FD where the FD method is expected to perform well. However, in land exploration seismology, there is a need to perform elastic wave modelling in area of complex topography such as foothills and thrust belts (Figure 9). Moreover, onshore targets often exhibit weathered layers with very low wave speeds in the near surface which require a locally-refined discretization for accurate modelling. In shallow water environment, a mesh refinement is also often required near the sea floor for accurate modelling of guided and interface waves near the sea floor. Accurate modelling of acoustic and elastic waves in presence of complex boundaries of arbitrary shape and the local adaptation of the discretization to local features such as weathered near surface layers or sea floor were two of our motivations behind the development of a discontinuous finite element method on unstructured meshes for acoustic and elastic wave modelling.

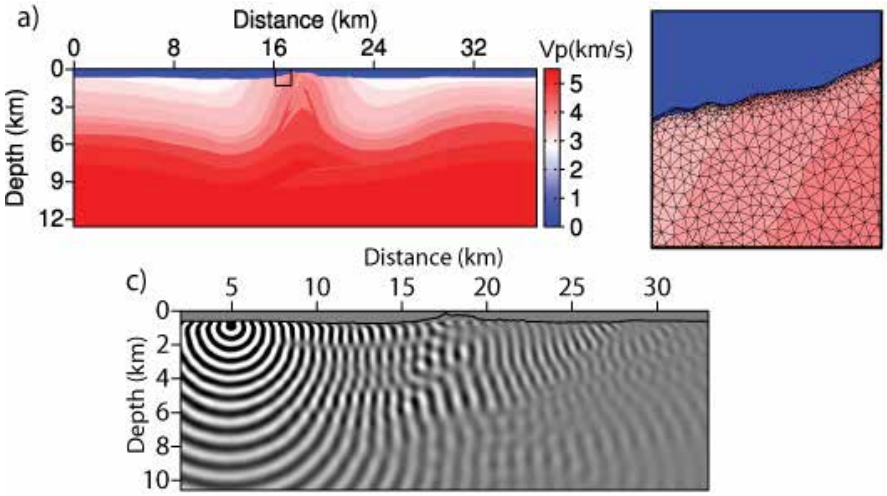


Fig. 9. Application of the DG method in seismic exploration. (a) Velocity model representative of a foothill area affected by a hilly relief and a weathered layer in the near surface. (b) Close-up of the unstructured triangular mesh locally refined near the surface. (c) Example of monochromatic pressure wavefield.

4.1 *hp*-adaptive Discontinuous Galerkin discretization

In the finite-element framework, the wavefields are approximated by means of local polynomial basis functions defined in volume elements. In the following, we adopt the nodal form of the DG formulation, assuming that the wavefield vector is approximated in triangular or tetrahedral elements for 2D and 3D problems, respectively:

$$\vec{u}_i(\omega,x,y,z,t) = \sum_{j=1}^{d_i} \vec{u}_{ij}(\omega,x_j,y_j,z_j)\varphi_{ij}(\omega,x,y,z), \tag{19}$$

where $\vec{u}$ is the wavefield vector of components $\vec{u} = (p, v_x, v_y, v_z)$. i is the index of the element in an unstructured mesh. $\vec{u}_i(\omega, x, y, z)$ denotes the wavefield vector in the element i and (x, y, z) are the coordinates inside the element i. In the framework of the nodal form of the DG method, φ_{ij} denotes Lagrange polynomial and d_i is the number of nodes in the element i. The position of the node j in the element i is denoted by the local coordinates (x_j, y_j, z_j).
In the following, the first-order acoustic velocity-pressure system, equation 1, will be written in a pseudo-conservative form:

$$\mathcal{M}\vec{u} = \sum_{\theta \in \{x,y,z\}} \partial_\theta \left(\mathcal{N}_\theta \vec{u} \right) + \vec{s}, \tag{20}$$

where

$$\mathcal{M} = \begin{pmatrix} -i\omega/\kappa & 0 & 0 & 0 \\ 0 & -i\omega\rho & 0 & 0 \\ 0 & 0 & -i\omega\rho & 0 \\ 0 & 0 & 0 & -i\omega\rho \end{pmatrix} \tag{21}$$

$$\mathcal{N}_x = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \mathcal{N}_y = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \mathcal{N}_z = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \tag{22}$$

$\mathcal{N}_\theta \vec{u}$ are linear fluxes and $\vec{s}$ is the source vector.
The first step in the finite-element formulation is to obtain the weak form of the first-order acoustic velocity-stress system by multiplying equation 20 by a test function φ_{ir} and integration over the element volume V_i

$$\int_{V_i} \varphi_{ir} \mathcal{M}_i \vec{u}_i dV = \int_{V_i} \varphi_{ir} \sum_{\theta \in \{x,y,z\}} \partial_\theta \left(\mathcal{N}_\theta \vec{u}_i \right) dV + \int_{V_i} \varphi_{ir} \vec{s}_i dV, \tag{23}$$

where $r \in [1, d_i]$. In the framework of Galerkin methods, we used the same function for the test function and the shape function, equation 19.
Integration by parts of the right hand side of equation 23 leads to:

$$\int_{V_i} \varphi_{ir} \mathcal{M} \vec{u}_i dV = -\int_{V_i} \sum_{\theta \in \{x,y,z\}} \partial_\theta \varphi_{ir} \left(\mathcal{N}_\theta \vec{u}_i \right) dV + \int_{S_i} \varphi_{ir} \left(\sum_{\theta \in \{x,y,z\}} \mathcal{N}_\theta n_\theta \right) \vec{u}_i dS + \int_{V_i} \varphi_{ir} \vec{s}_i dV, \tag{24}$$

where S_i is the surface of the element i and $\vec{n} = (n_x, n_y, n_z)$ is the outward pointing unit normal vector with respect to the surface i. We recognize in the second term of the right-hand side of equation 24 the numerical flux $\mathbf{f}_i$ defined by:

$$\mathbf{n} \cdot \mathbf{f}_i = \sum_{\theta \in \{x,y,z\}} \mathcal{N}_\theta n_\theta \vec{u}_i \tag{25}$$

A suitable expression $\mathbf{f}_{i/k}$ of the numerical flux $\mathbf{f}_i$ should guarantee the consistency between the values of the wavefield computed at a node shared by two neighbor elements i and k.

In this study we used centered fluxes for their energy conservation properties (Remaki; 2000):

$$\mathbf{f}_{i/k} = \mathbf{f}_i\left(\frac{\vec{u}_i + \vec{u}_k}{2}\right) \tag{26}$$

Assuming constant physical properties per element and plugging the expression of the centered flux, equation 26, in equation 24 give:

$$\mathcal{M}_i \int_{V_i} \varphi_{ir}\vec{u}_i dV = -\int_{V_i} \sum_{\theta\in\{x,y,z\}} \partial_\theta \varphi_{ir}\left(\mathcal{N}_\theta \vec{u}_i\right) dV + \frac{1}{2}\sum_{k\in N_i}\int_{S_{ik}} \varphi_{ir}\mathcal{P}_{ik}\left(\vec{u}_i + \vec{u}_k\right) dS + \int_{V_i} \varphi_{ir}\vec{s}_i dV, \tag{27}$$

where $k \in N_i$ represents the elements k adjacent to the element i, S_{ik} is the face between elements i and k; and $\mathcal{P}$ is defined as follow:

$$\mathcal{P}_{ik} = \sum_{\theta\in\{x,y,z\}} n_{ik\theta}\mathcal{N}_\theta, \tag{28}$$

where $n_{ik\theta}$ is the component along the θ axis of the unit vector $\vec{n}_{ik}$ of the face S_{ik}.
Equations 27 shows that the computation of the wavefield in one element requires only information from the directly neighboring elements. This highlights clearly the local nature of the DG scheme. If we replace the expression of $\vec{u}_i$ and $\vec{u}_k$ by their decomposition on the polynomial basis, equation 19, we get:

$$\left(\mathcal{M}_i \otimes \mathcal{K}_i\right)\vec{\vec{u}}_i = -\sum_{\theta\in\{x,y,z\}}\left(\mathcal{N}_\theta \otimes \mathcal{E}_{i\theta}\right)\vec{\vec{u}}_i + \frac{1}{2}\sum_{k\in N_i}\left[\left(\mathcal{Q}_{ik}\otimes\mathcal{F}_{ik}\right)\vec{\vec{u}}_i + \left(\mathcal{Q}_{ik}\otimes\mathcal{G}_{ik}\right)\vec{\vec{u}}_k\right] + \left(\mathcal{I}\otimes\mathcal{K}_i\right)\vec{\vec{s}}_i \tag{29}$$

where the coefficients rj of the mass matrix $\mathcal{K}_i$, of the stiffness matrix $\mathcal{E}_i$ and of the flux matrices $\mathcal{F}_i$ and $\mathcal{G}_i$ are respectively given by:

$$\begin{aligned}
\left(\mathcal{K}_i\right)_{rj} &= \int_{V_i}\varphi_{ir}\varphi_{ij}dV, \quad j,r\in\left[1,d_i\right] \\
\left(\mathcal{E}_{i\theta}\right)_{rj} &= \int_{V_i}\left(\partial_\theta\varphi_{ir}\right)\varphi_{ij}dV, \quad j,r\in\left[1,d_i\right]\ \theta\in\{x,y,z\} \\
\left(\mathcal{F}_{ik}\right)_{rj} &= \int_{S_{ik}}\varphi_{ir}\varphi_{ij}dS, \quad j,r\in\left[1,d_i\right] \\
\left(\mathcal{G}_{ik}\right)_{rj} &= \int_{S_{ik}}\varphi_{ir}\varphi_{kj}dS, \quad r\in\left[1,d_i\right], j\in\left[1,d_k\right]
\end{aligned} \tag{30}$$

In equation 29, $\vec{\vec{u}}_i$ and $\vec{\vec{s}}_i$ gather all nodal values for each component of the wavefield and source. $\mathcal{I}$ is the identity matrix and $\otimes$ is the tensor product of two matrices **A** and **B**:

$$\mathbf{A}\otimes\mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \dots & a_{1m}\mathbf{B} \\ . & . & . \\ . & . & . \\ . & . & . \\ a_{n1}\mathbf{B} & \dots & a_{nm}\mathbf{B} \end{bmatrix} \tag{31}$$

where ($n \times m$) denotes the dimensions of the matrix **A**. The four matrices $\mathcal{K}_i$, $\mathcal{E}_i$, $\mathcal{F}_{ik}$ and $\mathcal{G}_{ik}$ are computed by exact numerical integration.
It is worth noting that, in equation 30, arbitrary polynomial order of the shape functions can be used in elements i and k indicating that the approximation orders are totally decoupled from one element to another. Therefore, the DG allows for varying approximation orders in the numerical scheme, leading to the p-adaptivity.
Equation 30 can be recast in matrix form as:

$$\mathbf{A}\,\mathbf{u} = \mathbf{s} \tag{32}$$

In contrast to the parsimonious FD formulation, we do not eliminate the auxiliary velocity wavefields from the system because the elimination procedure is a cumbersome task in the DG formulation.

4.2 Which interpolation orders?

For the shape and test functions, we used low-order Lagrangian polynomials of orders 0, 1 and 2, referred to as P_k, $k \in 0, 1, 2$ in the following (Brossier; 2009; Etienne et al.; 2009). Let's remind that our motivation behind seismic modelling is to perform seismic imaging of the subsurface by full waveform inversion, the spatial resolution of which is half the propagated wavelength and that the physical properties of the medium are piecewise constant per element in our implementation of the DG method. The spatial resolution of the FWI and the piecewise constant representation of the medium direct us towards low-interpolation orders to achieve the best compromise between computational efficiency, solution accuracy and suitable discretization of the computational domain. The P_0 interpolation (or finite volume scheme) was shown to provide sufficiently-accurate solution on 2D equilateral triangular mesh when 10 cells per minimum propagated wavelength are used (Brossier et al.; 2008), while 10 cells and 3 cells per propagated wavelengths provide sufficiently-accurate solutions on unstructured triangular meshes with the P_1 and the P_2 interpolation orders, respectively (Brossier; 2009). Of note, the P_0 scheme is not convergent on unstructured meshes when centered fluxes are used (Brossier et al.; 2008). This prevents the use of the P_0 scheme in 3D medium where uniform tetrahedral meshes do not exist (Etienne et al.; 2008). A second remark is that the finite volume scheme on square cell is equivalent to second-order accurate FD stencil (Brossier et al.; 2008) which is consistent with a discretization criterion of 10 grid points per wavelength (Virieux; 1986). Use of interpolation orders greater than 2 would allow us to use coarser meshes for the same accuracy but these coarser meshes would lead to an undersampling of the subsurface model during imaging. On the other hand, use of high interpolation orders on mesh built using a criterion of 4 cells par wavelength would provide an unnecessary accuracy level for seismic imaging at the expense of the computational cost resulting from the dramatic increase of the number of unknowns in the equation 32.
The computational cost of the LU decomposition depends on the numerical bandwidth of the matrix, the dimension of the matrix (i.e., the number of rows/columns) and the number of non-zero coefficients per row (n_z). The dimension of the matrix depends in turn of the number of cell (n_{cell}), of the number of nodes per cell (n_d) and the number of wavefield components (n_{wave}) (3 in 2D and 4 in 3D). The number of nodes in a 2D triangular and 3D tetrahedral element is given by Hesthaven and Warburton (2008):

$$2D\ mesh: n_d = \frac{(k+1)(k+2)}{2},\quad 3D\ mesh: n_d = \frac{(k+1)(k+2)(k+3)}{6}, \tag{33}$$

where k denotes the interpolation order (Figure 10).

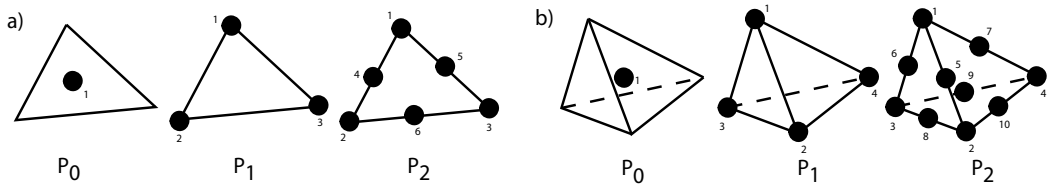


Fig. 10. Number of P_0, P_1, P_2 nodes in a triangular (a) and tetrahedral (b) element.

The numerical bandwidth is not significantly impacted by the interpolation order. The dimension of the matrix and the number of non zero elements per row of the impedance matrix are respectively given by $n_{wave} \times n_d \times n_{cell}$ and $(1+n_{neigh}) \times n_d \times n_{der} + 1$, where n_{neigh} is the number of neighbor cells (3 in 2D and 4 in 3D) and n_{der} is the number of wavefield components involved in the r.h.s of the velocity-pressure wave equation, equation 20. Table 4 outlines the number of non zero coefficients per row for the mixed-grid FD and DG methods. Increasing the interpolation order will lead to an increase of the number of non zero coefficients per row, a decrease of the number of cells in the mesh and an increase of the number of nodes in each element. The combined impact of the 3 parameters n_z, n_{cell}, n_d on the computational cost of the DG method makes difficult the definition of the optimal discretization of the frequency-domain DG method. The medium properties should rather drive us towards the choice of a suitable discretization. To illustrate this issue, we perform a numerical experiment with two end-member models composed of an infinite homogeneous and a two-layer model with a sharp velocity contrast at the base of a thin low-velocity layer. Both models have the same dimension (4 km x 4 km). The top layer of the two-layer model has a thickness of 400 m and a wave speed of 300 m/s, while the bottom layer has a wave speed of 1.5 km/s. During DG modelling, the models were successively discretized with 10 cells per minimum wavelength on an equilateral mesh for the P_0 interpolation, 10 cells per local wavelength on unstructured triangular mesh for the P_1 interpolation and 3 cells per local wavelength on unstructured triangular mesh for the P_2 interpolation. A fourth simulation was performed where P_1 interpolation is applied in the top layer while P_0 interpolation is used in the bottom layer. Table 5 outlines the time and memory requirement of the LU factorization and multi-r.h.s solve for the FD and DG methods. Among the different DG schemes, the P_2 scheme is the most efficient one in terms of computational time and memory for the two-layer model. This highlights the benefit provided by the decreasing of the number of elements in the mesh resulting from the h adaptivity coupled with a coarse discretization criterion of 3 cells per local wavelength. The mixed P_0-P_1 scheme performs reasonably well in the two-layer model although it remains less efficient than the P_2 scheme. In contrast, the performances of the P_0 and P_2 schemes are of the same order in the homogeneous model. This highlights that P_2 scheme does not provide any benefit if the h adaptivity is not required. The P_1 scheme is the less efficient one in homogeneous media because it relies on the same discretization criterion than the P_0 scheme but involves an increasing number of nodes per element. As expected, the FD method is the most efficient one in the homogeneous model thanks to the parsimonious formulation which

	FD^{2D}	$DG^{2D}_{P_0}$	$DG^{2D}_{P_1}$	$DG^{2D}_{P_2}$	FD^{3D}	$DG^{3D}_{P_0}$	$DG^{3D}_{P_1}$	$DG^{3D}_{P_2}$
n_d	1	1	3	6	1	1	4	10
n_z	9	5-9	13-25	24-48	27	6-16	21-61	51-151

Table 4. Number of nodes per element (n_d) and number of non-zero coefficients per row of the impedance matrix (n_z) for the FD and DG methods. Left: 2D case; Right: 3D case. n_z depends on the number of wavefield components involved in the r.h.s of the first-order wave equation, n_{der}, unlike the parsimonious FD method applied to the second-order wave equation.

Test	Resource	P_0	P_1	P_0-P_1	P_2	FD
Homog.	Cell/point numbers	113 097	136 724	116 363	12 222	9 604
	Degrees of freedom	339 291	1 230 516	417 477	219 996	9 604
	T_{LU} (s)	0.7	8.5	0.8	1.5	0.16
	Mem_{LU} (Gb)	1.34	5.84	1.62	1.49	0.1
	T_s (s)	11.6	40.9	13.6	7.2	0.5
Two-lay.	Cell/point numbers	2 804 850	291 577	247 303	32 664	232 324
	Degrees of freedom	8 414 550	2 624 193	1 416 243	587 952	232 324
	T_{LU} (s)	57.5	15.0	6.4	3.4	1.3
	Mem_{LU}(Gb)	31.68	11.44	5.58	3.02	1.18
	T_s (s)	274.3	83.3	46.8	18.9	2.7

Table 5. Computational ressources required for the forward problem solved with DGs P_0, P_1, P_0-P_1 and P_2 and optimized FD method in two simples cases, on 16 processors.
Nomenclature: *Homog*: homogeneous model. *Two − lay*: two-layer model. T_{LU}: time for LU factorization. Mem_{LU}: memory required by LU factorization. T_s: time for 116 r.h.s solve.

involves only the pressure wavefield and the optimized discretization criterion of 4 grid points per wavelength. The time and memory costs of the FD and P_2-DG methods are of the same order in the two-layer model. However, the P_2-DG method will be the method of choice as soon as sharp boundaries of arbitrary geometries will be present in the model due to the geometrical flexibility provided by the unstructured triangular mesh.

4.3 Boundary conditions and source implementation

Absorbing boundary conditions are implemented with unsplitted PML in the frequency-domain DG method (Brossier; 2009) following the same approach than for the FD method (see section *PML absorbing boundary conditions*). Free surface boundary condition is implemented with the method of image. A ghost cell is considered above the free surface with the same velocity and the opposite pressure components to those below the free surface. This allows us to fulfill the zero pressure condition at the free surface while keeping the correct numerical estimation of the particle velocity at the free surface. Using these particle velocities and pressures in the ghost cell, the pressure flux across the free surface interface vanishes, while the velocity flux is twice the value that would have been obtained by neglecting the flux contribution above the free surface (see equation 26). As in the FD method, this boundary condition has been implemented by modifying the impedance matrix accordingly without introducing explicitely the ghost element in the mesh. The rigid boundary condition is implemented following the same principle except that the same pressure and the opposite velocity are considered in the ghost cell. Concerning the source

excitation, the point source at arbitrary positions in the mesh is implemented by means of the Lagrange interpolation polynomials for $k \geq 1$. This means that the source excitation is performed at the nodes of the cell containing the source with appropriate weights corresponding to the projection of the physical position of the source on the polynomial basis. When the source is located in the close vicinity of a node of a triangular cell, all the weights are almost zero except that located near the source. In the case of the P_2 interpolation, a source close to the vertex of the triangular cell is problematic because the integral of the P_2 basis function over the volume of the cell is zero for nodes located at the vertex of the triangle. In this case, no source excitation will be performed (see equation 29). To overcome this problem specific to the P_2 interpolation, one can use locally a P_1 interpolation in the element containing the source at the expense of the accuracy or distribute the source excitation over several elements or express the solution in the form of local polynomials (i.e., the so-called modal form) rather than through nodes and interpolating Lagrange polynomials (i.e., the so-called nodal form). Another issue is the implementation of the source in P_0 equilateral mesh. If the source is excited only within the element containing the source, a checker-board pattern is superimposed on the wavefield solution. This pattern results from the fact that one cell out of two is excited in the DG formulation because the DG stencil does not embed a staggered-grid structure (the unexcited grid is not stored in staggered-grid FD methods; see Hustedt et al. (2004) for an illustration). To overcome this problem, the source can be distributed over several elements of the mesh or P_1 interpolation can be used in the area containing the sources and the receivers, while keeping P_0 interpolation in the other parts of the model (Brossier et al.; 2008). Of note, use of unstructured meshes together with the source excitation at the different nodes of the element contribute to mitigate the checker-board pattern in the in P_1 and P_2 schemes. The same procedure as for the source is used to extract the wavefield solution at arbitrary receiver positions.

4.4 Numerical examples

We present below two applications involving highly-contrasted media where the DG method should outperform the FD method thanks to the geometric flexibility provided by unstructured triangular or tetrahedral meshes to implement boundary conditions along interfaces of arbitrary shape.

4.4.1 Acoustic wave modelling in presence of cavities

We design a model that mimics a perfect 2D oceanic waveguide of dimension 20 000 m x 2 000 m. Applications of modelling ocean waveguide are for instance acoustic imaging of the oceanic currents, continuous monitoring of fish populations and localization of scattering sources. A free surface and a rigid surface explicit boundary conditions are implemented on the top and on the bottom of the water column to mimic the sea surface and the sea floor, respectively. A pressure source, located at position (x = 1000m; z = 1000m), propagates the direct wave in the homogeneous water layer as well as waves which are multi-reflected from the top and the bottom boundaries. Result of the simulation with the DG-P_2 scheme at 10 Hz is shown in Figure 11a. In a second simulation, we added a circular cavity of diameter 400 m in the center of the waveguide. A free surface boundary condition is implemented along the contour of the cavity. The unstructured triangular meshing around the cavity allows for an accurate discretization of the circular cavity (Figure 12).

Simulation in the waveguide with the cavity is shown in Figure 11b. Comparison with the simulation performed in the homogeneous waveguide (Figure 11a) highlights the strong interaction between the multi-reflected wavefield with the scattering source and the intrinsic non linearity of oceanic imaging resulting from complex wavepaths in the water column.

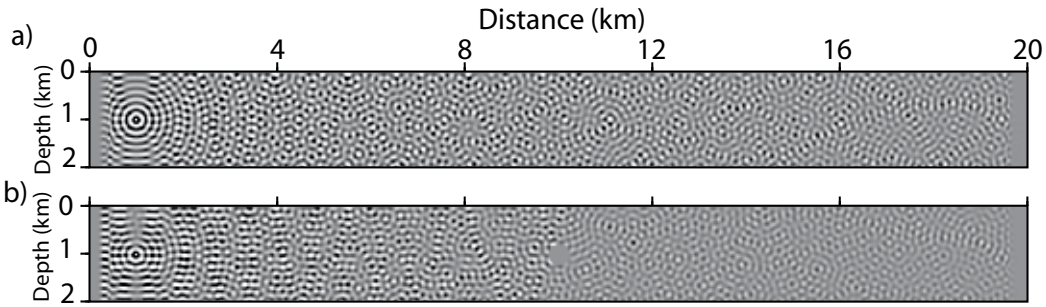


Fig. 11. Pressure wavefield in the oceanic waveguide without (a) and with (b) a circular cavity in the water column. Note that two 500-m layers of PML absorbing conditions are implemented at the two ends of the model.

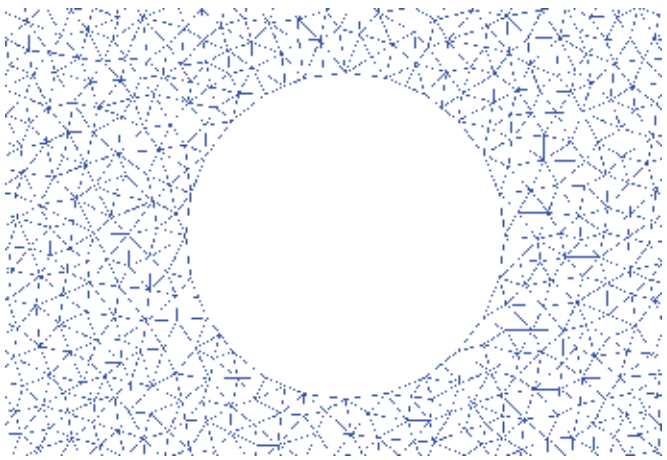


Fig. 12. Wave guide - Cavity model mesh: zoom on the cavity position.

4.4.2 Acoustic wave modelling in galleries

A second potential application of the DG method is the modelling of the air/solid contact in the framework of blast reduction in acoustic problems. The selected target illustrates the impact of the gallery design on blast reduction with application to military safety. The gallery geometry is delineated by the solid black lines in Figure 14. Due to the high wave speed contrast between the air and the solid, an adaptive mesh with a mesh refinement in the air layer was designed to minimize the number of degrees of freedom in the DG simulation (Figure 13). Figure 14(a-c) shows the horizontal velocity wavefield at the frequencies 50, 100 and 200 Hz resulting from an explosive source located near the entrance of the gallery. The wavefield in the main gallery is clearly attenuated thanks to the anti-blast first gallery and the multiple angles which hinders energy propagation.

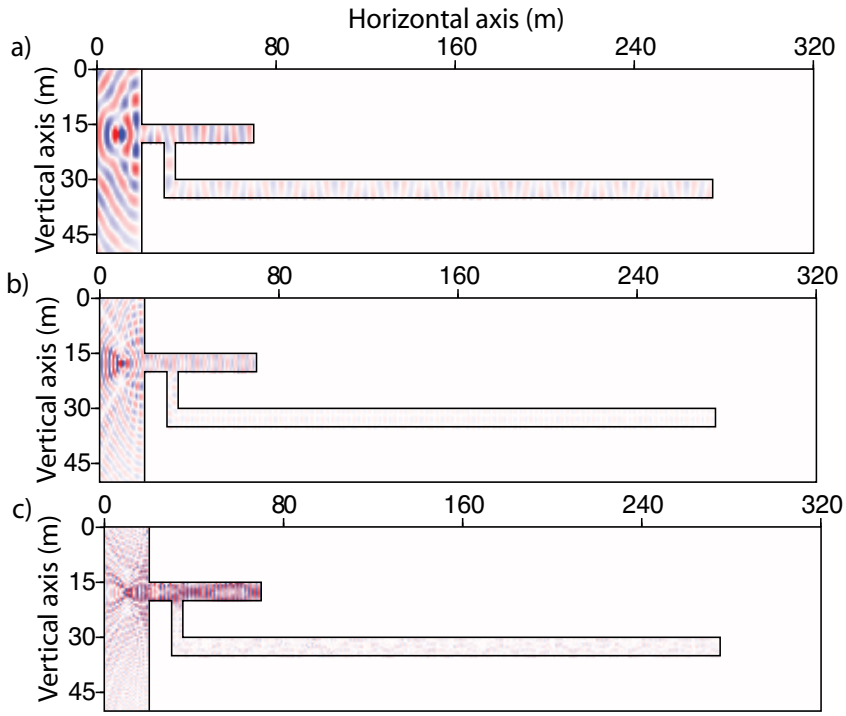


Fig. 13. (a) Gallery model geometry. Real part of the horizontal velocity wavefield at frequencies (b) 50 Hz, (c) 100 Hz and (d) 200 Hz.

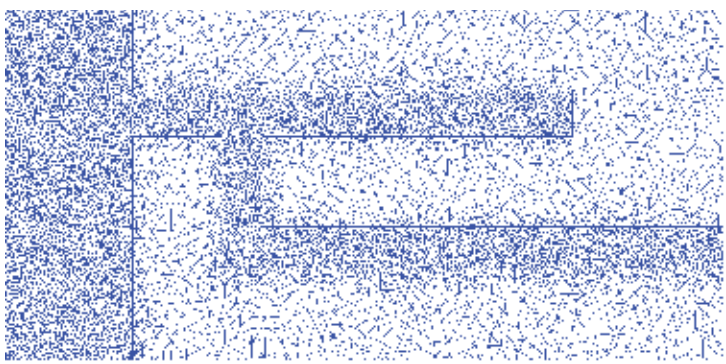


Fig. 14. Zoom on the gallery model mesh. Note the size of cells adapted to local wavespeed.

5. Conclusion and perspectives

We have reviewed two end-member numerical methods to perform visco-acoustic wave modelling in the frequency domain with sparse direct solvers. Two benefits of the frequency domain compared to the time domain are the straightforward and inexpensive implementation of attenuation effects by means of complex-valued wave speeds and the computational efficiency of multi-source modelling when a sparse direct solver is used to solve the linear system resulting from the discretization of the wave equation in the

frequency domain. The first discretization method relies on a parsimonious staggered-grid FD method based on a compact and accurate stencil allowing for both the minimization of the numerical bandwidth of the impedance matrix and the number of unknowns in the FD grid. The discretization criterion which can be used with this method is 4 grid points per minimum wavelength. We have shown the efficiency of the method for tackling 3D problems involving few millions of unknowns and few thousands of right-hand sides on computational platform composed of a limited number of processors with a large amount of shared memory. Since the FD lacks geometrical flexibility to discretize objects of complex geometries, we have developed a 2D discontinuous finite element method on unstructured triangular mesh. The DG method is fully local in the sense that each element is uncoupled from the next, thanks to the duplication of variables at nodes shared by two neighboring elements. This uncoupling allows for a flexible implementation of the so-called $h - p$ adaptivity, where the size of the element can be adapted to the local features of the model and the order of the interpolating polynomials can be adapted within each element. The price to be paid for the geometrical flexibility of the discretization is the increase of the number of unknowns compared to continuous finite element methods. We have illustrated the fields of application where the frequency-domain DG method should perform well. A first perspective of this work concerns the investigation of other linear algebra techniques to solve the linear system and overcome the limits of sparse direct solver in terms of memory requirement and limited scalability. Use of domain decomposition methods based on hybrid direct-iterative solvers should allow us to tackle 3D problems of higher dimensions. A second perspective is the improvement of the frequency-domain DG method to make possible the extension to 3D. One possible improvement is the use of heterogeneous medium properties in each element of the mesh to allow for higher-order interpolation orders. Another field of investigation concerns the numerical flux, which is a central ingredient of the DG method. Although we used centered fluxes for their energy conservation properties, other fluxes such as upwind fluxes should be investigated for improved accuracy of the scheme.

6. Acknowledgments

This study was partly funded by the SEISCOPE consortium (http://seiscope.oca.eu), sponsored by BP, CGG-Veritas, ENI, Exxon-Mobil, Shell and Total. The linear systems were solved with the MUMPS direct solver (http://graal.enslyon. fr/MUMPS). The mesh generation in DG modelling was performed with Triangle (http://www.cs.cmu.edu/ quake/triangle.html). Fill-reducing ordering was performed with METIS (http://glaros.dtc.umn.edu/gkhome/views/metis). Access to the high-performance computing facilities of SIGAMM (http://crimson.oca.eu), CINES (http://www.cines.fr) and CIMENT (OSUG) computer centers provided the necessary computer ressources. We would like to thank also Pr. G. Nolet (Geoazur) for access to the Thera cluster.

7. References

Amestoy, P. R., Guermouche, A., L'Excellent, J. Y. and Pralet, S. (2006). Hybrid scheduling for the parallel solution of linear systems, *Parallel Computing* 32: 136–156.

Aminzadeh, F., Brac, J. and Kunz, T. (1997). *3-D Salt and Overthrust models*, SEG/EAGE 3-D Modelling Series No.1.

Ben-Hadj-Ali, H., Operto, S. and Virieux, J. (2008). Velocity model building by 3D frequency-domain, full-waveform inversion of wide-aperture seismic data, *Geophysics* 73(5): VE101–VE117.

BenJemaa, M., Glinsky-Olivier, N., Cruz-Atienza, V. M. and Virieux, J. (2009). 3D Dynamic rupture simulations by a finite volume method, *Geophys. J. Int.* 178: 541–560.

BenJemaa, M., Glinsky-Olivier, N., Cruz-Atienza, V. M., Virieux, J. and Piperno, S. (2007). Dynamic non-planar crack rupture by a finite volume method, *Geophys. J. Int.* 171: 271– 285.

Berenger, J.-P. (1994). A perfectly matched layer for absorption of electromagnetic waves, *Journal of Computational Physics* 114: 185–200.

Boore, D. M. (1972). Finite-difference methods for seismic wave propagation in heterogeneous materials, *in* B. B. A. Ed. (ed.), *Methods in computational physics*, Vol. 11, Academic Press, Inc.

Brossier, R. (2009). *Imagerie sismique à deux dimensions des milieux visco-élastiques par inversion des formes d'onde: développements méthodologiques et applications.*, PhD thesis, Université de Nice-Sophia-Antipolis.

Brossier, R., Virieux, J. and Operto, S. (2008). Parsimonious finite-volume frequency-domain method for 2-D P-SV-wave modelling, *Geophys. J. Int.* 175(2): 541–559.

Chaljub, E., Komatitsch, D., Vilotte, J.-P., Capdeville, Y., Valette, B. and Festa, G. (2007). Spectral element analysis in seismology, *in* R.-S. Wu and V. Maupin (eds), *Advances in Wave Propagation in Heterogeneous Media*, Vol. 48 of *Advances in Geophysics*, Elsevier - Academic Press, pp. 365–419.

Chew, W. C. and Weedon, W. H. (1994). A 3-D perfectly matched medium from modified Maxwell's equations with stretched coordinates, *Microwave and Optical Technology Letters* 7: 599–604.

Collino, F. and Monk, P. (1998). Optimizing the perfectly matched layer, *Computer methods in Applied Mechanics and Engineering* 164: 157–171.

Collino, F. and Tsogka, C. (2001). Application of the perfectly matched absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media, *Geophysics* 66: 294–307.

Dablain, M. A. (1986). The application of high-order differencing to the scalar wave equation, *Geophysics* 51: 54–66.

de la Puente, J., Ampuero, J.-P. and Käser, M. (2009). Dynamic Rupture Modelling on Unstructured Meshes Using a Discontinuous Galerkin Method, *J. Geophys. Res.* 114: B10302.

Demmel, J.W. (1997). *Applied numerical linear algebra*, SIAM, Philadelphia.

Dolean, V., Fol, H., Lanteri, S. and Perrussel, R. (2008). Solution of the time-harmonic Maxwell equations using discontinuous Galerkin methods, *J. Comput. Appl. Math.* 218(2): 435–445.

Dolean, V., Lanteri, S. and Perrusel, R. (2007). A domain decomposition method for solving the three-dimensional time-harmonic Maxwell equations discretized by discontinuous Galerkin methods, *J. Comput. Phys.* 227(3): 2044–2072.

Drossaert, F. H. and Giannopoulos, A. (2007). A nonsplit complex frequency-shifted PML based on recursive integration for FDTD modelling of elastic waves, *Geophysics* 72(2): T9–T17.

Duff, I. S., Erisman, A. M. and Reid, J. K. (1986). *Direct methods for sparse matrices*, Clarendon Press, Oxford, U. K.

Duff, I. S. and Reid, J. K. (1983). The multifrontal solution of indefinite sparse symmetric linear systems, *ACM Transactions on Mathematical Software* 9: 302–325.

Dumbser,M. and K¨aser,M. (2006). An Arbitrary High Order Discontinuous Galerkin Method for ElasticWaves on Unstructured Meshes II: The Three-Dimensional Isotropic Case, *Geophysical Journal International* 167(1): 319–336.

Erlangga, Y. A. and Nabben, R. (2008). Multilevel projection-based nested krylov iteration for boundary value problems, *SIAM - J. Scientific Computing* 30(3): 1572–1595.

Erlangga, Y. A., Vuik, C. and Oosterlee, C. (2006). A novel multigrid based preconditioner for heterogeneous Helmholtz problems, *SIAM - Journal of Scientific Computing* 27: 1471–1492.

Etienne, V., Brossier, R., Operto, S. and Virieux, J. (2008). A 3D Parsimonious Finite-Volume Frequency-Domain Method for ElasticWave Modelling, *Expanded Abstracts*, EAGE.

Etienne, V., Virieux, J. and Operto, S. (2009). A massively parallel time domain discontinuous Galerkin method for 3D elastic wave modelling, *SEG Technical Program Expanded Abstracts, Houston* 28(1): 2657–2661.

Faccioli, E. F., Paolucci, R. and Quarteroni, A. (1997). 2D and 3D elastic wave propagation by a pseudo-spectral domain decomposition method, *J. Seismol.* 1: 237–251.

Futterman,W. (1962). Dispersive body waves, *Journal of Geophysics Research* 67: 5279 5291.

Guermouche, A., L'Excellent, J. Y. and Utard, G. (2003). Impact of reordering on the memory of a multifrontal solver, *Parallel computing* 29: 1191–1218.

Haidar, A. (2008). *On the parallel scalability of hybrid linear solvers for large 3D problems*, PhD thesis, Institut National Polytechnique de Toulouse - CERFACS.

Hesthaven, J. S. and Warburton, T. (2008). *Nodal Discontinuous Galerkin Method. Algorithms, Analysis, and Application,* Springer.

Hicks, G. J. (2002). Arbitrary source and receiver positioning in finite-difference schemes using kaiser windowed sinc functions, *Geophysics* 67: 156–166.

Holberg, O. (1987). Computational aspects of the choice of operators and sampling interval for numerical differentiation in large-scale simulation of wave phenomena, *Geophys. Prospecting* 35: 629–655.

Hustedt, B., Operto, S. and Virieux, J. (2004). Mixed-grid and staggered-grid finite difference methods for frequency domain acoustic wave modelling, *Geophys. J. Int.* 157: 1269–1296.

Jensen, F. B., Kuperman,W. A., Porter, M. B. and Schmidt, H. (eds) (1994). *Computational ocean acoustics.*, AIP series in modern Acoustics and signal processing.

Jo, C. H., Shin, C. and Suh, J. H. (1996). An optimal 9-point, finite-difference, frequency-space 2D scalar extrapolator, *Geophysics* 61: 529–537.

Karypis, G. and Kumar, V. (1999). A fast and high quality multilevel scheme for partitioning irregular graphs, *SIAM Journal on Scientific Computing* 20(1): 359 – 392.

Kolsky, H. (1956). The propagation of stress pulses in viscoelastic solids, *Philosophical Magazine*1: 693–710.

Komatitsch, D. and Martin, R. (2007). An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation, *Geophysics* 72(5): SM155– SM167.

Komatitsch, D. and Vilotte, J. P. (1998). The spectral element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures, *Bull. Seismol. Soc. Am.* 88: 368–392.

Kurzak, J. and Dongarra, J. (2006). Implementation of the mixed-precision high performance LINPACK benchmark on the CELL processor, *Technical report ut-cs-06-580, University of Tennessee. http://icl.cs.utk.edu/iter-ref/* .

Kuzuoglu, M. and Mittra, R. (1996). Frequency dependence of the constitutive parameters of causal perfectly matched anisotropic absorbers, *IEEE Microwave and Guided Wave Letters* 6: 447–449.

Langou, J., Luszczek, P., Kurzak, J., Buttari, A., and Dongarra, J. (2006). LAPACK working note 175: exploiting the performance of 32 bit floating point arithmetic in obtaining 64 bit accuracy, *Technical report, University of Tennessee. http://icl.cs.utk.edu/iter-ref/* .

LeVeque, R. J. (2002). *Finite Volume Methods for Hyperbolic Problems*, Cambridge Texts in Applied Mathematics.

Liu, J. W. H. (1992). The multifrontal method for sparse matrix solution: theory and practice, *SIAM review* 34(1): 82–109.

Lombard, B. and Piraux, J. (2004). Numerical treatment of two-dimensional interfaces for acoustic and elastic waves, *J. Compu. Physics* 195: 90–116.

Lombard, B., Piraux, J., Gelis, C. and Virieux, J. (2008). Free and smooth boundaries in 2-D finite-difference schemes for transient elastic waves, *Geophys. J. Int.* 172: 252–261.

Luo, Y. and Schuster, G. T. (1990). Parsimonious staggered grid finite-differencing of the wave equation, *Geophysical Research Letters* 17(2): 155–158.

Marfurt, K. (1984). Accuracy of finite-difference and finite-elements modelling of the scalar and elastic wave equation, *Geophysics* 49: 533–549.

Mattsson, K., Ham, F. and Iaccarino, G. (2009). Stable boundary treatment for the wave equation on second-order form, *J. Sci. Comput.* 41(3): 366–383.

Moczo, P., Kristek, J., Vavrycuk, V., Archuleta, R. and Halada, L. (2002). 3D heterogeneous staggered-grid finite-difference modelling of seismic motion with volume harmonic and arithmetic averaging of elastic moduli and densities, *Bull. Seismol. Soc. Am.* 92: 3042–3066.

MUMPS-team (2009). *MUMPS - MUltifrontal Massively Parallel Solver users' guide - version 4.9.2 (November 2009)*, ENSEEIHT-ENS Lyon, http://www.enseeiht.fr/apo/MUMPS/ or http://graal.ens-lyon.fr/MUMPS.

Operto, S., Virieux, J., Amestoy, P., L' Éxcellent, J.-Y., Giraud, L. and Ben-Hadj-Ali, H. (2007). 3D finite-difference frequency-domain modelling of visco-acoustic wave propagation using a massively parallel direct solver: A feasibility study, *Geophysics* 72(5): SM195– SM211.

Operto, S., Virieux, J., Ribodetti, A. and Anderson, J. E. (2009). Finite-difference frequency-domain modelling of visco-acoustic wave propagation in two-dimensional TTI media, *Geophysics* 74 (5): T75–T95.

Pitarka, A. (1999). 3D elastic finite-difference modelling of seismic motion using staggered grids with nonuniform spacing, *Bull. Seism. Soc. Am.* 89(1): 54–68.

Plessix, R. E. (2007). A Helmholtz iterative solver for 3D seismic-imaging problems, *Geophysics* 72(5): SM185–SM194.

Priolo, E., Carcione, J. M. and Seriani, G. (1994). Numerical simulation of interface waves by high-order spectral modelling techniques, *J. acoust. Soc. Am.* 95: 681–693.

Remaki, M. (2000). A new finite volume scheme for solving Maxwell's system, *COMPEL* 19(3): 913–931.

Riyanti, C. D., Erlangga, Y. A., Plessix, R. E., Mulder, W. A., Vuik, C. and Oosterlee, C. (2006). A new iterative solver for the time-harmonic wave equation, *Geophysics* 71(E): 57–63.

Riyanti, C. D., Kononov, A., Erlangga, Y. A., Vuik, C., Oosterlee, C., Plessix, R. E. and Mulder, W. A. (2007). A parallel multigrid-based preconditioner for the 3D heterogeneous high-frequency Helmholtz equation, *Journal of Computational physics* 224: 431–448.

Roden, J. A. and Gedney, S. D. (2000). Convolution PML (CPML): An efficient FDTD implementation of the CFS-PML for arbitrary media, *Microwave and Optical Technology Letters* 27(5): 334–339.

Saad, Y. (2003). *Iterative methods for sparse linear systems,* SIAM, Philadelphia.

Saenger, E. H., Gold, N. and Shapiro, S. A. (2000). Modelling the propagation of elastic waves using a modified finite-difference grid, *Wave motion* 31: 77–92.

Sen, M. K. and Stoffa, P. L. (1995). *Global Optimization Methods in Geophysical Inversion*, Elsevier Science Publishing Co.

Seriani, G. and Priolo, E. (1994). Spectral element method for acoustic wave simulation in heterogeneous media, *Finite elements in analysis and design* 16: 337–348.

Sirgue, L., Etgen, J. T. and Albertin, U. (2008). 3D Frequency Domain Waveform Inversion using Time Domain Finite Difference Methods, *Proceedings 70th EAGE, Conference and Exhibition, Roma, Italy*, p. F022.

Smith, B. F., Bjørstad, P. E. and Gropp, W. (1996). *Domain decomposition: parallel multilevel methods for elliptic partial differential equations,* Cambridge University Press.

Sourbier, F., Haidar, A., Giraud, L., Operto, S. and Virieux, J. (2008). Frequency-domain full-waveform modelling using a hybrid direct-iterative solver based on a parallel domain decomposition method: A tool for 3D full-waveform inversion?, *SEG Technical Program Expanded Abstracts* 27(1): 2147–2151.

Stekl, I. and Pratt, R. G. (1998). Accurate viscoelastic modelling by frequency-domain finite difference using rotated operators, *Geophysics* 63: 1779–1794.

Taflove, A. and Hagness, C. (2000). *Computational electrodynamics: the finite-difference time-domaine method*, Artech House, London, United Kindom.

Takeuchi, N. and Geller, R. J. (2000). Optimally accurate second-order time-domain finite-difference scheme for computing synthetic seismograms in 2-D and 3-D media, *Phys. Earth planet. Inter.* 119: 99–131.

Virieux, J. (1986). P-SV wave propagation in heterogeneous media, velocity-stress finite difference method, *Geophysics* 51: 889–901.

Virieux, J. and Operto, S. (2009). An overview of full waveform inversion in exploration geophysics, *Geophysics* 74(6): WCC127–WCC152.

Virieux, J., Operto, S., Ben-Hadj-Ali, H., Brossier, R., Etienne, V., Sourbier, F., Giraud, L. and Haidar, A. (2009). Seismic wave modelling for seismic imaging, *The Leading Edge* 28(5): 538–544.

Yee, K. S. (1966). Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, *IEEE Trans. Antennas and Propagation* 14: 302–307.

8

Shear Elastic Wave Refraction on a Gap between Piezoelectric Crystals with Uniform Relative Motion

Nick Shevyakhov[1] and Sergey Maryshev[2]
[1]*Kotel'nikov Institute of Radio Engineering and Electronics of Russian Academy of Sciences, Ul'yanovsk Branch*
[2]*Moscow Institute of Physics and Technology*
Russia

1. Introduction

The ability of acoustic waves to pass through a slot or vacuum (air) gap between piezoelectric crystals which are not contacting with each other, for the first time has noted Kaliski (Kaliski, 1966). In those years the interest of a researchers to the phenomenon first of all was connected to the being widely discussed problem of generation and amplification of ultrasonic waves that propagate along boundary of a piezoelectric medium adjoining semi-conductor with a drift current (Gulyaev, 2005). More important it turned out to use Kaliski's idea of acoustic wave passage through a gap of piezoelectric crystals in metrological purposes. Firstly for development of contactless measurements of electro-acoustic fields in crystals, at which are excluded or minimized the distortions caused by own loading action of transducer. Secondly for search of effective ways of contactless excitation of acoustic oscillations in solids.

Because of anisotropy and of weakly expressed transversal piezoelectricity the case of cubic piezoelectric crystals that has considered Kaliski for shear waves of horizontal polarization (Kaliski, 1966) have not shown the proper efficiency of wave passage through a gap even with very small thickness and under condition of almost sliding incidence. Therefore in subsequent this phenomenon due to similarity to tunnel transition in the quantum mechanics (Landau & Lifshitz, 1991) named by acoustic tunneling (Balakirev & Gilinskii, 1982), began to be considered for more suitable crystals of tetragonal and hexagonal systems. The being reviewed cycle of investigations for case of strictly plane boundaries (a Balakirev & Gorchakov, 1977; Balakirev & Gorchakov, 1986), was finished (Balakirev et al., 1978) by experimental detection of effect.

By common result of the quoted works was the conclusion that the efficiency of acoustic tunneling is caused essentially by electromechanical coupling factor of crystals $\mathcal{K}$, and with growth of thickness of a gap is very decreasing. The passage of an acoustic wave through a gap will be especially appreciable at angle of incidence $\alpha \sim \pi/2 - \mathcal{K}^2(1+\mathcal{K}^2)^{-1}$. So, even for such strong piezoelectric, as $BaTiO_3$, we have $\mathcal{K}^2 < 0.4$ (Royer & Dieulesaint, 2000) with a following from here estimation $\alpha > 75°$. Therefore the opportunity of acoustic tunneling to using is very being complicated.

The first attempt of overcoming this difficulty of practical realization of the phenomenon of acoustic tunneling was connected with known opportunity to control by coupling of acoustic and electric oscillation in crystal with high electrostriction by an external electrical field (Gulyaev, 1967; a Gulyaev & Plessky, 1977). In particular, for shear waves with horizontal polarization of displacement (SH-waves) that propagate across applied field, its action is similar to piezoelectricity of 6mm (4mm)-class crystals with piezoelectric modulus $e_{15}=aE_0/2$, where a is the coefficient of electrostriction, and E_0 is the strength of electric field. However the simple reproduction of the above results for this case takes place (Filippov, 1985) as the increase of piezoelectric activity of crystal with escalating of an electrical field almost up to a voltage of dielectric breakdown only a few decrease the suitable angles of incidence. The case, when the incident shear wave has vertical polarization and the external field is lying in a plane of incidence, also was considered by Filippov (Filippov, 1985). It is more interesting as in such conditions the acceptable for practical purposes angles of incidence can be lowered up to forty degrees.

Results received in (Filippov, 1985), have encouraged the researchers of the phenomenon of acoustic tunneling, but have not brought the complete satisfaction because of necessity in using a source of a high voltage. The attention has been addressed to other opportunity to increase the efficiency of tunneling of waves through a gap not only on intermediate angles of incidence, but also small ones. In its basis is laid the account of resonant properties of a gap as a waveguide of the slotted electroacoustic wave (Balakirev & Gorchakov, 1977 b; Gulyaev & Plessky, 1977 b). For achievement of declared object it was necessary to change resonant properties of a gap appreciably. As an effective way it was offered the using of piezocrystals with a periodic shape of surfaces (Gulyaev & Plessky, 1978) or with periodic inertial loading in form of guideway layer from other dielectric material (Gulyaev et al., 1978).

At a geometrical resonance of incident wave with the period of profile or loading impedance of boundaries the effective excitation of the appropriate mode of a slotted electroacoustic wave took place. In a result the complete passage through a gap, possible on conditions of excitation even at normal incidence, will be achieved. As in a gap there are two modes of slotted electroacoustic wave (Gulyaev & Plessky, 1977 b), for the given configuration of slotted structure it was possible to determine two frequencies ensuring for a wave the complete passage through a gap. In case of guidway boundary layers of an other dielectric with periodic inertial loading (Gulyaev et al., 1978) slotted electroacoustic waves of a gap are being replaced, as a matter of fact, by surface Love waves (Royer & Dieulesaint, 2000), which connect through a gap by an electrical field. The advantage of use of surface Love waves before slotted electroacoustic waves consists in much stronger boundary localization and, as a consequence, in their ability to form on appreciably smaller distances along guidway boundary. Due to this the resonant tunneling of waves through a gap "adjusted" on Love waves, can be carried out with the appreciably smaller apertures of an incident acoustic beam. Idea to take advantage of resonant properties of a gap for achievement of complete passage of a wave through a gap experimentally was realized in work (Grigor'evskii, 1987) for waves of vertical polarization, when resonant modes of a gap with a periodic profile of boundaries are surface Releigh-type waves. It is necessary to note, that in this experiment the passage of a wave through a gap of piezoelectrics with rectangular grooves was not quite complete. The authors have explained it by partial transformation in transversal waves of a longitudinal wave, which is falling normally on a gap with periodic grooves.

The concept of acoustic tunneling is successfully applied now to interpretation of transfer effects of wave disturbances between phononic crystals (Qui et al, 2005; Van Der Biest et al, 2005; Pennec et al, 2009). Strictly speaking, there is not here the obvious analogy to acoustic tunneling of waves through a slot between piezoelectric crystals because of absence of a vacuum gap. Instead of it in respect to tunneling phonons consider the forbidden zone of phononic crystals, and role of the electric-field coupling between piezoelectric crystals begin to play the allowed states, which arise in the forbidden zone because of infringements of Bragg interference requirements by discrepancy of the periods of lattices or by introduction of artificial defects of periodic structure of phononic crystals.

Other direction of modern researches of acoustic tunneling, which directly continues early works (Balakirev & Gorchakov, 1977 a; Balakirev & Gorchakov, 1986; Balakirev et al, 1978; Filippov, 1985), is connected with taking into account the relative longitudinal displacement of piezoelectric crystals divided by a gap. As we know, earlier in all works on acoustic tunneling in piezoelectric layered structures with a gap the crystals always relied fixed. There are, however, some reasons, on which the acoustic tunneling in conditions of relative longitudinal displacement of piezoelectric crystals represents doubtless interest. So, in practice the relative moving of bodies is one of the main occasions for using of contactless ways of introduction of acoustic oscillations. On the other hand, in sphere of high technologies (robotics, mechatronics) the important place occupies monitoring relative moving of elements of designs that, in particular, means development of sensor controls using piezoelectric effect. At last, the relative longitudinal displacement of piezoelectrics in slot-type structures is possible to consider and as the additional factor of the processing of signal information by standard means of acoustoelectronics.

Present article is written on materials of the publications (Gulyaev et al, 2007 a; Gulyaev et al, 2007 b; Maryshev & Shevyakhov, 2007), concerning only of case of shear waves of horizontal polarization in piezoelectric crystals of some classes of crystallographic symmetry with ideal plane boundaries of a gap, which are not subjected to any periodic impedance loading. The appropriate generalizations, for example, account not only electrical, but also magnetic connection of crystals (piezomagnetics) by fields through a gap are represented by matter of the nearest future. Confirming it we shall refer to work (Vilkov et al, 2009), in which tunneling of shear magnetoelastic waves through a gap of ferromagnetic crystals testing relative longitudinal displacement recently was considered.

2. Tunneling of a shear wave through a gap of piezoelectric crystals with relative longitudinal motion

2.1 Shear elastic wave in a moving crystal

The typical geometry of boundary problem of acoustic tunneling through a gap of pair piezoelectric crystals with relative longitudinal motion is submitted on Fig. 1. On it one of crystals (bottom) moves with the given constant velocity V, whereas other (top) is in rest. Generally crystals can differ in the parameters, have various orientations of crystallographic axes and belong to various classes of crystal symmetry. However, it is important, that falling on a gap on the part of immobile crystal the acoustic wave was a piezoactive wave, i.e. was accompanied in its deformations by electric fields, and the surfaces of crystals – boundaries of a gap, were not covered with metal electrodes.

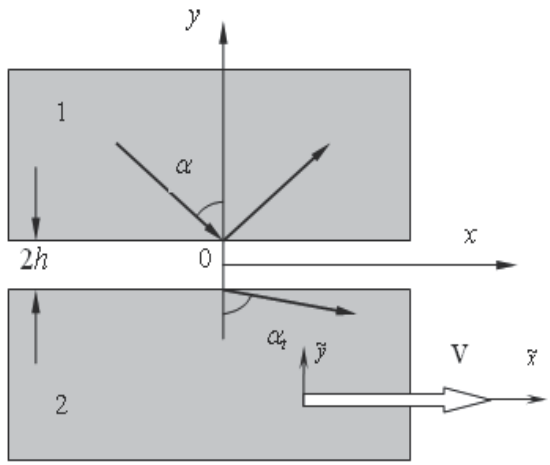


Fig. 1. Geometry of acoustic wave tunneling through a gap of piezoelectric crystals with relative longitudinal motion

From elementary reasons the following way of consideration of acoustic tunneling in conditions of relative longitudinal motion of piezoelectric crystals arises. It is necessary to connect with each of crystal own system of coordinates. For immobile crystal "1" it is the laboratory system of coordinates $x0yz$. For a moving crystal "2" it is the passing coordinate system $\tilde{x}0\tilde{y}\tilde{z}$. The propagation of acoustic waves in own coordinate systems of crystals, where both of them are immobile, is being described, obviously, by the standard manner (Balakirev & Gilinskii, 1982; Royer & Dieulesaint, 2000). However, because of coupling by electric fields through a gap it is impossible already to consider these wave processes as isolated ones and refraction of waves by a gap must be described in one common coordinate system. For this purpose any of own coordinate systems of crystals is suitable, but more preferably to use the laboratory system of coordinates, as according with logic of subject matter just with the one are connected the acoustic radiator and detector of a reflected wave. Thus, we need to describe waves, passing into a moving crystal, with a point of the observer of laboratory system of reference. In language of mathematics it means that in equations for moving crystal we transfer from coordinates $\tilde{x}, \tilde{y}, \tilde{z}$ to coordinates x, y, z.

If to accept reasonable restriction $V<<c$, where c is velocity of light, the transfer of disturbances by electric fields through a gap can rely instantaneous. It will be first and foremost is in accordance with usually used quasistatic approximation for determination of the electric fields, which accompany the acoustic waves in piezoelectric crystals (Balakirev & Gilinskii, 1982; Royer & Dieulesaint, 2000). Secondly, we then may be limited by mechanical relativity and to use for connection of coordinate systems $x0yz$ and $\tilde{x}0\tilde{y}\tilde{z}$ Galilean transformation

$$x = \tilde{x} + V\tilde{t}\ , \quad y = \tilde{y}\ , \quad z = \tilde{z}\ , \quad t = \tilde{t}\ . \tag{1}$$

Here $\tilde{t}$ and t is the time, which has equal duration in both systems of reference.

Let's consider identical piezoelectric crystals of a class 6 (4, 6mm, 4mm, ∞m) with common orientation of axes of symmetry of high order 6 (4), along coordinate directions z and $\tilde{z}$. In shear waves of horizontal polarization the elastic displacement u_j (here and everywhere are

lower j=1, 2 is the number of a crystal) also are parallel these directions: $\mathbf{u}_1=\mathbf{u}_1(x,y,t)\,||\,z$ and $\mathbf{u}_2=\mathbf{u}_2(\tilde{x},\tilde{y},\tilde{t})\,||\,\tilde{z}$. Therefore according to the equations of state for a piezoelectric material the working components of stress tensor T_{ik} and vector of an electrical induction $\mathbf{D}$ for a moving crystal have in the passing system of coordinates the form (Balakirev & Gilinskii, 1982; Royer & Dieulesaint, 2000)

$$T_{xz}^{(2)}=T_{zx}^{(2)}=\lambda\frac{\partial u_2}{\partial\tilde{x}}+e_{15}\frac{\partial\varphi_2}{\partial\tilde{x}}-e_{14}\frac{\partial\varphi_2}{\partial\tilde{y}},\quad T_{yz}^{(2)}=T_{zx}^{(2)}=\lambda\frac{\partial u_2}{\partial\tilde{y}}+e_{15}\frac{\partial\varphi_2}{\partial\tilde{y}}+e_{14}\frac{\partial\varphi_2}{\partial\tilde{x}}, \tag{2}$$

$$D_x^{(2)}=e_{15}\frac{\partial u_2}{\partial\tilde{x}}+e_{14}\frac{\partial u_2}{\partial\tilde{y}}-\varepsilon\frac{\partial\varphi_2}{\partial\tilde{x}},\quad D_y^{(2)}=e_{15}\frac{\partial u_2}{\partial\tilde{y}}-e_{14}\frac{\partial u_2}{\partial\tilde{x}}-\varepsilon\frac{\partial\varphi_2}{\partial\tilde{y}}. \tag{3}$$

The similar expressions, but already in laboratory system of coordinates, turn out for the immobile crystal. They directly follow from the formulas (2), (3), if in them to change number of a crystal j=2 for number j =1 and instead of coordinates $\tilde{x}$, $\tilde{y}$ to use accordingly laboratory coordinates x, y. In the formulas (2), (3) φ is the potential of an electrical field, λ is shear modulus, e_{15} and e_{14} is the piezoelectric modules of longitudinal and transverse piezoelectric effect, ε is the permittivity of a crystal.

By the initial equations for electroelastic fields of SH-waves propagating in a plane $\tilde{x}0\tilde{y}$ of a moving piezoelectric crystal, are the equations

$$\rho\frac{\partial^2 u_2}{\partial\tilde{t}^{\,2}}=\frac{\partial T_{ik}^{(2)}}{\partial\tilde{x}_k},\quad \frac{\partial D_i^{(2)}}{\partial\tilde{x}_i}=0, \tag{4}$$

where ρ is density. First of them represents the equation of elastic medium motion, and second expresses the fact of absence of free carriers of a charge in piezoelectric crystal and in quasi-static approximation with high accuracy replaces with itself complete system of the Maxwell equations for determination of an electrical field. For immobile piezoelectric in view of the above-stated replacement of number of a crystal and use of laboratory coordinates we have the similar equations. Let's remind that partial derivatives on spatial variables in the equations (4) are summarized on a repeating index, forming tensor convolutions.

Result of substitution of expressions (2), (3) in the equations (4) will be well known (Balakirev & Gilinskii, 1982; Royer & Dieulesaint, 2000) the equations of piezocrystal acoustics for waves of a SH-type

$$\rho\frac{\partial^2 u_2}{\partial\tilde{t}^{\,2}}=\lambda\tilde{\nabla}^2 u_2+e_{15}\tilde{\nabla}^2\varphi_2,\quad \frac{e_{15}}{\varepsilon}\tilde{\nabla}^2 u_2=\tilde{\nabla}^2\varphi_2. \tag{5}$$

Similarly for immobile crystal is received

$$\rho\frac{\partial^2 u_1}{\partial t^2}=\lambda\nabla^2 u_1+e_{15}\nabla^2\varphi_1,\quad \frac{e_{15}}{\varepsilon}\nabla^2 u_1=\nabla^2\varphi_1. \tag{6}$$

The difference between the equations (5), (6) is defined by differences in pairs of differential operators: $\partial/\partial\tilde{t}$, $\partial/\partial t$ and $\tilde{\nabla}^2=\partial^2/\partial\tilde{x}^2+\partial^2/\partial\tilde{y}^2$, $\nabla^2=\partial^2/\partial x^2+\partial^2/\partial y^2$. By rule of indirect differentiation of functions with many variables the connection between them it is possible to open, using relations

$$\frac{\partial}{\partial\tilde{x}}=\left(\frac{\partial x}{\partial\tilde{x}}\right)\frac{\partial}{\partial x}+\left(\frac{\partial y}{\partial\tilde{x}}\right)\frac{\partial}{\partial y}+\left(\frac{\partial t}{\partial\tilde{x}}\right)\frac{\partial}{\partial t},$$
$$\frac{\partial}{\partial\tilde{y}}=\left(\frac{\partial x}{\partial\tilde{y}}\right)\frac{\partial}{\partial x}+\left(\frac{\partial y}{\partial\tilde{y}}\right)\frac{\partial}{\partial y}+\left(\frac{\partial t}{\partial\tilde{y}}\right)\frac{\partial}{\partial t}, \qquad (7)$$
$$\frac{\partial}{\partial\tilde{t}}=\left(\frac{\partial x}{\partial\tilde{t}}\right)\frac{\partial}{\partial x}+\left(\frac{\partial y}{\partial\tilde{t}}\right)\frac{\partial}{\partial y}+\left(\frac{\partial t}{\partial\tilde{t}}\right)\frac{\partial}{\partial t}.$$

As from (1) follows, that $\frac{\partial x}{\partial\tilde{x}}=1,\frac{\partial y}{\partial\tilde{y}}=1,\ \frac{\partial x}{\partial\tilde{t}}=V,\frac{\partial t}{\partial\tilde{t}}=1$ and all others derivative, included in (7) are equal to zero, on the basis (7) we come to equalities

$$\frac{\partial}{\partial\tilde{x}}=\frac{\partial}{\partial x},\quad \frac{\partial}{\partial\tilde{y}}=\frac{\partial}{\partial y},\quad \frac{\partial}{\partial\tilde{t}}=V\frac{\partial}{\partial x}+\frac{\partial}{\partial t}.$$

From here it is visible, that in the equations (5) transitions from coordinates of passing system of reference to laboratory coordinates are reduced to the following replacement of the differential operators: $\tilde{\nabla}^2\to\nabla^2$, $\partial/\partial\tilde{t}\to\partial/\partial t+V\partial/\partial x$. On this basis the equations (5) can give a form

$$\rho\left(\frac{\partial}{\partial t}+V\frac{\partial}{\partial x}\right)^2 u_2=\lambda\nabla^2u_2+e_{15}\nabla^2\varphi_2,\quad \frac{e_{15}}{\varepsilon}\nabla^2u_2=\nabla^2\varphi_2. \qquad (8)$$

We shall be interested in propagation of plane monochromatic SH-waves in moving piezoelectric from a position of the observer to laboratory system of reference, not accepting, while, in attention limitation of the sizes of a crystal. Then according to second of the equations (8) we have $\varphi_2=u_2(e_{15}/\varepsilon_2)$, where $u_2\sim\exp[i(\mathbf{k}_2\mathbf{r}-\omega t)]$ is the solution of the wave equation

$$\left(\frac{\partial}{\partial t}+V\frac{\partial}{\partial x}\right)^2 u_2=c_t^2\nabla^2u_2,\quad c_t=\sqrt{\frac{\lambda^*}{\rho}},\quad \lambda^*=\lambda+\frac{e_{15}^2}{\varepsilon}. \qquad (9)$$

Further, noticing, that $\mathbf{V}\|x$ and using replacements $\frac{\partial}{\partial t}=-i\omega$, $V\frac{\partial}{\partial x}=\mathbf{V}\frac{\partial}{\partial\mathbf{r}}=i\mathbf{V}\mathbf{k}_2$, $\nabla^2=-c_t^2\mathbf{k}_2^2$, we come to next dispersion relation for SH-waves in a moving crystal

$$c_t^2\mathbf{k}_2^2=(\omega-\mathbf{k}_2\mathbf{V})^2. \qquad (10)$$

In expressions (9), (10) c_t is the velocity of shear waves in a piezoelectric material, λ^* is the shear modulus modified by piezoelectric effect.

The formula (10) establishes connection of a SH-wave frequency in laboratory system of reference ω with wave number k_2, and also shows dependence of phase velocity of a wave $v_2=\omega/k_2$ from a direction of propagation in relation to a direction of a crystal motion. Thus, a consequence of a crystal motion concerning the observer is the anisotropy of propagation of SH-waves. If the left side of equality (10) to transfer to the right, the dispersion relation will accept a form of a difference of two squares with a zero right part:

$$(\omega-\mathbf{k}_2\mathbf{V})^2-(\mathbf{k}_2c_t)^2=[(\omega-\mathbf{k}_2\mathbf{V})+c_t\sqrt{\mathbf{k}_2^2}][(\omega-\mathbf{k}_2\mathbf{V})-c_t\sqrt{\mathbf{k}_2^2}]=0\ .$$

Accordingly, it will break up to two independent equations

$$\omega=\mathbf{k}_2\mathbf{V}\pm c_t\sqrt{\mathbf{k}_2^2}\ . \tag{11}$$

The presence of two various dispersion branches has basic meaning for understanding of specificity of acoustic tunneling of waves through a gap of piezoelectrics, undergoing the relative longitudinal displacement. In the beginning we shall notice, that the ray or group velocity of waves can be defined on known (Balakirev & Gilinskii, 1982; Royer & Dieulesaint, 2000) to the formula $\mathbf{V}_g=\partial\omega/\partial\mathbf{k}_2$. From (11) differentiation ω on $\mathbf{k}_2$ and taking into account, that $k_2=\sqrt{\mathbf{k}_2^2}$, we receive

$$\mathbf{V}_g=\mathbf{V}\pm c_t\mathbf{n}_2\,, \tag{12}$$

where the value $\mathbf{n}_2=\mathbf{k}_2/k_2$ is the vector of wave normal of a SH-wave. On the other hand, expression (11) it is possible to write as $\omega=\mathbf{k}_2\mathbf{V}\pm c_tk_2$ and after division on k_2 to receive expression for phase velocity

$$\mathrm{v}_2=\mathbf{V}\mathbf{n}_2\pm c_t\ . \tag{13}$$

Multiplying both sides (13) on $\mathbf{n}_2$, we come to a conclusion, that the phase velocity of a wave $\mathbf{v}_2=\mathrm{v}_2\mathbf{n}_2$ coincides with its group velocity $\mathbf{V}_g$ and for the observer of laboratory system of reference represents expected result of Galilean addition of velocity of wave propagation concerning a crystal with velocity of moving of the crystal.

Pair of signs in expressions (11) - (13) should not cause bewilderment, as the propagation of plane monochromatic waves along any elected direction in a crystal can occur by a counter manner. For an immobile crystal direct and return propagation of waves ($+\mathbf{n}_2$ is the wave normal for a wave direct, and $-\mathbf{n}_2$ - for a wave of return propagation) are made equally with velocity c_t. The crystal motion brings in a difference to their propagation, indicating about acquisition by a crystal of such quality, as nonreciprocity of propagation. An evident picture of nonreciprocity of wave propagation because of a crystal motion demonstrate on Fig. 2, 3 polar curves of the reduced phase velocity

$$\frac{\mathrm{v}_2}{c_t}=\beta\cos\theta\pm1\,,\quad \beta=\frac{V}{c_t}\,, \tag{14}$$

where θ is the angle between vectors $\mathbf{n}_2$ and $\mathbf{V}$. At construction the polar curves we were guided by a rule to correlate to waves of direct propagation orientation wave normal in side from pole and, opposite, for waves of return propagation to consider as it oriented along a direction of wave propagation in the side of a pole. We agree also to represent the polar curves of phase velocity of waves of direct propagation by continuous lines, and polar curves of phase velocity of waves of return propagation - dashed lines. Let's notice, that the equality (13), resulting to the formula (14), represents balance of projections of velocities participating in Galilean addition, on a direction of wave propagation. In this connection the value v_2 for waves of return propagation turned out negative, and at construction of the polar curves its modulus was used.

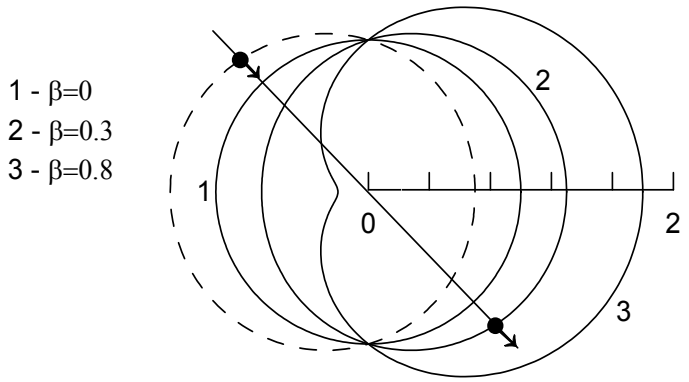


Fig. 2. Polar curves of phase velocity of SH-wave propagation at subsonic velocities of a crystal motion

The accepted way of graphic representation of polar curves excludes mess in definition of types of waves (direct or return propagation) and choice of the appropriate orientation wave normal. For an example, on Fig. 2 thin straight line allocates a direction of propagation of a SH-wave, which in the top point is crossed with dashed polar curve of return propagation ($v_2<0$) for $\beta=0.3$. Thus, this point we correlate a return wave with wave normal, as shown by arrow directed to a pole. The same wave, but only direct propagation, we have the right to connect with the bottom point laying in crossing of a line of propagation with curve 2, which is mirror reflection of dashed polar curve concerning a vertical line passing through a pole. Last circumstance is a geometrical consequence of rearrangement by places of waves of direct and return propagation at inversion of velocity of a crystal motion of what it is uneasy to be convinced by substitution $\beta\rightarrow-\beta$ in (13), (14).

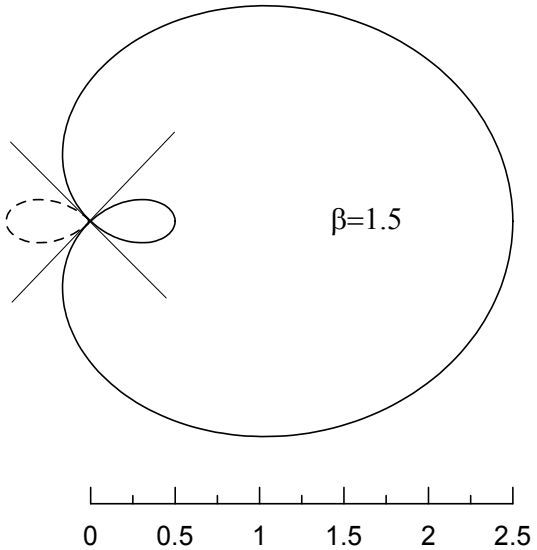


Fig. 3. Polar curve of phase velocity of a SH-wave propagation at supersonic crystal motion

At subsonic velocities of a crystal motion ($\beta<1$, Fig. 2) circular at $\beta=0$ polar curve 1, identical at direct and return propagation, is horizontal stretched (is compressed) for waves of direct (return) propagation in sector of polar angles $|\theta|<\pi/2$. In sector of angles $|\theta|>\pi/2$ takes place opposite. Such deformation of polar curves, reflecting property of nonreciprocity of SH-wave propagation in a moving crystal, is expressed, naturally, the more strongly, than above velocity of a crystal. At supersonic velocities ($\beta>1$, Fig. 3) the change of polar curves by motion of a crystal is complicated by an opportunity of mutual transformation of waves of direct and return propagation. On mathematical language it will be expressed by change of sign of phase velocity v_2 in the formula (14). So, at $\beta>1$ and angles $\theta>\pi/2$ the first term $\beta\cos\theta$ of the right side of expression for the phase velocity of a wave of direct propagation will be negative and, since an angle of transformation $\theta^*_+=\arccos(-1/\beta)$, begins to surpass in magnitude unit. Thus instead of former values $v_2>0$ we shall receive, as the certificate of the come transformation of a wave in a wave of return propagation, $v_2<0$. On the polar curves of phase velocity the range of transformation will settle down of a symmetrically horizontal axis and it will found in sector of obtuse (acute) angles between two thin straight lines, crossed in a pole, on Fig. 3 for waves of direct (return) propagation.

The parts of polar curves of phase velocity appropriate to the transformed waves, look like petals. On Fig. 3 such petals appropriate to transformation of a wave of direct propagation to a wave of return propagation, is shown by a dashed line. Instead of it the directly propagating waves receive mirror imaged concerning a vertical and shown continuous line a petal of waves, which are transformed by a crystal motion from waves of return propagation in waves of direct propagation. Certainly, that in the crystal, i.e. in a passing system of coordinates any transformation of waves does not occur. It appears possible only with transition to a position of the observer of laboratory system of reference, and in this sense is effect typically of a relativistic nature. As at transformation of a wave there is an inversion its wave normal, this phenomenon can be classified as specific, relativistic version of the known phenomenon of conjugation of wave front (Fisher, 1983; Brysev et al, 1998; Fink et al, 2000). But if in a basis of the processes, described in the literature, the parametrical effects put, nonlinear first of all, here conjugation of wave front is provided with the linear laws of Galilean kinematics.

2.2 Refractive properties of a gap

After we have established characteristics of shear wave propagation caused by relative motion of a crystal, it is possible to begin definition of those waves, which arise in crystals on the different sides of a gap under action of a wave, falling on it. As shown in Fig. 1, we shall believe, that the incidence of a shear wave on a gap occurs on the side of immobile crystal. Then, it is necessary to understand frequency ω as a frequency of incident wave. Standard for the wave refraction problems (Balakirev & Gilinskii, 1982; Royer & Dieulesaint, 2000) the need of phase conjugation of harmonic fields on boundaries of a gap $y=\pm h$ follows from boundary conditions (will be discussed more in details in section 2.3) and means identical concurrence of phases of oscillations in all arising waves and near-boundary electrical fields with a phase of oscillations of incident wave. Or else, if the incident wave has a phase multiplier $\exp[i(k_x x-\omega t)]$, the same phase multiplier will characterize oscillations with change of longitudinal coordinate x and time t in all other arising waves. Accordingly, the law of wave refraction is formulated as equality of frequencies of waves to frequency of incident wave

$$\omega_{\mathrm{i}} = \omega_{\mathrm{R}} = \omega_{\mathrm{T}} \equiv \omega\,, \tag{15}$$

and as the predefiniteness of projections of wave vectors to a direction of boundaries of a gap by a projection of a wave vector of incident wave

$$k_{\|}^{(\mathrm{i})} = k_{\|}^{(\mathrm{R})} = k_{\|}^{(\mathrm{T})} \equiv k_x\,. \tag{16}$$

In expressions (15), (16) indexes i, R, T show an belonging of examined parameter according to incident, reflected (arising in immobile crystal) and refracted (arising in a moving crystal) to waves.
In view of equality (6) for waves in immobile crystal we have a dispersion relation

$$c_t^2\mathbf{k}_1^2 = \omega^2\,. \tag{17}$$

From two its possible branches $c_tk_1=\pm\omega$ for incident wave we, actually, elect a branch of directly propagating wave $c_tk_1=\omega$. It specifies a positive sign in (15). Thus, in the subsequent transformations with use of expressions (15), (16) we accept, that $\omega>0$, $k_1>0$ and accordingly $k_x=k_1\sin\alpha>0$, where α is the angle of incidence (see Fig. 1).
As $\mathbf{k}_1=\mathbf{n}_1\omega/c_t$, where $\mathbf{n}_1$ is the vector of a wave normal, the refractve curve, described by a vector $\mathbf{k}_1=\mathbf{k}_1(\mathbf{n}_1)$ in the incident plane $x0y$, has for waves in immobile crystal the form of a circle of radius ω/c_t. In particular, the incident wave has the wave normal $\mathbf{n}_1^{(i)}=(\sin\alpha, -\cos\alpha)$. In view of (16) a wave normal $\mathbf{n}_1^{(\mathrm{R})}=(n_x^{(\mathrm{R})}, n_y^{(\mathrm{R})})$ any other wave arising in immobile crystal, also is characterized by value $n_x^{(\mathrm{R})}=\sin\alpha$ and noticing further, that $n_x^{(\mathrm{R})2}+n_y^{(\mathrm{R})2}=1$, we have $n_y^{(\mathrm{R})}=\pm\cos\alpha$. The negative sign here actually is already used for an incident wave, so on reasons connected with causality, we are compelled to stop the choice on a positive sign. Thus, in immobile crystal in addition to the incident wave there is only one reflected wave with the wave normal $\mathbf{n}_1^{(\mathrm{R})}=(\sin\alpha, \cos\alpha)$, which is propagated in side from the boundary $y=h$. It is obvious, that in complete conformity with Mandelstam' principle of radiation the following from dispersion relation (17) the expression for the group velocity of waves in immobile crystal $\mathbf{V}_{\mathrm{g}}^{(1)}=\mathbf{n}_1c_t$ is confirmed with ability of this wave to take aside energy from boundary.
For waves arising in a moving crystal under action of incident wave, it is easier all to proceed from expression (13) and formula for a wave vector $\mathbf{k}_2=\omega\mathbf{n}_2^{(\mathrm{T})}/\mathrm{v}_2$, where $\mathbf{n}_2=(n_x^{(\mathrm{T})}, n_y^{(\mathrm{T})})$ - vector of wave normal. Meaning, that in examined case $\mathbf{V}\mathbf{n}_2=\mathrm{V}n_x^{(\mathrm{T})}=\mathrm{V}\sin\alpha_t$ (α_t - angle of refraction), we receive

$$\mathbf{k}_2 = \begin{cases} \mathbf{k}_2^{+} = \dfrac{\omega}{c_t}\dfrac{\mathbf{n}_2}{1+\beta\sin\alpha_t}, & \beta\sin\alpha_t > -1, \\ \mathbf{k}_2^{-} = \dfrac{\omega}{c_t}\dfrac{\mathbf{n}_2}{\beta\sin\alpha_t - 1}, & \beta\sin\alpha_t > 1\,. \end{cases} \tag{18}$$

According to (18) there are two refraction branches, appropriate to signs "plus" ($\mathbf{k}_2^{+}$) and "minus" ($\mathbf{k}_2^{-}$) in the formulas (12) - (14). The conditions of existence of the branches express the mentioned above requirement of positive values of wave numbers $k_2^{\pm}>0$ at the elected way of representation of problem solution in laboratory system of reference by means of waves of direct propagation $\omega>0$. In this sense the formula (18) does not add the new information that was received in the previous section, and only translates its in the terms of wave vectors more convenient for consideration of refractive effects.

The first refraction branch with wave number k_2^+ we arrange to name as a usual branch, as for it the waves in a moving crystal represent waves of a direct propagation irrespective of a choice of system of reference. Really, if, using (1) to compare phases of oscillations of a wave in passing $\exp[ik_x\tilde{x}-\Omega\tilde{t}]$ and laboratory system of reference $\exp[ik_x x-\omega t]$, for frequency of a wave in passing system of reference it is not difficult to receive expression

$$\Omega = \omega - \mathbf{k}_2\mathbf{V} \,. \tag{19}$$

It shows Doppler shift of frequency of a wave and at substitution $\mathbf{k}_2^+$ from (18) determines always positive values of frequencies $\Omega_+=\Omega(\mathbf{k}_2^+)=\omega(1+\beta\sin\alpha_t)^{-1}$. On the contrary, at the substitution in (19) $\mathbf{k}_2^-$, we receive $\Omega_-=\Omega(\mathbf{k}_2^-)=-\omega(\beta\sin\alpha_t-1)^{-1}$ and for the second refraction branch we have $\Omega_-<0$, whereas $\omega>0$. Thus, in case of this refraction branch, the waves, refracted in a moving crystal, are in relation to the crystal waves with the reversed wave front, but are perceived in laboratory system of reference as waves of direct distribution. Therefore it is possible to name a refraction branch $\mathbf{k}_2^-$ as a reverse refraction branch.

As against known results (Fisher, 1983; Brysev et al, 1998; Fink et al, 2000) the phenomenon of conjugation of wave front, examined by us, has of a purely kinematic origin. It is caused by drift action of a medium moving at a transonic velocity along the wave incident from the immobile crystal, which exhaustively compensates the reverse propagation of a refracted wave relative to the crystal and eventually provides its spatial synchronism (by means of electrical fields induced via the gap) with waves that are true of direct propagation in the immobile piezoelectric crystal.

On Fig. 4, 5 solid lines show typical refraction curves of direct propagating waves which are described by the ends of wave vectors $\mathbf{k}_2$ from (18) at change of a direction of a vector wave normal $\mathbf{n}_2$ in a plane of incidence. They correspond to two qualitatively different cases of SH-wave refraction by a gap at subsonic ($\beta<1$, Fig. 4) and very supersonic ($\beta>2$, Fig. 5) velocities of relative crystal motion. Simultaneously with it the dashed circles represent on Fig. 4, 5 dependences $\mathbf{k}_1(\mathbf{n}_1)$ for SH-waves in immobile crystal. At $\beta<1$ takes place only usual refraction (refraction curve is marked "plus"). The incident wave with a wave vector

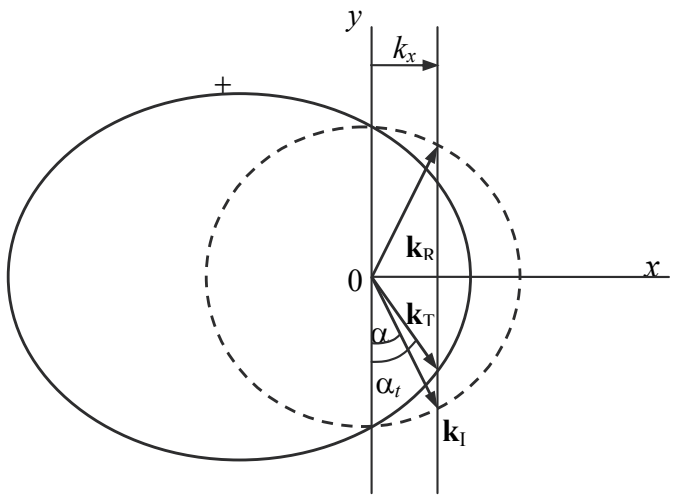


Fig. 4. Polar curves of refraction for the case $\beta<1$.

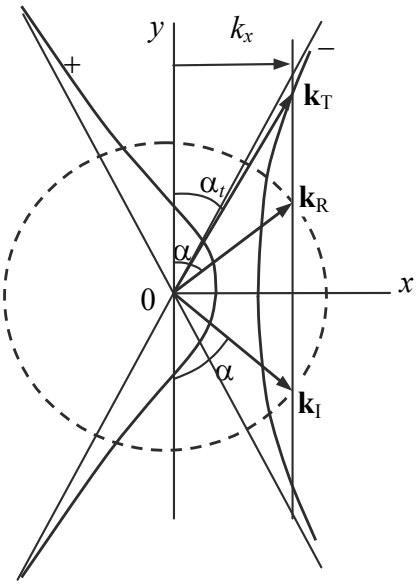


Fig. 5. Polar curves of refraction for the case β>2.

$\mathbf{k}_I=(k_1\sin\alpha, -k_1\cos\alpha)$ defines valid (16) identical in all other waves a horizontal projection k_x. The wave vectors reflected $\mathbf{k}_R$ and refracted in a moving crystal $\mathbf{k}_T$ of waves will be, therefore, are directed from the origin 0 to points of crossing of appropriate refraction curves by a thin vertical line cutting on a horizontal a segment, equal k_x, so that the energy was removed by waves on a direction of their propagation from boundaries of crystals. Thus, we have $\mathbf{k}_R=(k_1\sin\alpha, k_1\cos\alpha)$, $\mathbf{k}_T=(k_2\sin\alpha_t, -k_2\cos\alpha_t)$.

In case of β>1 branch usual refraction exists in intervals $0<\theta<\theta_1^*$ and $\theta_2^*<\theta<2\pi$ of polar angle $\theta=\pi/2-\alpha_t$, where $\theta_2^*=2\pi-\theta_1^*$, $\theta_1^*=\arccos(-1/\beta)$. In addition to it, as shown in Fig. 5, in the sector of angles $|\theta|<\arccos(1/\beta)$ there is a branch inversed refraction, marked by sign "minus". However, if β<2, its curve lays more to the right of a dashed circle for a refraction curve of immobile crystal. For this reason appropriate inversed refraction of a wave are not capable to be raised in a moving crystal by incident wave and refraction picture does not differ that is submitted on Fig. 4. At velocities of relative motion of crystals is twice higher sound usual refraction will be replaced, as shown in Fig. 5, inversed refraction. It will take place, since the angle of incidence α_0, at which

$$\sin\alpha_0 = \frac{1}{\beta-1}. \tag{20}$$

In order to conclude this condition in expression (18) for wave number of the inversed wave k_2^- it is necessary to accept $\alpha_t=\pi/2$ and to take into account following from (16) equality $k_2^-=k_1\sin\alpha$. In passing we shall notice, that in a regime of sliding propagation $\alpha_t=\pi/2$ difference of longitudinal projections k_x of wave vectors for inversed and usual refracted waves is given by the formula

$$\Delta k_x = k_x^- - k_x^+ = \frac{2k_1}{(\beta-1)^2}. \tag{21}$$

From (21) we have $\Delta k_x>0$ at any finite values of quantity β. On geometry this fact means absence of crossing of usual and inversed refraction curves. Physically it shows existence of the refracted wave always in a form of single wave, fist (at $\alpha<\alpha_0$) as usual, and then (at $\alpha>\alpha_0$, if $\alpha_0\in[0, \pi/2]$),), - as the inversed wave. As the transition from usual to inversed refraction is reached by change of a sign $\cos\alpha_t$ (at an invariance of all other parameters of a wave), at construction of the solution there is a temptation to describe it in the terms of usual refraction, not resorting to consideration of two separate solutions. By an implicit manner such opportunity contains in refractive relations. Really, at usual refraction from (16), (18) the expression turns out

$$\cos\alpha_t = \frac{\sqrt{(1-\beta\sin\alpha)^2-\sin^2\alpha}}{1-\beta\sin\alpha}. \tag{22}$$

According to the requirement $k_2^-{>}0$, that is equivalent also to following from (16), (18) condition $\beta\sin\alpha>1$, the actual inclusion by the formula (22) case not only usual, but also inversed refraction ($\cos\alpha_t\to-\cos\alpha_t$) is obvious. Thus, not ordering beforehand to $\cos\alpha_t$ of a negative sign, i.e. describing refraction of a SH-wave in a moving crystal as usual, with use of the formula (22) it is possible automatically to take into account transition to inversed refraction.

2.3 Solution of a boundary problem

The connection between crystals is carried out by electrical fields penetrating through a gap. Therefore it is necessary to consider the equations (6), (8) together with the Laplace equation for potential φ of an electrical field in a gap

$$\nabla^2\varphi = 0. \tag{23}$$

It is got, if, considering a gap as very rarefied material medium with permeability ε_g, instead of the equations (4) to use in laboratory system of coordinates the equation $\nabla\mathbf{D}=0$, where $\mathbf{D}=\varepsilon_g\mathbf{E}$ is the induction, and $\mathbf{E}=-\nabla\varphi$ is the strength of a field. According to the equation (6) and accepted on a Fig. 1 picture of incidence, for immobile crystal we have

$$\begin{aligned} &u_1 = U\exp[i(k_x x-\omega t)][\exp(-ik_y^{(1)}y)+R\exp(ik_y^{(1)}y)], \quad \varphi_1 = \frac{e_{15}}{\varepsilon}u_1+\Phi_1, \\ &\Phi_1 = F_1\exp[i(k_x x-\omega t)\exp(-k_x y), \quad k_x = \frac{\omega}{c_t}\sin\alpha, \quad k_y^{(1)} = \frac{\omega}{c_t}\cos\alpha. \end{aligned} \tag{24}$$

In the moving crystal on base of equations (8) and stated above idea to consider the tunneling wave as a single wave of usual refraction, we have

$$\begin{aligned} &\varphi_2 = \frac{e_{15}}{\varepsilon}u_2+\Phi_2, \quad \Phi_2 = F_2\exp[\,i(k_x x-\omega t)]\exp(-k_x y), \\ &u_2 = UT\exp[\,i(k_x x-\omega t)]\exp(-ik_y^{(2)}y), \\ &k_y^{(2)} = \sqrt{k_2^2-k_x^2} = k_1\sqrt{(1-\beta\sin\alpha)^2-\sin^2\alpha}. \end{aligned} \tag{25}$$

To the expressions (24), (25) we shall add expression for an electric field potential in a gap

$$\varphi = \exp[i(k_x x - \omega t)][C\exp(k_x y) + D\exp(-k_x y)]\,. \tag{26}$$

This expression follows from the equation (23).
In the formulas (24) - (26) values Φ_j represent potentials of fields of near-boundary electrical oscillations, U is the known amplitude of incident wave. The coefficients of reflection (R) and passage of incident wave through the gap (T), and also amplitude of potentials of near-boundary electrical oscillations F_1, F_2, C, D are subject still to determination. With this purpose we use boundary conditions of a problem, which mean a continuity of electrical potentials, y-components of an electrical induction and absence of shear stresses T_{zy} at $y=\pm h$. As the values $D_y^{(2)}$, $T_{yz}^{(2)}$ included in boundary conditions, do not contain derivative on time, they will not change at transitions from passing system of reference to laboratory system of reference. In result the boundary conditions will accept in laboratory system of reference the form

$$\begin{gathered}
\left(\frac{e_{15}}{\varepsilon}u_j + \Phi_j\right)\Bigg|_{y=(-1)^{j+1}h} = \varphi\Big|_{y=(-1)^{j+1}h}\,, \\
\left(e_{14}\frac{\partial u_j}{\partial x} + \varepsilon\frac{\partial \Phi_j}{\partial y}\right)\Bigg|_{y=(-1)^{j+1}h} = \varepsilon_g\frac{\partial \varphi}{\partial y}\Big|_{y=(-1)^{j+1}h}\,, \\
\left[\lambda^*\frac{\partial u_j}{\partial y} + \frac{e_{14}e_{15}}{\varepsilon}\frac{\partial u_j}{\partial x} + \left(e_{15}\frac{\partial \Phi_j}{\partial y} + e_{14}\frac{\partial \Phi_j}{\partial x}\right)\right]\Bigg|_{y=(-1)^{j+1}h} = 0\,.
\end{gathered} \tag{27}$$

After substitution (24) - (26) in (27) and solution of forming system of the nonhomogeneous algebraic equations we shall receive representing for us interest coefficients

$$R = \frac{\dfrac{k_y^{(1)}k_y^{(2)}}{k_x^2} + \left[\Delta_a\Delta_s + i\dfrac{(k_y^{(1)} - k_y^{(2)})}{2k_x}(\Delta_a + \Delta_s)\right]}{\dfrac{k_y^{(1)}k_y^{(2)}}{k_x^2} - \left[\Delta_a\Delta_s + i\dfrac{(k_y^{(1)} - k_y^{(2)})}{2k_x}(\Delta_a + \Delta_s)\right]}, \tag{28}$$

$$T = -\frac{\dfrac{i\bar{\varepsilon}\mathcal{K}^2}{(1+\bar{\varepsilon}^2)\sinh(\xi) + 2\bar{\varepsilon}\cosh(\xi)}\left(\dfrac{k_y^{(2)} + k_y^{(1)}}{k_x}\right)}{\dfrac{k_y^{(1)}k_y^{(2)}}{k_x^2} - \left[\Delta_a\Delta_s + i\dfrac{(k_y^{(1)} - k_y^{(2)})}{2k_x}(\Delta_a + \Delta_s)\right]}, \tag{29}$$

where we have

$$\begin{gathered}
\Delta_a = \frac{\mathcal{K}^2 - \bar{\varepsilon}\tanh(\xi/2)\mathcal{K}_\perp^2}{1 + \bar{\varepsilon}\tanh(\xi/2)}\,, \quad \Delta_s = \frac{\mathcal{K}^2\tanh(\xi/2) - \bar{\varepsilon}\mathcal{K}_\perp^2}{\bar{\varepsilon} + \tanh(\xi/2)}\,, \\
\mathcal{K}^2 = \frac{e_{15}^2}{\varepsilon\lambda + e_{15}^2}\,, \quad \mathcal{K}_\perp^2 = \frac{e_{14}^2}{\varepsilon\lambda + e_{15}^2}\,.
\end{gathered} \tag{30}$$

In these formulas, $\mathcal{K}^2$ and $\mathcal{K}_\perp^2$ are the square coefficients of electromechanical coupling for the longitudinal and transverse piezoeffect respectively, $\xi=k_x h$ is wave half-width of the gap, and $\bar{\varepsilon}=\varepsilon/\varepsilon_g$. In a particular case $\beta=0$ when the relative longitudinal motion of piezoelectric crystals is absent, we have $k_y^{(1)}=k_y^{(2)}\equiv k_y$, $k_y/k_x=\tan\theta$ ($\theta=\pi/2-\alpha$ is the glancing angle) the expression (28)-(30) leads in earlier known results (Balakirev & Gilinskii ,1982).

2.4 Discussion of results

The main attention we shall concentrate here on angular spectra of coefficients of reflection and passage of waves through a gap. For the beginning we shall notice, that in limiting cases $h\to\infty$ and $\varepsilon_g\to\infty$ ($\bar{\varepsilon}\to 0$) the expressions (28) - (30) show absence of passage $T\to 0$. In the first case it is caused by the disappearance of coupling of crystals by electrical fields through a gap in process of increase of its thickness. In the second case takes place a shielding of fields of a gap due to metallization of crystal surfaces.

Typical behaviour of angular dependences of modules of reflection coefficient $|R|$ and the passage coefficient $|T|$, calculated on the formulas (28) - (30) for pair of crystals $LiIO_3$ with parameters $\mathcal{K}^2=0.38$, $\mathcal{K}_\perp^2=0.002$, $\bar{\varepsilon}=8.2$, demonstrate Fig. 6 and 7. As can be seen, a general tendency in the case of usual refraction is a decrease in the extent of wave tunneling into the moving crystal with increasing angle of incidence. This trend is more pronounced in the angular dependences of the reflection coefficient R. Indeed, even at relatively small velocities, the opposite (antiparallel) relative longitudinal displacement (RLD) ($\beta=-0.05$, see curve 1 in Fig. 6) lead to extension of the wedge of transparency (depicted by the dashed curve in the region of large α) by more than a half toward greater angles ($|R|_{\min}>0.6$). However, a nearly complete extension of this wedge (Fig. 7, curve 3) takes place only for

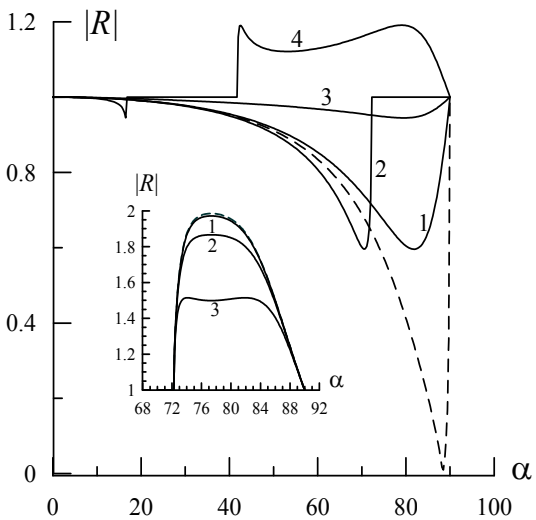


Fig. 6. Plots of reflection coefficient $|R|$ versus angle of incidence α for a pair of piezoelectric $LiIO_3$ with an extremely thin ($\xi=10^{-6}$) gap for an RLD velocity of $\beta=-0.05$ (1), 0.05 (2), –2.5 (3), and 2.5 (4). The inset shows the angular dependence of the reflection coefficient in the case of reverse refraction for $\beta=2.05$ and various gap thicknesses $\xi=10^{-3}$ (1), 10^{-2} (2), 0.06 (3), and 10^{-6} (dashed curve).

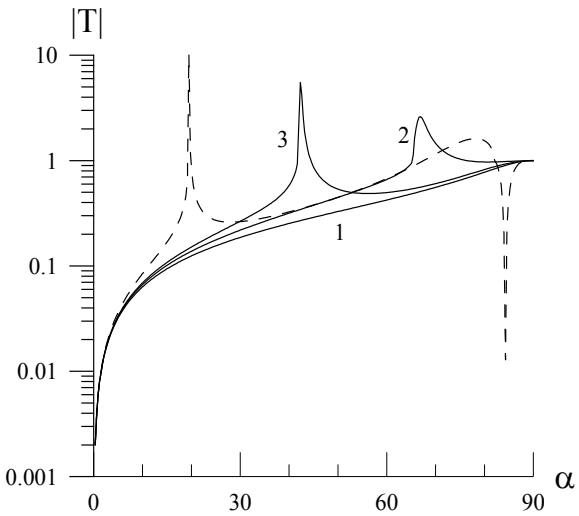


Fig. 7. Plots of the transmission coefficient $|T|$ versus angle of incidence α for a pair of piezoelectric $LiIO_3$ crystals with a thin ($\xi=10^{-3}$) gap for an RLD velocity of β=–0.5 (1), 0.1 (2), 0.5 (3) and 2.01 (dashed curve).

ultrahigh velocities of the opposite RLD (β<0, $|\beta|>2$). However, a comparison of curves 1 – 3 in Fig. 7 shows that no significant decrease in the transmission of waves through the gap takes place and the possibility of practical application of the effect of wave tunneling is retained.

In the case of parallel RLDs (β>0) the transparency wedge under the usual refraction conditions is not only extended with increasing β, but is additionally shifted toward smaller incidence angles by the appearing region of total reflection (Fig. 6, curves 2). The angular dependences of transmission (Fig. 7, curves 2 and 3) show well-pronounced peaks at the limiting angles α^* of total reflection ($\sin\alpha^*=(1+\beta)^{-1}$). The left sides of these peaks apparently correspond to the conditions of effective tunneling of incident wave into the moving crystal. However, it should be taken into account that, in view of the proximity to α^*, the tunneling waves will have very small transverse components ($k_y^{(2)}\geq 0$) of the wave vector. Thus, the effective tunneling of waves into the moving crystal is possible, but only for small (or very small) angles of refraction for moderate (Fig. 7, curve 3) and even small (Fig. 7, dashed curve) angles of incidence.

In the latter case, ultra-high RLD velocities (β>2) are necessary, which make possible the reverse refraction. As for the phenomenon of tunneling as such, the region of reverse refraction $\alpha>\alpha^{**}$ ($\sin\alpha^{**}=(\beta-1)^{-1}$, $\alpha^{**}\sim 82°$ for the dashed curve in Fig. 7) does not present much interest because formula (13) implies "closing" of the gap for $k_y^{(1)}+k_y^{(2)}=0$ with significant decrease in the transmission coefficient $|T|$ in the vicinity of the corresponding incidence angle. On the other hand, there is an attractive possibility of enhancement of the reflected wave for $|R|>1$ (see Fig. 6, curve 4 and the inset to Fig. 6, curves 1 – 3), which is related to the fact that the wave in a moving crystal in the case of reverse refraction propagated (as indicated by dashed arrow in Fig. 1) toward the gap and carries the energy in the same direction. Naturally, an increase in the gap width leads to decrease in electric coupling between crystals and in the enhancement of reflection (see the inset to Fig. 6, curves 1 – 3).

3. Tunneling of shear waves by a vacuum gap of piezoelectric 6- and 222-class crystal pair at the uniform relative motion

In this section we consider the effect of tunneling of shear waves in the layered structure of piezoelectric crystals with a gap for the crystal pair of 6 (6mm, 4, 4mm, ∞m) and 222 (422, 622, $\bar{4}$2m, $\bar{4}$3m, 23) class symmetry, undergoing relative longitudinal motion. This case allows, to estimate influence of elastic and electric anisotropy on tunneling of SH-waves in a moving crystal in conditions of difference of its symmetry from symmetry of an immobile crystal. We assume that the shear wave falls on the part of the immobile crystal of a class 6. Now, instead (8) we shall have in laboratory system of reference the equations

$$\rho_2\left(\frac{\partial}{\partial t}+V\frac{\partial}{\partial x}\right)^2 u_2=\lambda_{55}^{(2)}\frac{\partial^2 u_2}{\partial x^2}+\lambda_{44}^{(2)}\frac{\partial^2 u_2}{\partial y^2}+(e_{14}^{(2)}+e_{25}^{(2)})\frac{\partial^2\varphi_2}{\partial x\partial y},$$
$$(e_{14}^{(2)}+e_{25}^{(2)})\frac{\partial^2 u_2}{\partial x\partial y}=\varepsilon_1^{(2)}\frac{\partial^2\varphi_2}{\partial x^2}+\varepsilon_2^{(2)}\frac{\partial^2\varphi_2}{\partial y^2}. \qquad (31)$$

The equations (6) remain in force, but with a clause, that in them all parameters of a crystal are marked by an index "1", i.e. $\rho\to\rho_1$, $\lambda\to\lambda_{55}^{(1)}$, $e_{15}\to e_{15}^{(1)}$ and $\varepsilon\to\varepsilon_1^{(1)}$.
Following from (6), (31) the dispersion relation of SH-waves and Snell's condition (16) allow to establish the refraction low in form of the inverse dependence

$$\sin\alpha=\frac{v_1\sin\alpha_t}{V\sin\alpha_t\pm v_{2\|}\sqrt{(\sin^2\alpha_t+a\cos^2\alpha_t)(1+Q^2(\alpha_t)\sin^2\alpha_t)}}. \qquad (32)$$

Here $v_1=(\lambda_{55}^{(1)*}/\rho_1)^{1/2}$ is the velocity of SH-waves in immobile crystal, $\lambda_{55}^{(1)*}=\lambda_{55}^{(1)}+e_{15}^{(1)2}/\varepsilon_1^{(1)}$, $v_{2||}=(\lambda_{55}^{(2)}/\rho_2)^{1/2}$ is the velocity of SH-wave propagation in a moving crystal along [100]-direction (axis $\tilde{x}$), $a=\lambda_{44}^{(2)}/\lambda_{55}^{(2)}$ is the elastic anisotropy factor of moving crystal. Function $Q^2(\alpha_t)$, determinated by equality

$$Q^2(\alpha_t)=\frac{(e_{14}^{(2)}+e_{25}^{(2)})^2\cos^2\alpha_t}{(\lambda_{55}^{(2)}\sin^2\alpha_t+\lambda_{44}^{(2)}\cos^2\alpha_t)(\varepsilon_1^{(2)}\sin^2\alpha_t+\varepsilon_2^{(2)}\cos^2\alpha_t)}, \qquad (33)$$

is the square of electromechanical coupling factor for SH-waves propagating in (001)-plane of a crystal.
The expression (32) shows that at subsonic velocities of crystal motion there exists only usual refraction, corresponding to the top sign. It is not accompanied by the inversion of wave fronts and has the top threshold of incident angle α^*, such that $\sin\alpha^*=v_1/(V+v_{2||})$. At the supersonic velocities of crystal motion $V>v_{2||}$ total reflection for the usual refraction ($\alpha^*<\alpha<\alpha^{**}$) becomes possible even at smaller rigidity of a moving crystal. Second refraction branch appropriate to the bottom sign in formula (32) and accompanied by the inversion of wave fronts, is possible only at supersonic velocities of crystal motion and additional condition $V>v_1+v_{2||}$. The bottom threshold of this branch α^{**} exceeds the value α^* is determined by equality $\sin\alpha^{**}=v_1/(V-v_{2||})$. On Fig. 8, 9 the curves usual and inversed refraction, received by calculation under the formulas (32), (33) for pair of crystals $Pb_5Ge_3O_{11}$ – Rochell salt with parameters taken from (Royer & Dieulesaint, 2000; Shaskolskaya, 1982) are submitted accordingly.

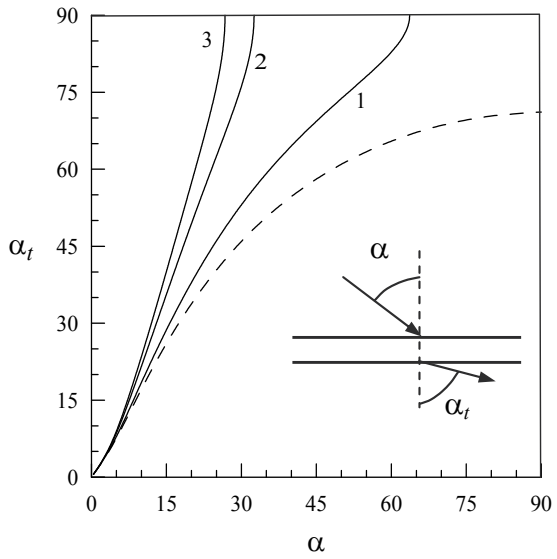


Fig. 8. Curves usual refraction of a wave by a gap $Pb_5Ge_3O_{11}$ – Rochell salt: 1 – $\beta=V/v_{2||}=0.5$, 2 – β=1.5, 3 – β=1.8 (β=0 – dashed curve).

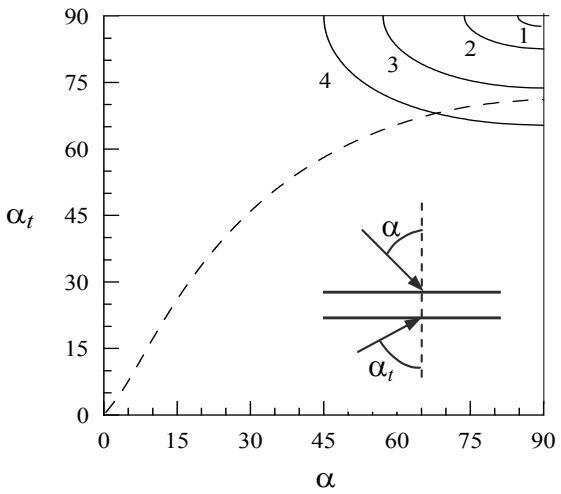


Fig. 9. Curves reversed refraction of a wave by a gap $Pb_5Ge_3O_{11}$ – Rochell salt: 1 – $\beta=V/v_{2||}=2.35$, 2 – β=2.4, 3 – β=2.5, 4 – β=2.6 (β=0 – dashed curve).

The solutions of the equations (6), (20) will keep the form (24), (26), and instead of (25) from the equations (31) we shall receive

$$u_2 = U\exp(i\phi)[T\exp(-ik_y^{(2)}y) + A\exp(sy)],$$
$$\varphi_2 = \frac{ik_x s(e_{14}^{(2)} + e_{25}^{(2)})}{\varepsilon_2^{(2)}s^2 - \varepsilon_1^{(2)}k_x^2}UA\exp(i\phi)\exp(sy) - \frac{ik_x k_y^{(2)}(e_{14}^{(2)} + e_{25}^{(2)})}{\varepsilon_2^{(2)}k_y^{(2)2} + \varepsilon_1^{(2)}k_x^2}UT\exp(i\phi)\exp(-ik_y^{(2)}y). \quad (34)$$

The values $k_y^{(2)}$ and s in expressions (34) are accordingly imaginary $q=-ik_y^{(2)}$ (for solution (34) in writing we chose the case of usual refraction) and real $q=s$ a root of the characteristic equation $[(\omega-k_xV)^2v_{2||}^{-1}a^{-1}+q^2-ak_x^2](bk_x^2-q^2)+Q_0^2k_x^2q^2=0$, where $b=\varepsilon_1^{(2)}/\varepsilon_2^{(2)}$ is the factor of electric anisotropy of a crystal, and $Q_0=Q(0)$. As against the solution (25) for pair of identical hexagonal crystals the near-boundary oscillations any more are not only electrical. They are the connected electro-elastic oscillations, which are made with amplitude A and phase $\phi=k_xx-\omega t$.

The physical sense of boundary conditions will not change. For the top boundary $y=h$ on former it is possible to use conditions (27). On the bottom boundary $y=-h$ their change will be caused by the appropriate differences of the state equations for 222-class crystals from the equations (2), (3) (Royer & Dieulesaint, 2000). After substitution of expressions (24), (26), (34) in boundary conditions and solutions of system of the algebraic equations we shall receive expressions for amplitude coefficients. For example, in the case of a very thin gap ($k_xh\to0$) we have

$$R = \frac{1-\tan\alpha Q_1^2\Psi}{1+\tan\alpha Q_1^2\Psi}, \quad T = -\frac{2(e_{15}^{(1)} + ie_{14}^{(1)})\tan\alpha}{\lambda_{44}^{(2)}(1+i\Delta)(1+\tan\alpha Q_1^2\Psi)}\frac{e_{25}^{(2)}bk_x^2 - e_{14}^{(2)}[k_y^{(2)2} + 2k_x^2(1-b)]}{k_y^{(2)2} + k_x^2[b + Q_0^2f_2(1+f_2)^{-1}]}.$$

The value Ψ characterizes mutual piezoelectric connection of crystals through a gap and is defined by equalities

$$\Psi = i\frac{f_1^2\varepsilon_1^{(1)}\Gamma - \varepsilon_2^{(2)}}{\varepsilon_2^{(2)} + \varepsilon_1^{(1)}\Gamma}, \Gamma = \frac{k_xs(1+i\Delta)(1+f_2)}{bk_x^2(1+isk_y^{-1}\Delta) + f_2s^2(1-ik_y^{(2)}s^{-1}\Delta)}, \Delta = \frac{(s^2 - bk_x^2)(1+f_2) - k_x^2f_2Q_0^2}{(k_y^{(2)2} + bk_x^2)(1+f_2) + k_x^2f_2Q_0^2}.$$

There are $f_1=e_{14}^{(1)}/e_{15}^{(1)}$, $f_2=e_{14}^{(2)}/e_{25}^{(2)}$, and $Q_1^2=e_{15}^{(1)2}/[\varepsilon_1^{(1)}\lambda_{55}^{(1)*}]$.

The numerical accounts show, that elastic and electrical anisotropy of a moving crystal does not cause essential changes in angular spectra of reflection and passage of SH-waves through a gap. The distinctions of symmetry of the crystals in addition to their relative motion are reduced by efficiency of acoustic tunneling. Thus, the assumption, that in a slot structure of crystals, from which one with strong longitudinal, and another with strong transverse piezoelectric effect, is possible appreciable shift of effective acoustic tunneling in area of moderate incident angles, has not found confirmation. The amplitude A of near-boundary electro-elastic oscillations is usually small and does not vary almost under influence of crystal motion. In a considered case of crystals of various classes of symmetry amplification the reflected wave in conditions inversed refraction (superreflection) also takes place. However, similarly to acoustic tunneling the superreflection appears well appreciable only at sliding angles of incidence.

4. Conclusion

In this article we have touched upon the poorly investigated problem of refraction of acoustic waves by a gap of piezoelectric crystals with relative longitudinal motion. By the basic result was the conclusion about existence not only usual, but also so-called inversed refraction, capable to replace the usual refraction at superfast motion of a crystal with velocity twice above velocity of a sound. We have shown, that if usual refraction underlies representations about the tunneling of acoustic waves through a gap, with the inversed refraction the opportunity of amplification of reflection is connected.

Both these phenomena, however, provide essential changes of a level of the reflected signals because of a crystal motion (it is interesting to applications), only at the sliding angles of incidence. It is represented, therefore, most urgent search of conditions and means, which would allow to advance in area of moderate or small angles of incidence. With this purpose, as we have found out, is unpromising to use anisotropy of elastic and electrical properties of a moving crystal or distinction in classes of symmetry of crystals.

We believe that there are two approaches to the decision of a problem. It is, first, search and use of hexagonal piezoelectric crystals with equally strong both longitudinal, and trasverse piezoactivity. Secondly, it is the application already of known piezoelectric materials, but having not a plane, and periodically profiled boundaries of a gap. It is doubtless, that the appropriate theoretical researches of effects acoustic refraction by a gap of piezoelectric crystals with relative motion are required. In particular, it is desirable to consider a case of refraction of piezoactive acoustic waves of vertical polarization. We hope, that present article will serve as stimulus for the further study of acoustic refraction in layered structures of piezoelectric crystals with relative motion.

5. References

Balakirev, M. K. & Gorchakov, A. V. (1977 a). The leakage of elastic wave through a gap between piezoelectrics. *Fiz. Tverd. Tela*, Vol. 19, No 2, pp. 571-572, ISSN 0367-3294.

Balakirev, M. K. & Gorchakov, A. V. (1977 b). Connected surface wave in piezoelectrics. *Fiz. Tverd. Tela*, Vol. 19, No 2, pp. 613-614, ISSN 0367-3294.

Balakirev, M. K.; Bogdanov, S. V & Gorchakov, A. V. (1978). Tunneling of ultrasonic wave through a gap between jodat lithium crystals. *Fiz. Tverd. Tela*, Vol. 20, No 2, pp. 588-590, ISSN 0367-3294.

Balakirev, M. K. & Gilinskii, I. A. (1982). *Waves in Piezoelectric Crystals,* Nauka, Novosibirsk.

Balakirev, M. K. & Gorchakov, A. V. (1986). Volume wave refraction and surface waves in the piezoelectric – gap – piezoelectric system. *Surface*, No 5, pp. 80-85, ISSN 0207-3528 .

Brysev, A.P.; Krutyansky, L.M. & Preobrazhensky, V.L. (1998) Wave phase conjugation of ultrasonic beams. *Phys. Uspekhi*, Vol. 41, No 8, pp. 793-806, ISSN 1063-7869.

Filippov, V. V. (1985). Leakage of an elastic wave through a slot between two media, caused by electrostriction. *Techn. Phys.*, Vol. 55, No 5, pp. 975-979, ISNN 0044-4642.

Fink, M; Casserau, D.; Derode, A; Prada, C.; Roux, P; Tanter, M; Thomas, J.-L. & Wu, F (2000) Time-reversed acoustics. *Rep. Progr. Phys.*, Vol. 63, No 12, pp. 1893-1995, ISSN 0034-4885.

Fisher, R.A. (1983) *Optical Phase Conjugation*, Academic Press, New York.

Grigor'evskii, V.I.; Gulyaev, Yu.V.; Kotelyanskii, I.M.; Pashkevich, G.A. & Plessky, V.P. (1987). Acousti wave transmission through a slot in a piezoelectric. *Akust. Zhurn.*, Vol. 33, No 2, pp. 360-362, ISSN 0320-7919.

Gulyaev, Yu.V. (1967). To the question on electron-phonon interaction proportional to applied external field. *Fiz. Tverd. Tela*, Vol. 9, No 6, pp. 1816-1818, ISSN 0367-3294.

Gulyaev, Yu. V. & Plessky, V.P. (1977 a). Shear surface acoustic wave in dielectrics in the presence of an electric field. *Phys. Lett.*, Vol. 56A, No 6, pp. 491-492, ISSN 0375-9601.

Gulyaev, Yu. V. & Plessky, V.P. (1977 b). Surface waves propagating along a slot in piezoelectric materials. *Akust. Zhurn.,* Vol. 23, No 5, pp. 716-723, ISSN 0320-7919.

Gulyaev, Yu. V. & Plessky, V.P. (1978). Resonant penetration of bulk acoustic wave through the vacuum gap between the piezoelectrics. *Fiz. Tverd. Tela*, Vol. 20, No 1, pp. 133-136, ISSN 0367-3294.

Gulyaev, Yu. V.; Nikitov, S.A. & Plessky, V.P. (1978). Penetration of the acoustic wave through the gap between the piezoelectrics, covered by dielectric layers. *Fiz. Tverd. Tela*, Vol. 20, No 5, pp. 1580-1581, ISSN 0367-3294.

Gulyaev, Yu. V. (2005). Acoustoelectronics (historical review). *Phys. Uspekhi,* Vol. 48, No 8, pp. 847-855, ISNN 1063-7869.

Gulyaev, Yu. V.; Maryshev, S. N. & Shevyakhov, N. S. (2007 a). Shear wave passage through a vacuum gap between hexagonal piezoelectric crystals with relative longitudinal displacement. *Techn. Phys. Lett.,* Vol. 33, No 9, pp. 799-803, ISSN 1063-7850.

Gulyaev, Yu. V.; Maryshev, S. N. & Shevyakhov, N. S. (2007 b). Tunneling of shear waves by a vacuum gap o piezoelectric 6- and 222-class crystal pair under the uniform relative motion. Proceedings of the XIX Session of the Russian Acoustic Society, pp. 27-30, ISBN 8-85118-383-8, Nizhny Novgorod, September 2007, GEOS, Moscow.

Kaliski, S. (1966). The passage of an ultrasonic wave across a contactless junction between two piezoelectric bodies. *Proc. Vibr. Probl.* , Vol. 7, No 2, pp. 95-104, ISSN 0001-6705.

Landau, L. D. & Lifshitz, E.M. (1991). *Quantum Mechanics: Non-Relativistic Theory*, Pergamon Press, ISBN 0-08-020-940-8, New York.

Maryshev, S. N. & Shevyakhov, N. S. (2007). Acoustic wave tunneling in a gap between hexagonal piezoelectric crystals with relative longitudinal displacement. *Techn. Phys. Lett.,* Vol. 33, No 5, pp. 368-372, ISSN 1063-7850.

Pennec, Y.; Djafari-Rouhani, B.; Larabi, H.; Vasseur, J. & Hladky-Hennion, A.-Chr. (2009). Phononic crystals and manipulation of sound. *Phys. Status Solidi C*, Vol. 6, No. 9, pp. 2080-2085, ISSN 1610-1642.

Qiu, Ch. L.; Zhengyou, M. & Jun, S.J. (2005). Mode-selecting acoustic filter by using resonant tunneling of two-dimensional double phononic crystals. *Appl. Phys. Letters.*, Vol. 87, No 10, pp. 104101-104101-3, ISSN 0003-6951.

Royer, D. & Dieulesaint Eu. (2000). *Elastic Waves in Solids I*. Springer, ISBN 3-540-65932-3, Berlin-Heidelberg-New York.

Shaskolskaya, M.P. (1982). *Acoustic Crystals: A Handbook.* Nauka, Moscow.

Van Der Biest, F.; Sukhovich, A.; Tourin, A.; Page, J. H.; Van Tiggelen, B. A.; Liu, Z. & Fink, M. (2005). Resonant tunneling of acoustic waves through a double barrier

consisting of two phononic crystals. *Europhys.Letters.*, Vol. 71, No 1, pp. 63-69, ISSN 0295-50075.

Vilkov, E.A.; Moiseev, A.V. & Shavrov, V.G. (2009). Magnetoelastic wave tunneling via a gap between ferromagnetic crystals with relative longitudinal displacement. *Techn. Phys. Lett.,* Vol. 35. No 9. pp. 876-879, ISSN 1063-7850.

9

Surface Acoustic Wave Based Wireless MEMS Actuators for Biomedical Applications

Don W. Dissanayake, Said Al-Sarawi and Derek Abbott
The School of Electrical and Electronic Engineering
The University of Adelaide
Australia. SA 5005

1. Introduction

MEMS technologies have made it possible to fabricate small size, and high performance implantable devices to meet critical medical and biological needs such as site specific *in–vivo* drug delivery, Lab–on–a–Chip (LoC), micro total analysis systems, and polymerase chain reaction (PCR).

Actuators are one of the important components in Bio-MEMS, especially for fluid manipulation. The design of a suitable actuator device to pump the fluid at the microscale, for accurate operation, is of great importance. Many types of microactuators have been developed to match different requirements for various applications (Tsai & Sue, 2007; Varadan & Varadan, 2000). With miniaturisation, physical scaling laws inherently favour some technologies and phenomena over others. In some cases, technologies that can be made by micromachining work well at the microscopic scale, but have no analogy or usefulness in the macroscopic scale. Moreover most of these actuators are too complicated to fabricate within a micropump structure.

Notably, Surface Acoustic Wave (SAW) devices are used to develop micromachines such as ultrasonic micromotors and fluid transfer methodologies such as flexural micropumps (Wixforth, 2003; Strobl et al., 2004). Currently available microfabrication technologies such as photolithography and X–ray lithography with a combination of other processes have enabled the use of SAW devices for a variety of self–contained MEMS applications, which have advanced functionality and performance. The key benefits of these micromachines are; their small size, ease of production, and low–cost. The use of SAWdevices for micro actuation applications provides the great benefit of controlling and interrogation of devices remotely, without direct physical user intervention (Dissanayake et al., 2007; Varadan & Varadan, 2000; Jones et al., 2008).

In this chapter, SAW based novel batteryless and low–powered, secure, and wireless interrogation as well as actuation mechanisms for implantable MEMS devices such as actuators are introduced and investigated. This approach is based on SAW technology and significantly different from currently existing techniques, as the proposed method consists of dual functionality; the secure interrogation and actuation. Consequently such a microactuator can be embedded in a microfluidic device to modulate the fluid flow using less power compared to other mechanisms, such as piezoelectric micropumps. In Section 2, the use of SAW devices for micro actuation is presented and discussed. Section 3 explains

the operation of the SAW device based microactuator. The underlying theoretical model is then elaborated in Section 4 and followed by Section 5, which presents a method to derive the electric potential for electrostatic actuation. Section 6 shows a theoretical boundary condition analysis for the proposed model. Section 7 presents detailed Finite Element Modelling (FEM) of the actuator. Then simulation results are discussed in Section 8, and followed by the conclusion in Section 9.

2. SAW device based microactuator

SAW devices are widely used in MEMS applications, which require secure, wireless, and passive interrogation Jones et al. (2008). These devices are recognised for their versatility and efficiency in controlling and processing electrical signals. They are based on propagation of acoustic waves in elastic solids and the coupling of these waves to electric charge signals via an input and an output Inter Digital Transducers (IDT) that are deposited on the piezoelectric substrate. As shown in Figure 1, a SAW device consisting of a solid substrate with input and output IDTs Jones et al. (2008). An IDT is an array of narrow and parallel electrodes connected alternately to two bus bars made out of thin–film metal. The purpose of placing a set of IDTs on a SAW device is to provide a coupling between the electrical signal received (or transmitted) and the mechanical actuation of the piezoelectric substrate material. Since SAW devices are mostly used for wireless applications, a micro–antenna is need to be attached to the input IDT.

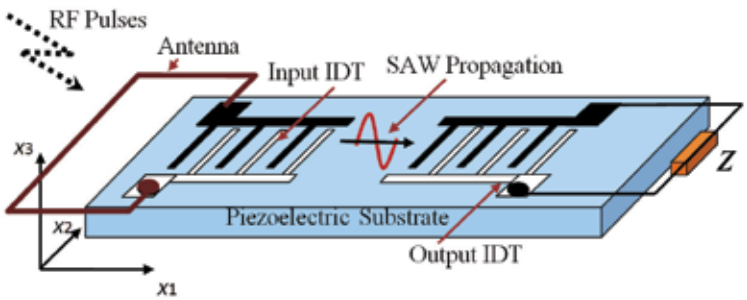


Fig. 1. Standard SAW device consist of a piezoelectric substrate, input IDT, and an output IDT. Input IDT is connected to a micro-antenna for wireless communication, and a load is connected to the output IDT for measurements.

Acoustic waves in these devices are propagating as surface waves, and hence can be perturbed easily by modifications to the substrate surface. Such features have enabled a large number of resonant sensors for applications such as chemical sensors Ruppel et al. (2002), gyroscopes Varadan & Varadan (2000), and accelerometers Subramanian et al. (1997). SAW devices also find application in oscillators, pulse compressors, convolvers, correlators, multiplexers and demultiplexers Ruppel et al. (2002).

SAW device related technology has been utilised to design and develop MEMS based microaccelerometers and gyroscopes for military and similar applications (Varadan & Varadan, 2000; Subramanian et al., 1997). The technology used in those applications is similar to the capacitor effect generated by programmable tapped delay lines, which use the principle of air gap coupling (Milstein & Das, 1979) between the SAW substrate and a silicon superstrate; a silicon layer superimposed on the SAW device. These capacitors are then used

to control the amount of RF coupling from the input IDT on the SAW substrate to the output terminal on the silicon chip (Subramanian et al., 1997). It is a well known method to use a sandwich structure of semiconductor on piezoelectric substrate to form the so called space–charge coupled SAW devices and SAW convolvers (Milstein & Das, 1979).
Such an approach can be utilised in the design of a SAW based microactuator. The proposed approach for the actuator design is converse to the method used by Varadan et. Al (Subramanian et al., 1997) for the microaccelerometer design. Being an elastic deformation wave on a piezoelectric substrate, the SAW induces charge separation. Thus it carries an electric field with it, which exists both inside and outside the piezoelectric substrate and decays according to Laplace's equation. In this SAW device based actuator, a thin conductive plate is placed on top of the output IDT, which is separated by an air–gap. The conductive plate does not alter the mechanical boundary conditions of the SAW substrate, but causes the surface to be equipotential and the propagating electric potential to be zero at the surface of the conductive plate. As a result, an electrostatic force is generated between the conductive plate and the output IDT in the SAW device causing micro deformations in the conductive plate.

3. Proposed microactuator operation

Figure 2 depicts the wireless interrogation unit for the SAW based microactuator. The actuator is made of a conductive material or alternatively, it can be made of a material such as Silicon (Si) or Silicon Nitride (Si_3N_4) and the bottom surface of the microactuator can be coated with a thin conductive material such as Gold, Platinum or Aluminium. The SAW substrate is made out of 128-YX-Lithium Niobate ($LiNbO_3$), as it is best suited for Rayleigh wave propagation.

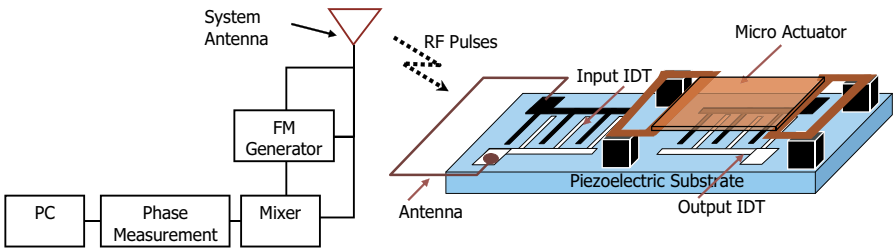


Fig. 2. Wireless interrogation unit for SAW device based actuator. The microactuator is placed on top of the output IDT of a SAW device. SAW device consists of a piezoelectric substrate, input IDT, and output IDT. Input IDT is connected to a micro-antenna for wireless interrogation.

Effectively, the output IDT and the conductive plate are used to generate an air–gap coupled SAW based electrostatic actuator. The device operation is as follows. The input IDT generates Rayleigh waves using inverse piezoelectric effect based on the RF signal that is being fed to the SAW device through the microstrip antenna. The Output IDT regenerates the electric signal using the piezoelectric effect of the SAW device. As it was explained in Section 2, the generated electrostatic field between this propagating electric potential wave and the conductive plate on top of the output IDT creates a compulsive and repulsive force between the two. Since the conductive plate is a thin flexural plate, it bends as a function of the applied electrostatic field enabling its use as a microactuator.

4. Theoretical analysis of the electrostatic actuation

In this section, a detailed theoretical analysis is carried out to analyse the generated electrostatic force at the conductive actuator. In achieving this, first the behavior of the propagating SAW on piezoelectric media is analysed. Then the analysis is focused towards deriving a general expression for the electric potential regenerated at the output IDT of the SAW device. Additionally, a boundary condition analysis is also presented, considering a specific scenario; the placement of the conductive actuator on top of the output IDT as shown in Figure 3.

4.1 Problem description

A variety of approaches have been followed by other researchers to investigate the SAW propagation in different piezoelectric substrates. These include, but not limited to, Potassium Niobate ($KNbO_3$), Lithium Tantalate ($LiTaO_3$), Lead Potassium Niobate ($Pb_2KNb_5O_{15}$ or PKN) (Dvoesherstov et al., 2000; Dvoesherstov & Chirimanov, 1999), and various cuts in Lithium Niobate $LiNbO_3$ (Gardner et al., 2001). However, the requirement of a detailed analysis of the use of 128-XY $LiNbO_3$, specific to the novel SAWbased actuator is of great importance. Therefore a descriptive theoretical analysis was carried out and presented in following sections.

In a SAW device, IDT patterns can be considered as a periodic structure. When an acoustic wave propagates on the surface through the periodic structure, it is partially reflected at each IDT finger. Depending on the operating frequency of the acoustic wave, the reflected parts interfere constructively or destructively (Zaglmayr et al., 2005). Generally, these reflections are considered to be very small and therefore, in this analysis the effect of the reflections is discarded and it is assumed that a surface wave propagate through each IDT finger only once. As can be seen from Figure 3, the direction of periodicity is denoted by x_1, the surface normal direction by x_3, and their perpendicular direction by x_2 following a *right-handed* coordinate system. The dimensional extension of electrodes in x_2 direction (length of the IDT fingers) is much larger in comparison to the periodicity. Additionally, a homogenous material topology is assumed in x_2 direction for this analysis.

4.2 Rayleigh waves

In general, SAWs consist of different types of propagation modes that are considered to be 3D. It is highly important to choose a suitable propagation mode for the SAW device especially when it is designed for microfluidic applications. Therefore, different acoustic modes need to be analysed. For example, one would expect horizontal shear SAW mode to be more useful than Rayleigh SAW mode. This is because when a SAW device is in vitro, Rayleigh SAW mode is rapidly damped out, due to the pressure generated by the surrounding fluid (Jones et al., 2008). In contrast to the general approach of using the SAW device surface to transfer the fluid, the proposed approach uses an actuator (conductive plate), which is on top of the SAW substrate. As a result the fluid flow is isolated from the SAW substrate and hence it prohibits the interaction of the fluid with the SAW mode. Moreover, Rayleigh SAW mode is best suited for space-charge related applications as most of the energy in this mode is concentrated within one wavelength of the substrate (Milstein & Das, 1979), as shown in Figure 4. Hence, considering the measures taken to isolate the SAW substrate from the fluid in this approach, Rayleigh SAW mode was chosen.

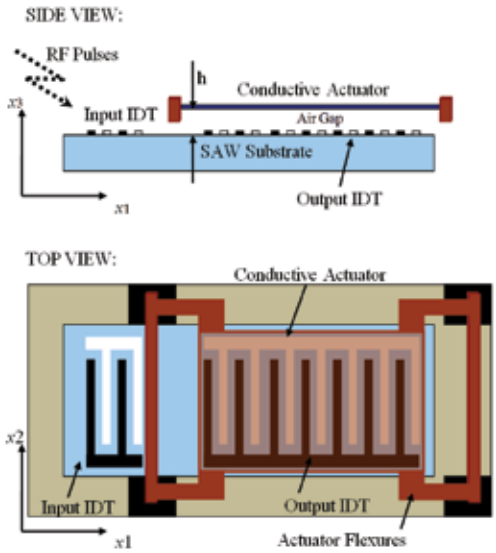


Fig. 3. A concept drawing of the SAW actuator model. Side view: Air–gap separated conductive actuator is placed above the output IDT of the SAW device (not to scale). Top view: The actuator consists of flexures to reduce the spring constant. SAW substrate is a rotated crystal cut to make sure the wave propagation in x_1 direction and h is the height of the air gap.

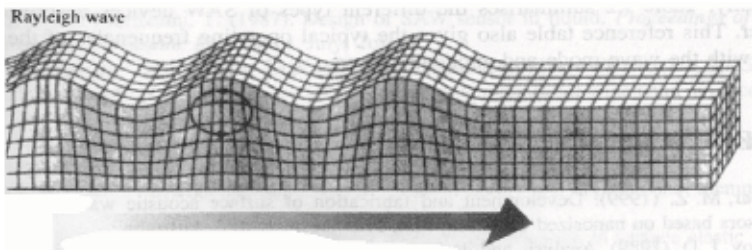


Fig. 4. Rayleigh Wave Mode which shows elliptical particle displacements (Gardner et al., 2001).

Rayleigh waves propagates near the surface of the piezoelectric substrate, the amplitude decreases rapidly within depth in x_3 direction, and becomes negligibly small within the depth of a few wavelengths from the surface (Zaglmayr et al., 2005). This behaviour is further elaborated and mathematically represented in coming sections.

The orientation of the piezoelectric material affects the SAW propagation on the substrate. Therefore, a certain crystal cut is required to direct the surface wave towards a desired direction. Here for this analysis, 128-XY $LiNbO_3$ is chosen as the substrate material due to its high electromechanical coupling coefficient (Čiplys & Rimeika, 1999; Gardner et al., 2001). This orientation facilitate the Rayleigh mode to propagate in x_1 direction and contained in a sagittal plane define by x_1–x_3 plane. Therefore, instead of performing a full three-dimensional analysis, a model reduction is performed in the geometric domain and the analysis is carried out within the sagittal plane.

In order to derive an expression for the resultant electrostatic force, the following assumptions are made of the model and the analysis, as well as simplifications to both are mentioned. In piezoelectric materials, the frequency of the induced electric field wave is sufficiently small enough to reasonably assume the electromagnetic coupling effects to be negligible. This means that local perturbations are felt almost instantaneously through out the substrate (Gantner et al., 2007; Kannan, 2006; Maugin, 1985).
Additionally, if l is the largest characteristic dimension of the actuator structure and c is the speed of light, the generated electromagnetic coupling effects can be safely discarded, if the operating frequency of the device is much less than the ratio c/l (Zaglmayr et al., 2005; Horenstein et al., 2000). Here, for $l \sim 3$ mm, $c/l \approx 100$ GHz. Therefore, since the chosen signal frequency of SAW is in the range of 50 – 100 MHz ($\ll c/l$), electromagnetic coupling effects can be safely discarded, and hence the electric field can be treated as quasi–static.

4.3 Piezoelectric equations

In piezoelectric analysis, quasi–static electric field is achieved by setting the permeability to zero, corresponding to an infinite speed of the electromagnetic wave. Using Maxwell's equations, electromagnetic quasi–static approximation can be written as

$$E_i = -\frac{\partial \Phi}{\partial x_i}, \tag{1}$$

$$\nabla \cdot D = 0, \tag{2}$$

where E is the electric field intensity, Φ is the electric potential, D is the electric flux density (dielectric displacement) and x_i (for i = 1, 2, 3) shows the direction of interest as shown in Figure 3. Here, Equation 2 elaborate the fact that piezoelectric materials are insulators hence there are no free volume charges (Zaglmayr et al., 2005). Moreover, from the mechanical equations of motion, the relationship between the mechanical displacement u, and the mechanical stress T is given by

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \sum_j \frac{\partial T_{ij}}{\partial x_i}, \tag{3}$$

where ρ is the density of the piezoelectric substrate. For a piezoelectric media, the relation between the mechanical strain S and mechanical displacement u can be written as

$$S_{ij} = \frac{1}{2}\left[\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right]. \tag{4}$$

In order to investigate the behaviour of the electric potential and the mechanical displacement of SAW, the constitutive equations for piezoelectric material are also need to be considered. The constitutive equations for piezoelectric material can be obtained by extending Hook's law and the electrostatic equation for the electric flux density. Therefore, the relationship between D, E, T and S in a SAW substrate can be written as

$$T_{ij} = \sum_k \sum_l c_{ijkl}^E S_{kl} - \sum_k e_{kij} E_k, \tag{5}$$

$$D_i = \sum_j \sum_k e_{ijk} S_{jk} + \sum_j \varepsilon_{ij}^S E_j, \tag{6}$$

where c_{ijkl} is the stiffness tensor for constant electric field, e_{ijk} is the piezoelectric coupling tensor (e_{kij} is the corresponding transpose tensor of e_{ijk}), ε_{ij} is the permittivity tensor for constant strain and i, j, k, l vary from 1 to 3. It should be noted that the mechanical stiffness matrix [$\boldsymbol{c}$] and the permittivity matrix [$\boldsymbol{\varepsilon}$] are symmetric. Additionally, the direct and converse piezoelectric effects are also symmetric, hence in piezoelectric matrix [$\boldsymbol{e}$], the coupling coefficients are equal for both the effects (Zaglmayr et al., 2005). These symmetry considerations allow the reduction of the corresponding matrices of the material tensors; [$\boldsymbol{c}$] to a 6×6 symmetric matrix, [$\boldsymbol{\varepsilon}$] to a 3×3 symmetric matrix, and [$\boldsymbol{e}$] to a 6×3 matrix.

Equations 1– 6 lead to a system of four coupled equations, which are represented by Equations 7 and 8.

$$\sum_j \sum_k \left[e_{kij} \frac{\partial^2 \Phi}{\partial x_j \partial x_k} + \sum_l c_{ijkl}^E \frac{\partial^2 u_k}{\partial x_j \partial x_l} \right] = \rho \frac{\partial^2 u_i}{\partial t^2}, \tag{7}$$

$$\sum_i \sum_j \left[\varepsilon_{ij}^S \frac{\partial^2 \Phi}{\partial x_i \partial x_j} - \sum_k e_{ijk} \frac{\partial^2 u_j}{\partial x_i \partial x_k} \right] = 0. \tag{8}$$

This relationship between mechanical displacement and electric potential is considered in a semi–infinite, isotropic and homogenous linear elastic space as shown in Figure 4. As was mentioned above, the piezoelectric crystal cut ensures the Rayleigh wave propagate in the x_1-x_3 sagittal plane as shown in Figure 3. An alternative definition to the sagittal plane, can be given as, the plane defined by the real wave vector (wave number) k and the unit vector normal to the surface of the SAW substrate (x_3 direction).

Various approaches have been proposed by researchers to solve wave propagation on anisotropic substrates (Zaglmayr et al., 2005; Gantner et al., 2007; Adler, 2000). The method of partial waves is considered to be a commonly used technique to analyse different SAW modes on anisotropic substrates such as piezoelectrics. Therefore, in this research the method of partial waves is used to solve this wave propagation phenomena for the SAW actuator model. As a result, plane wave solutions of the form given in Equations 9 and 10 are considered for the mathematical modelling of this device,

$$u_j^m(x_1, x_3, t) = \alpha_j^m e^{\mathrm{i}kb^m x_3} e^{\mathrm{i}k(x_1 - vt)}, \tag{9}$$

$$\Phi^m(x_1, x_3, t) = \alpha_4^m e^{\mathrm{i}kb^m x_3} e^{\mathrm{i}k(x_1 - vt)}, \tag{10}$$

where the α_j^m values are linear coefficients that depend on the decaying constant b_m is the phase velocity of the wave, k (= $2\pi/\lambda$) is the wave vector, λ is the wavelength, i is the standard imaginary unit (= $\sqrt{-1}$), m = 1, 2, 3, 4 and j = 1, 2, 3.

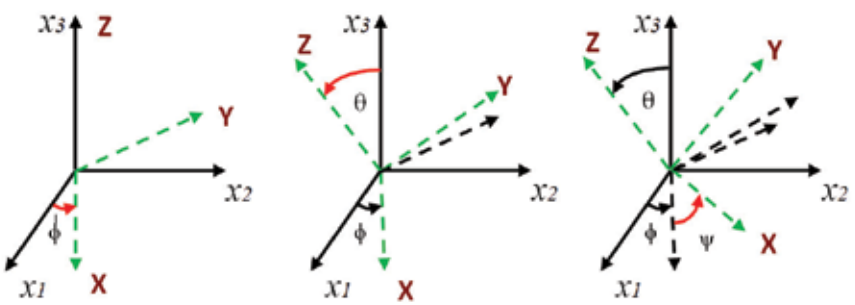
Equations 7 and 8, a linear system for the coefficients α_j^m can be obtained and solved. In solving these equations, it is necessary to transform the material parameters of the SAW substrate to match the coordinate system of the problem. This is discussed in Section 4.4.

4.4 Transformation of coordinates

As it was mentioned above, different crystal cuts consist of different material properties. Therefore it is significant to consider the specific orientation of the selected crystal for a specific analysis, to obtained the desired SAW mode. In the equations of motion the material parameters are expressed in terms of structural coordinate axes that are selected for convenient boundary condition and excitation requirements. However, the material parameters are presented in the form of [c], [ε], and [e] matrices as was mentioned in Section 4.3, and these are expressed according to the crystalline axes. Therefore it is necessary to transform the material parameters to match the coordinate system of the problem. Generally, the parameters are transformed using a transformation matrix [r] given by

$$[r]=\begin{bmatrix} \cos\psi.\cos\phi. & \cos\psi\sin\phi. & \sin\psi\sin\theta \\ -\cos\theta\sin\phi\sin\psi & +\cos\theta\cos\phi\sin\psi & \\ \sin\psi.\cos\theta & -\sin\psi\sin\phi & \cos\psi\sin\theta \\ -\cos\theta\sin\phi\cos\psi & +\cos\theta\cos\phi\cos\psi & \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix},$$

where the elements of this matrix are the direction cosines between the crystalline axis and the problem axis (Subramanian et al., 1997; Kannan, 2006; Wolfram MathWorld, 2009).
In this research, the X–convention is followed for specifying the order of rotation of the axes. The rotations specified by the Euler angle set that are related to the X-convention and the rotation order are explained in Figure 5.



(a) (b) (c)

Fig. 5. Transformation of coordinate system. Transformation of coordinate system using **X**–convention. Here $x_1x_2x_3$ is the structural coordinate system and XYZ is the crystal coordinate system. Here x_1 determines the wave propagation direction and x_3 is normal to the crystal surface. Initially both of these axes are parallel to each other. While the structural coordinate axes always remain same, the crystal axes XYZ are rotated by an angle ϕ around Z axis, then by an angle θ around X axis and finally by an angle ψ once again around Z axis. The direction cosines are then derived based on the Euler angles.

Once the transformation matrix is defined, the matrices that define the crystal properties ([c], [ε], and [e]) are then transformed using the Bond–Transformation procedure (Upadhyay, 2004). The advantage associated with the this procedure, for stiffness and

compliance, is that Bond-Transformation is directly applied on stiffness or compliance constants given in *abbreviated subscript notation*. Therefore the transformation laws of this procedure can be shown as

$$[\boldsymbol{c}_r] = [\boldsymbol{M}_B][\boldsymbol{c}][\boldsymbol{M}_B^T], \tag{11}$$

$$[\boldsymbol{e}_r] = [\boldsymbol{M}_B][\boldsymbol{e}][\boldsymbol{r}^T], \tag{12}$$

$$[\varepsilon_r] = [\boldsymbol{r}][\varepsilon][\boldsymbol{r}^T], \tag{13}$$

where $[\boldsymbol{M}_B]$ is the Bond–Transformation matrix given by

$$[\boldsymbol{M}_B] = \begin{bmatrix} r_{11}^2 & r_{12}^2 & r_{13}^2 & 2r_{12}r_{13} & 2r_{11}r_{13} & 2r_{11}r_{12} \\ r_{21}^2 & r_{22}^2 & r_{23}^2 & 2r_{22}r_{23} & 2r_{21}r_{23} & 2r_{21}r_{22} \\ r_{31}^2 & r_{32}^2 & r_{33}^2 & 2r_{32}r_{33} & 2r_{31}r_{33} & 2r_{31}r_{32} \\ r_{21}r_{31} & r_{22}r_{32} & r_{23}r_{33} & r_{22}r_{33}+r_{32}r_{23} & r_{21}r_{33}+r_{31}r_{23} & r_{21}r_{32}+r_{31}r_{22} \\ r_{11}r_{31} & r_{12}r_{32} & r_{13}r_{33} & r_{12}r_{33}+r_{32}r_{13} & r_{31}r_{13}+r_{11}r_{33} & r_{11}r_{32}+r_{31}r_{12} \\ r_{11}r_{21} & r_{12}r_{22} & r_{13}r_{23} & r_{12}r_{23}+r_{22}r_{13} & r_{11}r_{23}+r_{21}r_{13} & r_{11}r_{22}+r_{21}r_{12} \end{bmatrix}.$$

$[\boldsymbol{r}_T]$ denotes the transpose of the matrix $[\boldsymbol{r}]$. $[\boldsymbol{c}_r]$ is the transformed stiffness matrix, $[\boldsymbol{e}_r]$ is the transformed piezoelectric matrix and $[\boldsymbol{\varepsilon}_r]$ is the transformed permittivity matrix. These transformed matrices can be used to derive more specific partial wave equations for a certain SAW mode for a known crystal cut as shown in the following section.

4.5 Calculations for lithium niobate crystals

It is highly important to select a piezoelectric material with high electromechanical coupling in order to maximise the signal strength of the electric potential at the output IDT of the SAW device. Therefore, 128-YX-$LiNbO_3$ is chosen for this device, as it consists of the highest electromechanical coupling efficiency for Rayleigh SAW modes, compared to other commonly used piezoelectric substrates (Čiplys & Rimeika, 1999). Following the crystal transformation procedure described in Section 4.4, the coordinations of substrate is need to be transformed into a coordinate system defined by x_1, x_2 and x_3 as shown in Figure 3. This results in a new set of material parameters for the substrate, which is then used during the derivation of electric potential and mechanical displacement functions of the SAW mode.

The Euler angles associated with the crystal rotation for 128-YX-$LiNbO_3$ is $(0,38^o,0)$, and the rotated material parameters take the general form as presented below.

$$[\boldsymbol{c}_r] = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} & 0 & 0 \\ c_{12} & c_{22} & c_{23} & c_{24} & 0 & 0 \\ c_{13} & c_{23} & c_{33} & c_{34} & 0 & 0 \\ c_{14} & c_{24} & c_{34} & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{55} & c_{56} \\ 0 & 0 & 0 & 0 & c_{56} & c_{66} \end{bmatrix}, \quad [\boldsymbol{e}_r] = \begin{bmatrix} 0 & e_{12} & e_{13} \\ 0 & e_{22} & e_{23} \\ 0 & e_{32} & e_{33} \\ 0 & e_{42} & e_{43} \\ e_{51} & 0 & 0 \\ e_{61} & 0 & 0 \end{bmatrix}, \quad [\varepsilon_r] = \begin{bmatrix} \varepsilon_{11} & 0 & 0 \\ 0 & \varepsilon_{22} & \varepsilon_{23} \\ 0 & \varepsilon_{23} & \varepsilon_{33} \end{bmatrix}$$

In order to derive the specific plane wave equations for this crystal, partial wave Equations 9 and 10 are substituted into the Equations 7 and 8 with the rotated material parameters. As a result, the following Eigenvalue problem can be obtained.

$$\mathbf{M} \cdot \boldsymbol{\alpha} = 0. \tag{14}$$

Alternatively the above equation can be written as,

$$\begin{bmatrix} m_{11} - \rho v^2 & m_{12} & m_{13} & m_{14} \\ m_{12} & m_{22} - \rho v^2 & m_{23} & m_{24} \\ m_{13} & m_{23} & m_{33} - \rho v^2 & m_{34} \\ m_{14} & m_{24} & m_{34} & m_{44} - \rho v^2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} = 0. \tag{15}$$

The matrix coefficients in Equation 15 are shown below. It should be noted that for a 128-YXLiNbO$_3$ crystal c_{15}, c_{16}, c_{35}, c_{36}, c_{45}, and $c_{46} = 0$. Additionally, e_{11}, e_{31}, e_{41}, e_{53}, e_{63} and $\varepsilon_{13} = 0$. Therefore as shown below, the matrix coefficients in **M** become simpler and the Eigenvalue problem expressed in Equation 15 becomes easier to solve.

$$\begin{aligned}
m_{11} &= c_{55}b^2 + 2c_{15}b + c_{11} && \Rightarrow && m_{11} = c_{55}b^2 + c_{11} \\
m_{12} &= c_{45}b^2 + (c_{14} + c_{56})b + c_{16} && \Rightarrow && m_{12} = (c_{14} + c_{56})b \\
m_{13} &= c_{35}b^2 + (c_{13} + c_{55})b + c_{15} && \Rightarrow && m_{13} = (c_{13} + c_{55})b \\
m_{14} &= e_{53}b^2 + (e_{51} + e_{13})b + e_{11} && \Rightarrow && m_{14} = (e_{51} + e_{13})b \\
m_{22} &= c_{44}b^2 + 2c_{46}b + c_{66} && \Rightarrow && m_{22} = c_{44}b^2 + c_{66} \\
m_{23} &= c_{34}b^2 + (c_{36} + c_{45})b + c_{56} && \Rightarrow && m_{23} = c_{34}b^2 + c_{56} \\
m_{24} &= e_{43}b^2 + (e_{41} + e_{63})b + e_{61} && \Rightarrow && m_{24} = e_{43}b^2 + e_{61} \\
m_{33} &= c_{33}b^2 + 2c_{35}b + c_{55} && \Rightarrow && m_{33} = c_{33}b^2 + c_{55} \\
m_{34} &= e_{33}b^2 + (e_{31} + e_{53})b + e_{51} && \Rightarrow && m_{34} = e_{33}b^2 + e_{51} \\
m_{44} &= -(\varepsilon_{33}b^2 + 2\varepsilon_{13}b + \varepsilon_{11}) && \Rightarrow && m_{44} = -(\varepsilon_{33}b^2 + \varepsilon_{11})
\end{aligned}$$

In order to evaluate the non–trivial solution for Equation 15, the Eigenvalue problem, *det* (**M**) = 0, is required to be solved. This results in a system of characteristic equations for displacement amplitudes and electric potential in which the phase velocity v of the wave is used as unknown parameters. In general case, this system of characteristic equations is reduced to an eighth order polynomial in the decaying constant b for a given value of phase velocity. However, the resulting roots of b are either purely real or complex conjugate pairs. Since these roots leads to Rayleigh waves that decay with the depth along x_3, only the roots with negative imaginary parts are accepted to be consistent with the physical meaning of wave propagation in piezoelectric media (Gardner et al., 2001; Ippolito et al., 2002). There are four such roots for b (denoted as b_m for m = 1, 2, 3, 4) and for each such value, there exist a unique Eigenvector $\boldsymbol{\alpha}^m$. A general solutions is then obtained as a linear combination of partial waves such that each wave decays almost to zero as it shifts into the crystal depth approximately at a distance of several wavelengths from the surface ($x_3 = 0$). Furthermore, the solution consists of three displacement components u_j (j = 1, 2,3) and the electric potential Φ as described by the Equations 16 and 17.

$$u_j(x_1, x_3, t) = \left[\sum_m C_m \alpha_j^m e^{ikb^m x_3} \right] e^{ik(x_1 - vt)}, \tag{16}$$

$$\Phi(x_1, x_3, t) = \left[\sum_m C_m \alpha_4^m e^{ikb^m x_3} \right] e^{ik(x_1 - vt)}. \tag{17}$$

Additionally, during the derivation of the solution in the form of a SAW, it is assumed that the wave vector k is always parallel to the free surface, while the constant–phase surfaces are always normal to the free surface. Furthermore, it is considered that the multiplier in Equations 16 and 17, which depends on the x_3 coordinate, determines the displacement amplitude, and the wave properties of the solution are determined by the multiplier $e^{ik(x_1 - vt)}$. The weighting coefficients C_m of these plane waves are chosen to satisfy the mechanical and electrical boundary conditions at the surface of the piezoelectric substrate specific to this SAW based actuator model, which is discussed in detail in Section 6. But first, in Section 5, an expression is derived for the electric potential at the output IDT, using Equation 17.

5. Electric potential at output IDT

In order to determine the electrostatic field generated between the output IDT and the conductive plate, the evaluation of the electric potential at the output IDT is required. Here, once the plane wave equation is evaluated for the electric potential wave in the SAW device (Equation 17), an analysis is carried out to evaluate an expression for the electric potential at the output IDT. In achieving this, the following assumptions and simplification are made to the design apart from previously mentioned simplifications.

i. The crystal cut is best set so that the SAW is directly propagating in the x_1 direction to allow an orthogonal interaction between the SAW and the output IDT.
ii. The orientation of the IDT is such that a SAW is generated in the direction of maximum SAW–propagation speed.
iii. The acoustic wave is assumed to pass the output IDT once, and interact with no reflections caused by impedance mismatches (Gardner et al., 2001; Ippolito et al., 2002).
iv. The acoustic wave interact with the IDT in the near field where SAW can be treated as a travelling wave as was discussed before.

Due to the periodic nature of the propagating waves and the placement of the IDTs, the analysis is initially carried out only for a single period, and then extended to the whole structure. The single period placement of the output IDT is shown in Figure 6.

As explained in Figure 6, each finger in the IDT is assigned a negative or positive value that is determined by the finger's connection to either positive or negative bus bar. Therefore, the output IDT fingers are represented as square waves with the period defining wave length of the SAW and the duty cycle defining the finger width (Skinner et al., 2006). This width is $\frac{\lambda}{4}$ for a metallisation ratio of 0.5. For the SAW based electrostatic actuator model, initially a basic model of the IDT configuration is used.

While the output IDT is considered as a combination of square waves, the SAW is defined as a propagating plane wave as was elaborated in Section 4.5. In order to obtain the time response of the SAW interaction with the output IDT, a cross–correlation is applied to the

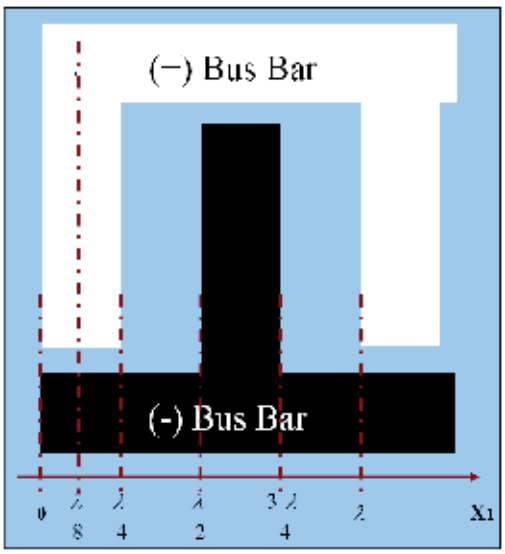


Fig. 6. Periodic IDT finger representation. IDT finger representation for one wavelength (λ) with a metalisation ratio of $\frac{1}{2}$. Each finger is connected to one of the two bus bars, with one bus bar assigned a positive value and the other assigned a negative value.

SAW and the IDT signals. Cross–correlation is a method by which two different function are compared over time with one time–fixed function and one time–shifted function. Within one wavelength of the IDT, the analysis has to be carried out in two parts, considering the space above the output IDT fingers ($0 \le x_1 \le \frac{\lambda}{4} \cup \frac{\lambda}{2} \le x_1 \le \frac{3\lambda}{4}$) and the space above the output IDT gap between fingers ($\frac{\lambda}{4} < x_1 < \frac{\lambda}{2} \cup \frac{3\lambda}{4} < x_1 < \lambda$). This is because the metal based IDT fingers consist of an equipotential distribution for a given time, and the gaps between the fingers consist of a space varying electric potential distribution in x_1 direction.

IDT Segment in Range ($0 \le x_1 \le \frac{\lambda}{4} \cup \frac{\lambda}{2} \le x_1 \le \frac{3\lambda}{4}$):

As shown in Figure 7, analysis of a single IDT finger will serve as the basis for the analysis for a single periodic structure and then for the entire IDT structure. Therefore, the cross-correlation between the SAW potential signal $\Phi(x_1, x_3, t)$, and the first finger connected to the positive bus bar ($0 \le x_1 \le \frac{\lambda}{4}$) can be expressed as

$$C_+(x_1, x_3, t) = (\Phi * g_{(+)})(x_1, x_3, t)$$

$$= \int_0^{T/4} \Phi(x_1, x_3, t+\tau) \cdot g_{(+)}(\tau) d\tau$$

$$= \frac{\sqrt{2}}{kv} \Phi(x_1 - \frac{\lambda}{8}, x_3, t). \quad (18)$$

Here, $g_{(+)}$ represents the positive square wave of the finger as can be seen from Figure 7.

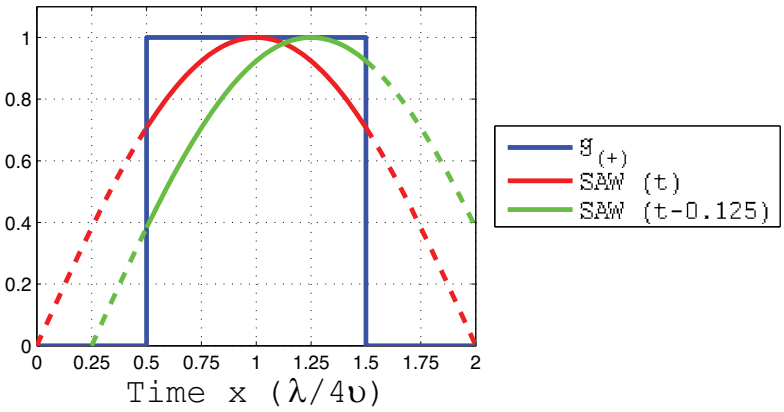


Fig. 7. SAW and single finger correlation. Superimposition of the SAW and a single finger of the output IDT. $g_{(+)}$ represents the equipotential behaviour of the conductive finger. Two SAWs are $T/8$ apart from one another.

The electric potential at this IDT finger can be considered as the average value of the correlated signal $C_+(x_1, x_3, t)$ over the finger width f_w, due to the equipotential nature of the metal based fingers. As the metalisation ratio is 0.5 in this analysis, $f_w = \lambda/4$ and this relate to $T/4$ in time scale, where T (= λ/v) is the time period of the SAW. Therefore, the electric potential generated at the first finger connected to the positive bus bar can be written as

$$\begin{aligned} V_+(x_1,x_3,t) &= \int_0^{\lambda/4} C_+(x_1,x_3,t)dx_1 \\ &= \frac{2T}{\pi^2}\Phi(\frac{\lambda}{8},x_3,t). \end{aligned} \tag{19}$$

By following an identical approach, the electric potential at the finger connected to the negative bus bar ($\frac{\lambda}{2} \le x_1 \le \frac{3\lambda}{4}$) can be derived and the result can be written as

$$V_-(x_1,x_3,t) = -\frac{2T}{\pi^2}\Phi(\frac{\lambda}{8},x_3,t). \tag{20}$$

IDT Segment in Range ($\frac{\lambda}{4} < x_1 < \frac{\lambda}{2} \cup \frac{3\lambda}{4} < x_1 < \lambda$):

Once the analysis is simplified by considering the aforementioned assumptions and simplifications, the electric potential at the gaps between the fingers can be considered to consist of the same electric potential of the propagating SAW as shown in Equation 17. Therefore

$$V_{gap}(x_1,x_3,t) = \Phi(x_1,x_3,t). \tag{21}$$

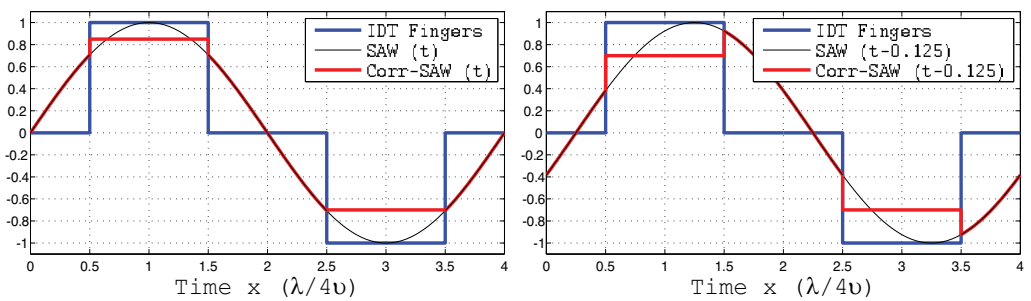
Based on the above analysis, the total electric potential generated by a single period of the output IDT can be expressed as

$$\Phi(x_1,x_3,t)=\begin{cases}\Psi, & for\ \ 0\le x_1\le\frac{\lambda}{4}\\ \Omega, & for\ \ \frac{\lambda}{4}<x_1<\frac{\lambda}{2}\\ -\Psi, & for\ \ \frac{\lambda}{2}\le x_1\le\frac{3\lambda}{4}\\ -\Omega, & for\ \ \frac{3\lambda}{4}<x_1<\lambda\end{cases}, \tag{22}$$

where $\Psi = V_+(x_1, x_3, t) = \frac{2T}{\pi^2}\,\Phi(\frac{\lambda}{8}, x_3, t)$, and $\Omega = V_{gap}(x_1, x_3, t) = \Phi(x_1, x_3, t)$. This is further elaborated graphically in Figure 8. Consequently, due to the periodic nature of the IDT, this expression can be easily extended to derive the electric potential for the full output IDT. Therefore, for an output IDT with N_p finger pairs, the total electrostatic potential at the output IDT can be expressed as

$$\Phi(x_1,x_3,t)=\begin{cases}\Psi, & for\ \ n\lambda\le x_1\le(\frac{1}{4}+n)\lambda\\ \Omega, & for\ \ (\frac{1}{4}+n)\lambda<x_1<(\frac{1}{2}+n)\lambda\\ -\Psi, & for\ \ (\frac{1}{2}+n)\lambda\le x_1\le(\frac{3}{4}+n)\lambda\\ -\Omega, & for\ \ (\frac{3}{4}+n)\lambda<x_1<(1+n)\lambda\end{cases}. \tag{23}$$

Here, $n = 0, 1, 2, ..., (N_p - 1)$ and Ψ and Ω are as explained before.
Once the general expressions is derived, a boundary condition analysis is carried out to specify values for the weighting coefficients in Equations 22 and 23. This analysis is further unfolded in next section.



(a) time = t (b) time = $t - T/8$

Fig. 8. SAW correlation and the electric potential at the output IDT. Correlation between SAW electric potential and the output IDT of the SAW device is demonstrated. For a periodic IDT structure, one finger pair is represented, hence one time period ($T = \lambda/v$) is considered. Equipotential IDT fingers are represented by square waves. (a) Electric potential of the propagating SAW, $SAW(t)$ is peaked at the center of output IDT fingers (considered at time t). (b) Electric potential of the propagating SAW is $T/8$ seconds delayed compared to $SAW(t)$. In both the cases, the cross-correlated electric signal consists of equipotentials across IDT fingers.

6. Boundary condition analysis

As it was highlighted above, the weighting coefficients in Equations 16 and 17 need to be determined based on the electrical and mechanical boundary conditions applicable for the SAW device based actuator model. Therefore, in this section a detailed analysis is presented to determine these weighting coefficients.

Mechanical boundary conditions:

At the output IDT area of the SAW substrate, the IDT mass is taken to be negligible for simplicity, so that the mechanical force acting on the SAW substrate can be discarded. Hence the surface is considered to be mechanically free. This can be mathematically expressed as

$$\sum_j T_{3j} = 0, \tag{24}$$

where T is the mechanical stress tensor and j = 1, 2, 3.

Electrical boundary conditions:

In the SAW based electrostatic actuator model, the electric potential generated at the output IDT region was investigated and evaluated in Section 5. Based on Equation 23, at the surface where the output IDT is deposited (x_3 = 0), the electrical boundary condition constitute an electric potential at the IDT fingers, and the electric potential at the gaps between the IDT fingers. Respectively, these regions can be noted as ($0 \le x_1 \le \frac{\lambda}{4} \cup \frac{\lambda}{2} \le x_1 \le \frac{3\lambda}{4}$) and ($\frac{\lambda}{4} < x_1 < \frac{\lambda}{2} \cup \frac{3\lambda}{4} < x_1 < \lambda$) which correspond to Figure 6 for one period of the output IDT structure. As explained in Section 5, the electric potential wave at the output IDT fingers act as an equipotential time varying wave throughout the positive and negative IDT fingers (independent of x_1). Whereas the electric potential wave at the gap between the fingers, still a time varying and a moving wave (a function of x_1).

However, in the SAW device based actuator model (Figure 3), the conductive plate is placed at a height h above the output IDT, and connected to the common ground of the device. Therefore, the electric potential approaches zero at $x_3 = h$. It should be noted that the electric potential above the SAW substrate satisfies Laplace's equation, which results in an exponentially decaying electric signal in x_3 direction, and can be written as

$$\Phi_L(x_1, x_3, t) = \left[A_L e^{kx_3} + B_L e^{-kx_3} \right] e^{ik(x_1 - vt)}, \tag{25}$$

where A_L and B_L are constants. Additionally, the electric potential and the electric flux density in x_3 direction are continuous at the surface (x_3 = 0) (Subramanian et al., 1997; Maugin, 1985), which can be mathematically expressed as

$$\begin{aligned} \Phi(x_3 = 0^-) &= \Phi(x_3 = 0) = \Phi(x_3 = 0^+), \\ D(x_3 = 0^-) &= D(x_3 = 0) = D(x_3 = 0^+). \end{aligned} \tag{26}$$

Equation 23 along with conditions highlighted in Equation 26 are considered together to eliminate both A_L and B_L in Equation 25. First, the electrical boundary condition at the conductive plate is considered. Since the conductive plate mounted at a distance h above the substrate, the electric potential at $x_3 = h$ becomes zero. Therefore, from Equation 25,

$$\begin{aligned} \Phi_L(x_1,h,t) &= \left[A_L e^{kh} + B_L e^{-kh}\right] e^{ik(x_1-vt)} = 0, \\ B_L &= -A_L e^{2kh}. \end{aligned} \tag{27}$$

Once B_L is eliminated, Equation 25 can be rewritten as,

$$\Phi_L(x_1,x_3,t) = A_L\left[e^{kx_3} - e^{2kh-kx_3}\right] e^{ik(x_1-vt)}. \tag{28}$$

To evaluate the constant A_L, electrical boundary condition at the surface of the substrate ($x_3 = 0$) is need to be considered. However, as shown in Equation 23, the electric potential at the output IDT area is a combination of various potentials due to the output IDT and the gap between them. Therefore, two different electric potential signals have to be considered in the analysis, one considering the space above the output IDT ($0 \le x_1 \le \frac{\lambda}{4} \cup \frac{\lambda}{2} \le x_1 \le \frac{3\lambda}{4}$) and the other, the space above the output IDT gap ($\frac{\lambda}{4} < x_1 < \frac{\lambda}{2} \cup \frac{3\lambda}{4} < x_1 < \lambda$). More importantly, this distinction does not affect the final result due to the similarity in the procedure followed for the derivation of A_L. Therefore the approach is shown only for one case.

Considering the plane wave Equation 23 for electric potential and Equation 28, and the continuity Equation 26, A_L can be evaluated as follows,

$$\begin{aligned} \Phi_L(x_1,0,t) &= \Phi(x_1,0,t), \\ A_L\left[1-e^{2kh}\right].e^{ik(x_1-vt)} &= \left[\sum_m C_m \alpha_4^m\right].e^{ik(x_1-vt)}, \text{ and} \\ A_L &= \frac{\sum_m C_m \alpha_4^m}{\left[1-e^{2kh}\right]}. \end{aligned} \tag{29}$$

The relationship between the electric flux density and the electric field can be written as $D = \varepsilon_0 E$. Considering this relation along with Equation 1, a relationship between the electric flux density and the electric potential in the gap in x_3 direction can be written as

$$D_3(x_1,x_3,t) = -\varepsilon_0 \frac{\partial \Phi_L(x_1,x_3,t)}{\partial x_3}. \tag{30}$$

Hence, using the above equation, the electric flux density at $x_3 = 0$ is calculated and written as follows.

$$D_3(x_1,0,t) = -k\varepsilon_0 \frac{\sum_m C_m \alpha_4^m \left[1+e^{2kh}\right]}{\left[1-e^{2kh}\right]} e^{ik(x_1-vt)}. \tag{31}$$

Additionally, another alternative expression for electric flux density at $x_3 = 0$ can be obtained from Equation 6. Therefore,

$$D_3(x_1,0,t) = \sum_j \sum_k e_{3jk} S_{jk} + \sum_j \varepsilon_{3j}^S E_j. \tag{32}$$

Ultimately, by equating Equations 31 and 32, another Eigenvalue problem is formulated, where the variables consist of the weighting coefficients. The resulting boundary conditions can be written out in matrix form as follows.

$$\begin{bmatrix} \mathbb{G} & 0 & \mathbb{H} & \mathbb{I} \\ \mathbb{J} & 0 & \mathbb{L} & \mathbb{N} \\ 0 & \mathbb{P} & 0 & 0 \\ \mathbb{Q} & 0 & \mathbb{R} & \mathbb{Y} \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} = 0, \tag{33}$$

where

$$\mathbb{G} = (c_{13} - c_{14}b^1)\alpha_1^1 + (-c_{14} + c_{11}b^1)\alpha_3^1 + (e_{13} - e_{22}b^1)\alpha_4^1$$
$$\mathbb{H} = (c_{13} - c_{14}b^3)\alpha_1^3 + (-c_{14} + c_{11}b^3)\alpha_3^3 + (e_{13} - e_{22}b^3)\alpha_4^3$$
$$\mathbb{I} = (c_{13} - c_{14}b^4)\alpha_1^4 + (-c_{14} + c_{11}b^4)\alpha_3^4 + (e_{13} - e_{22}b^4)\alpha_4^4$$
$$\mathbb{J} = c_{44}b^1\alpha_1^1 + (c_{44} - c_{14}b^1)\alpha_3^1 + e_{51}\alpha_4^1$$
$$\mathbb{L} = c_{44}b^3\alpha_1^3 + (c_{44} - c_{14}b^3)\alpha_3^3 + e_{51}\alpha_4^3$$
$$\mathbb{N} = c_{44}b^4\alpha_1^4 + (c_{44} - c_{14}b^4)\alpha_3^4 + e_{51}\alpha_4^4$$
$$\mathbb{P} = (c_{14} + c_{66}b^2)\alpha_2^2$$
$$\mathbb{Q} = e_{51}b^1\alpha_1^1 + (e_{51} + e_{22}b^1)\alpha_3^1 - (\varepsilon_{11}b^1 + \mathrm{i}\varepsilon_0\mathbb{X})\alpha_4^1$$
$$\mathbb{R} = e_{51}b^3\alpha_1^3 + (e_{51} + e_{22}b^3)\alpha_3^3 - (\varepsilon_{11}b^3 + \mathrm{i}\varepsilon_0\mathbb{X})\alpha_4^3$$
$$\mathbb{Y} = e_{51}b^4\alpha_1^4 + (e_{51} + e_{22}b^4)\alpha_3^4 - (\varepsilon_{11}b^4 + \mathrm{i}\varepsilon_0\mathbb{X})\alpha_4^4$$
$$\mathbb{X} = \frac{1 + e^{2hk}}{1 - e^{2hk}}.$$

Here the value of $\mathbb{X}$ depends on the position of the conductive plate above the SAW device (h), and takes the value of –1 for the case where the plate is at an infinite height above the substrate ($h \to \infty$). In this analysis, the phase velocity is explicitly present in Equations 16 and 17, and implicitly present in the roots b^m and in the linear coefficients α^m as mentioned before. Therefore a suitable value for phase velocity should be chosen for which the determinant of the coefficients in the Eigenvalue problem presented in Equation 33 vanishes, hence satisfy the associated boundary conditions.

It is important to realise that the above two Eigenvalue problems presented in Equations 15 and 33 are required to be solved simultaneously using iterative numerical procedures. This is due to the implicit dependency of decaying constants b^m on phase velocity v, and explicit dependency of Eigenvectors of linear coefficients α^m on v. Once a suitable phase velocity is found, the weighting coefficients C_m can be determined. As a result, a complete solution is obtained for the electric potential at the output IDT (Equation 23).

6.1 Electrostatic force generation

In electrostatic actuation, the electrostatic force applied on electrostatic plates can be described using the parallel plate capacitor effect (Tsai & Sue, 2007) as

$$F = \frac{1}{2}\frac{\varepsilon A \Phi^2}{(h - W_P)^2}, \tag{34}$$

where ε is the dielectric coefficient of the medium between the plates, A is the effective plate area, $W_P(x_1)$ is the instantaneous deflection of the actuator in x_3 direction, h is the initial plate spacing, and Φ is the applied electric potential between the plates.

In order to carry out the analysis to derive an expression for the resultant electrostatic force, the assumptions and simplifications mentioned in Section 4.2 are applied. Additionally, the electric field lines produced by the positive IDT fingers terminate either at the negative IDT fingers or at the conductive plate. For simplicity however, the effect of the electrostatic coupling between the IDT fingers, as well as the fringe capacitances (between the electrodes and the diaphragm), is discarded in this analysis.

Due to the periodic nature of the propagating waves and the placement of the IDTs, the electrostatic force analysis is initially carried out only for a single period, and then extended to the whole structure, similar to the electric potential calculation that is presented in Section 5. The single period placement of the output IDT is shown in Fig. 6.

Previously in Section 5, different segments of the output IDT were considered in analyzing the electric potential at output IDT. In this section, a similar approach is followed also in analysing the electrostatic force at the output IDT. Here, within one wavelength of the IDT, the analysis is carried out in two parts; one part considering the region above the output IDT electrodes ($0 \le x_1 \le \frac{\lambda}{4} \cup \frac{\lambda}{2} \le x_1 \le \frac{3\lambda}{4}$) and other, the region above the output IDT finger gaps ($\frac{\lambda}{4} < x_1 < \frac{\lambda}{2} \cup \frac{3\lambda}{4} < x_1 < \lambda$). This is due to the fact that was proven in Section 5; IDT fingers consisting of an equipotential distribution for a given time instance, while the gaps between the fingers consisting of a space varying electric potential distribution in x_1 direction as shown in Equation 23. In this analysis, an IDT with finger width of f_w and finger length of f_l is considered. For a metallisation ratio of 0.5 as in Fig. 6, the finger spacing is also f_w.

IDT Segment in Range ($0 \le x_1 \le \frac{\lambda}{4} \cup \frac{\lambda}{2} \le x_1 \le \frac{3\lambda}{4}$):

The electrostatic force generated by the electrode finger, which is connected to the positive bus bar $F_{(+)}$ can be evaluated considering Equations 23 and 34. As a result of the quadratic dependency of force to the applied electric potential, the force generated between the plate and the electrode finger, which is connected to the negative bus bar $F_{(-)}$ is equal to $F_{(+)}$. Therefore,

$$F_{(+)} = F_{(-)} = \frac{\varepsilon_0 f_l f_w}{2(h - W_P)^2} \left(\frac{2T}{\pi^2} \right)^2 \Phi^2 \left(\frac{\lambda}{8}, x_3, t \right), \tag{35}$$

where ε_0 is the dielectric coefficient of air.

IDT Segment in Range ($\frac{\lambda}{4} < x_1 < \frac{\lambda}{2} \cup \frac{3\lambda}{4} < x_1 < \lambda$):

A slightly different approach is needed to evaluate the electrostatic force generated between the conductive plate and finger gaps. This is because of the space varying electric potential distribution mentioned above. Each finger gap is divided into N_s subdivisions in x_1 direction, so that each subdivision has a width of $\frac{f_w}{N_s}$ and a length of f_l ($\approx$ aperture of the IDT). Combining the relevant range in Equations 23 and 34, and after some algebraic simplifications, the electrostatic force generated by each gap can be evaluated as

$$F_{(gap)} = \frac{\varepsilon_0 f_l f_w}{2N_s(h-W_P)^2} \cdot \left[\sum_j \Phi^2\left(\frac{\lambda}{4} + \frac{j\lambda}{4N_s}, x_3, t\right) \right] \tag{36}$$

for j = 1, 2, 3,.., N_s. Therefore, for a distance of single wavelength (λ), the total electrostatic force generated is

$$F_{(\lambda)} = 2\left[F_{(+)} + F_{(gap)} \right]. \tag{37}$$

Furthermore, the above results can be used to extend the analysis to the evaluation of the resultant electrostatic force ($F_{(tot)}$) generated by an output IDT with N_p pairs of fingers. From Equations 35 – 37,

$$F_{(tot)} = \frac{C_t}{(h-W_P)^2} \sum_j \left[\left(\frac{2T}{\pi^2} \right)^2 \Phi^2\left(\frac{\lambda}{8}, x_3, t\right) + \Phi^2\left(\frac{\lambda}{4} + \frac{j\lambda}{4N_s}, x_3, t\right) \right] \tag{38}$$

for j = 1, 2, 3,.., N_s and $C_t = \frac{\varepsilon_0 f_l f_w N_p}{N_s}$.

As the doubly–clamped actuator is deflected due to the applied electrostatic force, an elastic restoring force is developed in the actuator. At equilibrium, the kinetic energy becomes zero, and actuator's potential energy reaches to a maximum. Therefore, to determine the displacement achieved by the actuator, the calculated electrostatic force and the elastic restoring force need to be considered at their equilibrium point (Washizu, 1975; Hu et al., 2004). However, this become a complex problem to solve since both the forces ($F_{(+)}$ and $F_{(gap)}$) depend on the actuator's instantaneous displacement $W_p(x_1)$. Therefore, to obtain an accurate solution for $W_P(x_1)$, analytical methods or numerical analysis methods such as FEM are required.

7. Finite element modelling of the actuator

For the Finite Element Analysis (FEA) of the actuator, a coupled–filed analysis is required since electrostatic and solid interactions are involved. Two distinct coupled–field methods can be identified in ANSYS; (i) Direct-coupling method, and (ii) Load transfer method (ANSYS Incorporation, 2009).

The direct–coupling method involves just one analysis that uses a coupled–field element type containing all necessary degrees of freedom. The coupling is handled by calculating element matrices or element load vectors that contain all necessary terms. Whereas the load transfer methods involve two or more analysis with each belonging to a different field, and two fields are coupled by applying results from one analysis as loads in another analysis.

There are different types of load transfer analysis in ANSYS; (i) ANSYS Multi–field Solver (MFS and MFX), (ii) Physics file based load transfer, and (iii) Unidirectional load transfer (ANSYS Incorporation, 2009). Suitability of these methods for a certain analysis depends on the physics fields involved, and whether the load transfer is unidirectional or not. Therefore, it is crucial to chose the most appropriate method to analyse a given scenario in order to achieve more accurate results in a reasonable simulating time. However, for MEMS applications ANSYS Multi–field solver is highly appropriate as it is a solver for sequentially coupled field analysis. Therefore in this research, ANSYS MFS is used for FEA of the SAW device based actuator.

7.1 Preparation of the model for analysis

The steps that were followed in the design and modelling of this device is as follows. Initially the geometry is created, and then element and material properties are defined for the actuator and the air–gap. As depicted in Figure 9, SOLID95 and SOLID122 element types are used for the structural and electrostatic models respectively. SOLID95 element has capabilities such as plasticity, creep, stress stiffening, large deflection, and large strain capability hence highly suitable for the design of microactuators. Whereas, SOLID122 is a 3D, 20–node, charge based electric element, which has one degree of freedom (Voltage) at each node. It is designed to tolerate irregular shapes without much loss of accuracy. Moreover, SOLID122 elements have compatible voltage shapes and are well suited to model curved boundaries and applicable to 3D electrostatic and time–harmonic, quasi–static electric field analysis (ANSYS Incorporation, 2009). In this modelling, the effect of the output IDT is designed by coupling a set of nodes at the bottom of the air–gap to match the desired IDT pattern and assigning a Volt Degree–of– Freedom (DoF) to those nodes.

Next, the geometry is meshed to a fine level to accommodate for accurate micro level changes in the structure. Once the geometry is meshed, relevant electric and mechanical boundary conditions are applied. After setting the boundary conditions and constrains, a static analysis is carried out mainly to check for the convergence criteria. Once the results are converged in static analysis, then a model analysis is carried out to extract the natural frequencies of the conductive actuator. As a result, the operating mode for the actuator can be realised, and then a transient analysis is performed for a long enough time period that is dictated by the natural frequency mode of the actuator and the frequency of operation of the SAW device. This is an important step in the modelling process as it helps to decide on an optimal completion time for the transient analysis, since the transient simulations generally take a longer time to complete.

To simplify the analysis, the performance of the thin conductive plate with a smaller width was initially considered. Additionally, half–symmetry is exploited due to the symmetrical nature of the model. As a result, a reduced number of nodes and elements were generated for the model, and hence reduced simulation times and improved CPU usage were achieved.

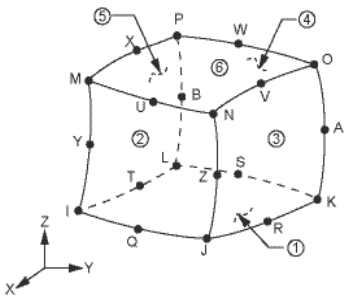


Fig. 9. SOLID95 and SOLID122 element geometries. 3D, 20–node elements used in the design of actuator and the air–gap (ANSYS Incorporation, 2009). SOLID95 element has capabilities such as plasticity, creep, stress stiffening, large deflection, and large strain capability. SOLID122 is a charge based electric element with one degree of freedom (Voltage) at each node. SOLID122 elements are well suited to model curved boundaries and applicable to 3D electrostatic and time–harmonic quasi–static electric field analysis (ANSYS Incorporation, 2009).

8. Simulations and results

8.1 Static analysis

Initially, the static analysis was carried out to determine the static displacement of the actuator. In order to mimic the effect of the electric potential wave generated at the output IDT of the SAW device, a set of interleaved electrodes were used and every alternative electrode was coupled, so that one set of electrodes act as the positive bus bar and the other as the negative bus bar. Hence, in the microactuator modelling, the whole SAW device was replaced at simulation level. Material properties of silicon were used for the doubly-clamped conductive plate, which in turn acts as a microactuator. The conductive plate dimensions were chosen to be 1000 μm × 2 μm × 10 μm (L×H×W). The gap between the electrodes and the conductive plate h was taken to be 10 μm and was considered to be filled with air. For static analysis, a 10 Volt input voltage was applied to the positive bus bar. The negative bus bar and the conductive plate was connected to a common ground to form the electrostatic field.

Initial FEA results are verified using a commonly used Rayleigh–Ritz method based analytical model. For comparison purposes, displacement versus voltage results were plotted and are shown in Figure 10. A good correlation can be observed between the analytical and simulation results for the microactuator. However, FEA results demonstrate slightly lower displacements for a given voltage. This is mainly because the full thickness of the actuator was considered in the simulated 3D model in FEA, whereas the actuator was modeled as a thin plate in the Rayleigh–Ritz method based analytical model. Therefore, the higher bending stiffness reduces the effective mid–beam displacement in the FEA model. It should be noted that the actuator displacement can be increased by reducing the gap between the conductive plate and the output IDT, reducing the thickness of the conductive plate, and reducing the stress level applied at the actuator by optimising the clamping mechanism.

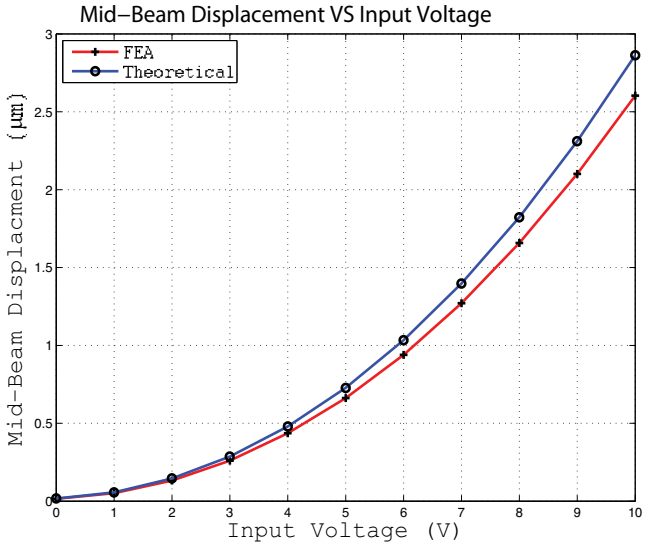
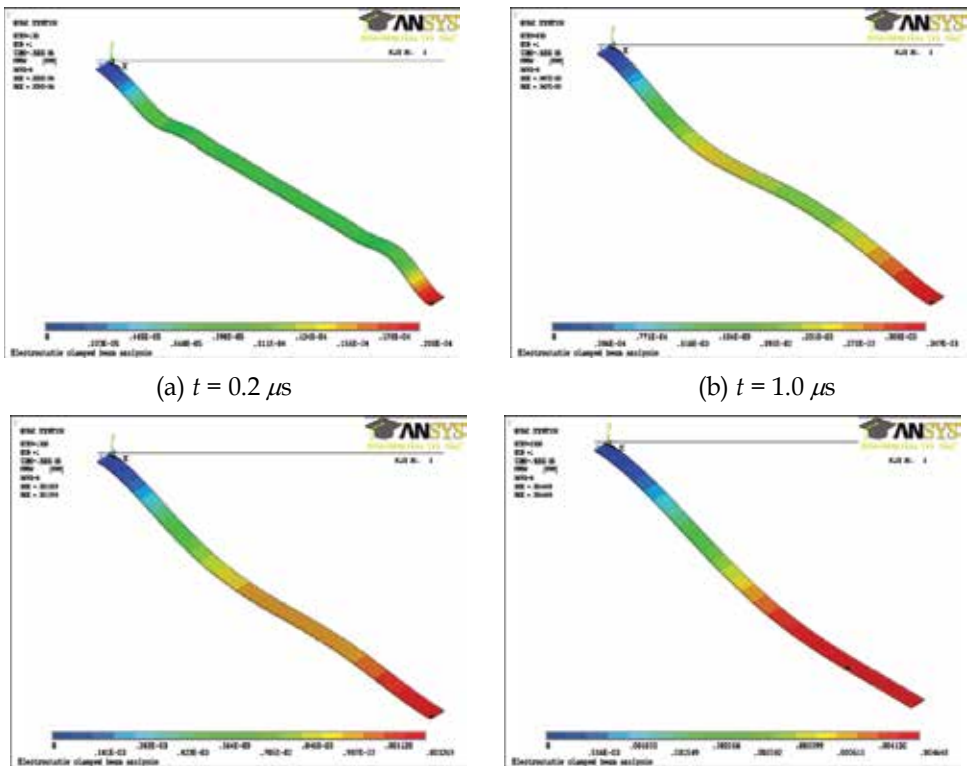


Fig. 10. Simulation and theoretical results. Comparison of simulated and theoretical results for the SAW actuator. Displacement VS Voltage plot for the mid-beam displacement in the conductive plate actuator above the SAW device.

Once the static analysis was completed more detailed transient analyses were performed in ANSYS to investigate the dynamic behavior of the actuator.

8.2 Transient analysis

It should be noted that when a conductive beam is subject to a dynamically changing electrostatic field, the displacement behaviour needs to be calculated analytically or numerically; using advanced simulation tools equipped with in built algorithms, such as ANSYS. This section presents the transient simulation results carried out for the conductive plate with the same dimensions mentioned in the static analysis above. Moreover, an AC sinusoidal wave with a frequency of 50 MHz and a peak voltage of 10 volts were used to emulate the electric potential wave at the output IDT as proven in Equation 23. The conductive plate is connected to ground so that the plate acts as an equipotential surface. However, the node density of the model, and the CPU processing power were found to be major constrains that restricted longer transient analysis (ex: 1000×T, where T is the period of SAW). Moreover, a higher node density was needed to effectively represent the output IDT in FEA model. By considering these factors, transient simulations were performed for 400×T during this analysis.



(a) t = 0.2 μs (b) t = 1.0 μs

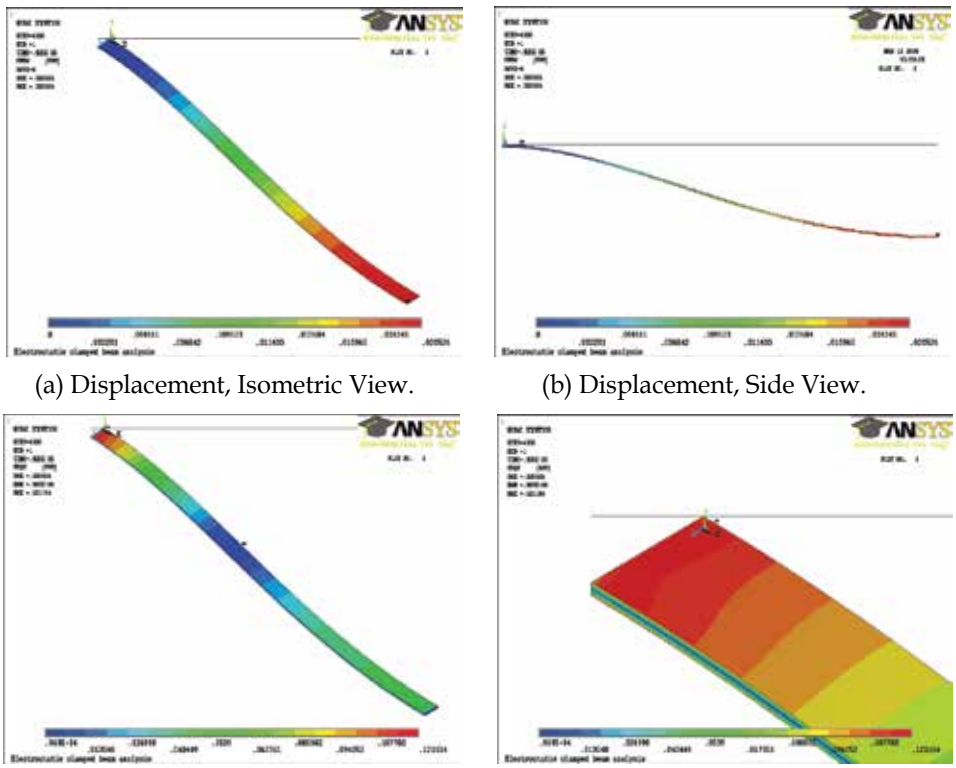
(c) t = 2.0 μs (d) t = 4.0 μs

Fig. 11. Transient analysis results for intermediate steps. Deflection results for the actuator performance at various time steps during the transient analysis. Half–symmetry is exploited due to the symmetrical nature of the model. The flexural behaviour is observed during stabilisation period.

Figures 11 – 12 depict the actuator displacements for different steps in transient analysis. As a thinner actuator is modelled in ANSYS, the flexural behavior of the actuator is first observed. As the time progresses, the deflection profile of the actuator is found to be similar to the profile obtained from the Rayleigh–Ritz method based analysis.

Figures 12 (c) and (d) depict the contour plot of the Von Mises stress distribution of the actuator. Here, Von Mises stress can be used to predict the yielding of any of the materials used, under any loading condition. The maximum Von Mises stress in this scenario is 0.121 MPa, which is much lower than the yield strengths of the selected material. This demonstrates that the actuator's deflection is well within the elastic range of the materials used.

As can be seen from these simulations, micro displacements are successfully obtained using SAW based actuation method. Figure 13 shows the mid–beam and the quarter–beam displacement variations over a simulation time of 400×T. Based on the static analysis however, it was shown that displacements up to ~3 μm can be achieved using SAW device



(a) Displacement, Isometric View. (b) Displacement, Side View.

(c) Von Mises stress, Isometric View. (d) Von Mises stress, Clamped edge.

Fig. 12. Transient analysis results for final step. Deflection and Von Mises stress analysis results for the actuator performance at t = 8.0 μs. Half–symmetry is exploited due to the symmetrical nature of the model. The maximum Von Mises stress in this scenario is 0.121 MPa, which is near the clamped edge. This is much lower than the yield strengths of the selected material, hence demonstrating that the actuator's deflection is well within the elastic range.

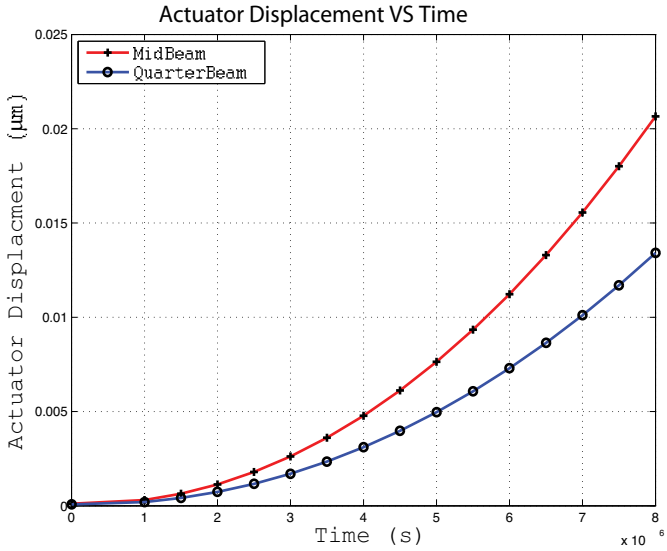


Fig. 13. Displacement VS Time plot of the mid–beam. Analysis carried out for 400×T, where T is the time period of the SAW signal. As the time increases the mid–beam displacement as well as the quarter–beam deflection increase at an increasing rate.

based actuation. As a result, it is proven that even after 400×T, still the dynamic displacement does not show any periodic nature but in the process of gaining more displacement. Based on these results, it is evident that the actual operating frequency of the conductive plate during actuation is a very much a scaled down version of the SAW frequency.

9. Conclusion

In this chapter, the use of a SAW device to generate microactuations was demonstrated. Detailed theoretical analysis explaining how the entire SAW device based actuator operation was carried out and boundary conditions applicable for presented design was used to derive the electric potential wave forms, hence the electrostatic field between the SAW device and the conductive plate. Displacement analysis of the conductive actuator was obtained. Static analysis results were generated using the ANSYS simulation tool, and compared with the theoretical results obtained by Rayleigh–Ritz method. A good correlation between the theoretical and simulated displacement curves were observed.

Once the static analysis was completed, the dynamic behaviour of the SAW device based electrostatic actuator was studied using transient analysis. This is more substantial in investigating the operating frequency of the conductive plate. Since the SAW frequency is in the range between 50 MHz–1 GHz it was crucial to verify the effective operating frequency of the conductive plate. Because of the time varying electrostatic field, it was found that the oscillating frequency of the actuator is much less than that of the SAW frequency. Therefore, the applicability of this SAW based secure and wireless interrogation for implantable MEMS devices is clearly demonstrated.

10. References

Adler, E. L. (2000). Bulk and surface acoustic waves in anisotropic solids, *International Journal of High Speed Electronics and Systems* 10(3): 653–684.

ANSYS Incorporation (2009). ANSYS Help Guide–V.11. *http://www.kxcad.net/ansys/ANSYS/ansyshelp/ index.htm* (visited on 25/05/2010).

Dissanayake, D.W., Tikka, A. C., Al-Sarawi, S. &Abbott, D. (2007). Radio frequency controlled microvalve for biomedical applications, *Proc. of SPIE–Smart Materials IV* 6413: Article 64130D: 1–13.

Dvoesherstov, M. Y. & Chirimanov, A. P. (1999). Numerical analysis of a surface and leaky surface acoustic wave in new piezoelectric $KNbO_3$, PKN, and LGN crystals, *Radiophysics and Quantum Electronics* 42(5): 431–438.

Dvoesherstov, M. Y., Petrov, S. G., Cherednik, V. I., & Chirimanov, A. P. (2000). Transformation of modes of surface acoustic waves in strong $KNbO_3$ and PKN piezoelectric crystals, *Radiophysics and Quantum Electronics* 43(5): 400–406.

Gantner, A., Hoppe, R. H.W., Köster, D., Siebert, K. G. & Wixforth, A. (2007). Numerical simulation of piezoelectrically agitated surface acoustic waves on microfluidic biochips, *Computing and Visualization in Science* 10(3): 145–161.

Gardner, J. W., Varadan, V. K. & Awadelkarim, O. O. (2001). *Microsensors, MEMS, and Smart Devices*, First edn, Tsinghua University Press, Beijing, chapter : Microsensors, Introduction to SAW devices, Surface acoustic waves in solids, IDT microsensor parameter measurement, IDT microsensor fabrication, IDT microsensors, pp. 227–396.

Horenstein, M. N., Perreault, J. A. & Bifano, T. G. (2000). Differential capacitive position sensor for planar MEMS structures with vertical motion, *Sensors and Actuators* 80: 53–61.

Hu, Y. C., Chang, C. M. & Huang, S. C. (2004). Some design considerations on the electrostatically actuated microstructures, *Sensors and Actuators A* 112: 155–161.

Ippolito, S. J., Kalantar-zadeh, K., Wlodarski, W. & Powell, D. A. (2002). Finite-element analysis for simulation of layered SAWdevices with XY LiNbO3 substrate, *Proc. of SPIE– Smart Structures, Devices, and Systems* 4935: 120–131.

Jones, I., Ricciardi, L., Hall, L., Hansen, H., Varadan, V., Bertram, C., Maddocks, S., Enderling, S., Saint, D., Al-Sarawi, S. & Abbott, D. (2008). Wireless RF communication in biomedical applications, *Smart Materials and Structures* 17: 015050: 1–10.

Kannan, T. (2006). *Finite element analysis of surface acoustic wave resonators*, Master's thesis, University of Saskatchewan.

Maugin, G. A. (1985). *Nonlinear electromachanical effects and applications*, First edn,World Sceintific Publishing Co. Pte. Ltd., chapter : Rayleigh Surface Waves, pp. 104–142.

Milstein, L. B. & Das, P. (1979). Surface AcousticWave Devices, *IEEE Communications Magazine* 17: 25–33.

Ruppel, C. C. W., Reindl, L. & Weigel, R. (2002). SAW devices and their wireless communication applications, *IEEE Microwave Magazine, ISSN 1527-3342* 3(2): 65–71.

Skinner, J. L., Cardinale, G. F., Talin, A. A. & Brocato, R.W. (2006). Effect of critical dimension variation on SAWcorrelator energy, *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control* 53(2): 497–501.

Strobl, C. J., Guttenberg, Z. V. & Wixforth, A. (2004). A Nano–and pico–dispensing of fluids on planar substrates using SAW, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 51(11): 1432–1436.

Subramanian, H., Varadan, V. K., Varadan, V. V. & Vellekoopz, M. J. (1997). Design and fabrication of wireless remotely readable MEMS based microaccelerometers, *Smart Materials and Structures* 6(6): 730–738.

Tsai, N. C. & Sue, C. Y. (2007). Review of MEMS–based drug delivery and dosing systems, *Sensors and Actuators A* 134: 555–564.

Upadhyay, S. K. (2004). *Seismic Reflection Processing: With Special Reference to Anisotropy*, First edn, Springer–Verlag, Berlin, chapter: Anisotropy Models of Sedimentary Sections and Characteristics of Wave Propagation, pp. 143–201.

Varadan, V. K. & Varadan, V. V. (2000). Microsensors, micromechanical systems (MEMS), and electronics for smart structures and systems, *Smart Materials and Structures* 9: 953–972.

Čiplys, D. & Rimeika, R. (1999). Measurements of electromechanical coupling coefficient for surface acoustic waves in proton-exchanged lithium niobate, *ULTRAGARSAS journal* 33(3): 14–20.

Washizu, K. (1975). *Variational methods in elasticity and plasticity*, Second edn, Pergamon Press Ltd., Oxford, chapter : Beams, Plates, pp. 132–182.

Wixforth, A. (2003). Acoustically driven planar microfluidics, *Superlattices and Microstructures* 6: 389–396.

Wolfram MathWorld (2009). Euler angles. *http://mathworld.wolfram.com/EulerAngles.html* (visited on 25/05/2010).

Zaglmayr, S., Sch¨berl, J. & Langer, U. (2005). *Progress in Industrial Mathematics at ECMI 2004*, Vol. 8 of *Mathematics in Industry*, Springer–Verlag, Berlin, chapter: Eigenvalue Problems in Surface Acoustic Wave Filter Simulations, pp. 75–99.

10

Surface Acoustic Wave Motors and Actuators: Mechanism, Structure, Characteristic and Application

Shu-yi Zhang and Li-ping Cheng
Lab of Modern Acoustics, Institute of Acoustics, Nanjing University
Nanjing 210093,
China

1. Introduction

Ultrasonic motors, as one kind of actuators, have attracted a lot of attention since it was proposed more than 20 years ago. In such kind of motors, the sliders (for linear motors) or the rotors (for rotary motors) are driven by the frictional forces between the sliders (rotors) and the stators when ultrasonic waves are propagating on the stators. Since then, the ultrasonic motors have been developed and applied successfully in wide fields, such as mechanical, optic, electronic, and automatic, as well as aeronautic and astronautic industries and technologies because of their unique advantages over conventional electro-magnetic ones, such as high driving forces and torques, easy controllability, quiet operation, non-electromagnetic induction, etc. (Sashida & Kenjo, 1993; Ueha & Tomikawa, 1993). Besides, with the rapid development of micro-electro-mechanical system (MEMS), miniature ultrasonic motors were developed (Dong et al., 2000; Zhang et al., 2006). However, the direct contact between the sliders (rotors) and the stators restricts the velocity and working lifetime of the motors, then a new kind of non-contact motors were presented, where a fluid is introduced between the stator and slider. Thus, instead of the frictional force, acoustic streaming excited by the acoustic wave on the stator and propagating in the fluid is used for driving the slider or rotor to move (Nakamura et al., 1990; Yamayoshi & Hirose, 1992; Hu et al., 1995; Cheng et al., 2007).

On the basis of conventional ultrasonic motors, several studies on new types of ultrasonic motors (actuators) with the driving forces coming from surface acoustic waves (SAWs) were presented (Moroney et al., 1989; Kurosawa et al., 1994). For the SAW motors, the SAWs are excited by interdigital transducers (IDTs) deposited on surfaces of piezoelectric substrates or thin films, and the SAW energies are concentrated in the thin layers near the surfaces of the substrates (for Rayleigh waves) or in the thin films (for Lamb waves). In addition to the characteristics of conventional ultrasonic motors, the SAW motors have more advantages, such as the high operation frequency, high speed, high energy density around the surfaces, and higher output force/torque, etc. Meanwhile, since the SAWs are excited by IDTs, which can be fabricated with planar technologies of semiconductor industries, the new types of motors are suitable for miniaturizing and integrating with integrated circuits and MEMS devices, etc.

To overcome the difficulties of the frictional drive and extend the applications of the motors, several kinds of non-contact SAW linear motors (actuators) were developed (Sano, et al., 1997), in which a fluid layer (or a drop) is introduced between the stator and slider (rotor) of the actuator. Then a SAW streaming excited by the IDT and propagating in the fluid covered on the surface of the stator, instead of the frictional force, is used to drive the slider (rotor), by which the required driving power of the actuators is reduced greatly and the lifetime can be extended (Shiokawa, et al. 1990; Takeuchi, et al., 1994; Gu, et al., 2008). The non-contact SAW actuators have been widely used in chemical and biochemical fields (Takeuchi et al., 2005).

In this chapter, the structures and characteristics of IDTs for exciting SAWs and the excited SAW modes on different substrates are introduced briefly. Then the structures of the stators and sliders (rotors), theories and characteristics of the conventional contact linear and rotary SAW motors are presented. In addition, the mechanisms, structures and characteristics of non-contact SAW actuators, as well as some applications of the motors (actuarors), are also described and discussed.

2. Generation and propagation mode of SAWs

2.1 Structure and characteristic of interdigital transducers

SAWs can be generated by many different types of transducers. Up to now, a most popular and effective type of the transducers is the interdigital transducer (IDT), which consists of two interlocking comb-shaped metallic electrode arrays. For the simplest structure, the metallic electrodes have the same length (aperture) and the same width $\lambda/4$ as that of the gap, where λ is the SAW wavelength, as shown in Fig.1(a). The IDT is deposited on a piezoelectric substrate by the photolithographic technology. When a RF voltage with the same frequency as that of the IDT is applied to the IDT, the electric field components change sign from gap to gap, so that a corresponding periodic mechanical strain field is produced through the piezoelectric effect of the substrate.

The IDT radiates acoustic waves in both forward and backward directions, but unidirectional radiation can be obtained with special interdigital arrays. The simplest one is to use two identical interdigital transducers separated by a distance $(n+1/4)\lambda$, where n is an integer; both transducers are driven from two generators having 90 degree phase difference between them, or by a single generator with a quarter-wavelength of electrical transmission line connecting both transducers. As a result, the generated waves traveling to the right from each transducer add up, while those traveling to the left cancel from each other. The unidirectionality increases the conversion efficiency of the transducer by 3 dB since waves radiate in only one direction instead of two directions, and the bandwidth is reduced by this operation. In addition, for undirectional transducers, the waves incident to the left transducer from the right are not as strongly reflected as from a bidirectional array (White, 1970).

For the substrates with a weak piezoelectric effect, if the nonlinear effect is neglected, the SAW vibration amplitude is approximately proportional to the electrode number N, but the bandwidth is inversely proportional to N of the IDT. Meanwhile, in order to obtain the SAW field with appropriate homogeneity, the length of the electrodes (aperture) of the IDTs should also be suitably enlarged if the size of the IDT has no limit.

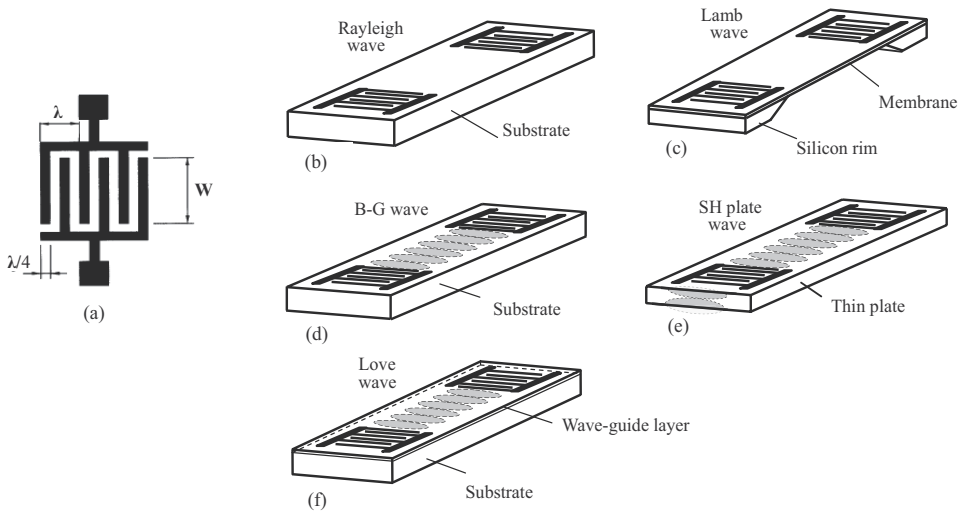


Fig. 1. Schematic diagram of transducer and SAW modes on different substrates: (a) interdigital transducer; (b) Rayleigh wave; (c) Lamb wave; (d) Bleustein-Gulyaev (B-G) wave; (e) shear horizontal plate wave; (f) Love wave

2.2 Mode and characteristic of SAWs

For different piezoelectric substrates, the IDT may excite waves with different modes, which depend on the materials and cut directions of the piezoelectric crystals or polarization directions of ceramics, as well as on the piezoelectric thin films with special growth directions on substrates. Generally, the wave modes are classified in five types as shown in Figs. 1(a)-1(f).

(i) In the Rayleigh wave mode shown in Fig.1(b), the surface particles in the sagittal plane of the substrate move in a retrograde elliptical trajectory relative to the SAW propagation direction, as shown in Fig.2(a). Besides, the amplitude of the Rayleigh mode decreases almost exponentially with the depth in the substrate, and the penetration depth of the wave is considered to be one wavelength range as shown in Fig.2(b). Therefore, the acoustic energy is concentrated in a thin layer beneath the surface with the depth about one wavelength of the Rayleigh wave. (ii) The Lamb wave propagates in a thin plate shown in Fig.1(c), so it is also called as the plate wave. There are two kinds of modes for Lamb wave, i.e., symmetric and anti-symmetric modes as shown in Fig.2(c), which may be considered as the composition of two Rayleigh waves propagating on both boundaries of a plate as the thickness of the plate is just over one wavelength. The symmetric and anti-symmetric modes of Lamb wave can be obtained by the composition of both Rayleigh modes with opposite phases and the same phases, respectively. (iii) For the Bleustein-Gulyaev (B-G) wave shown in Fig.1(d), it is a horizontally polarized surface wave propagating on an infinite piezoelectric substrate. (iv) For shear horizontal plate wave (SH plate wave) (see Fig.1(e)), the thickness of the substrate (thin plate) is half of the wavelength, and (v) for the Love wave (Fig.1(f)), the SH wave propagates in a thin layer covered on the substrate (White, 1970; Auld, 1973).

Therefore, for the last three kinds of the SAW modes, the vibrations of the surface particles are perpendicular to the wave propagation direction, but parallel to the surface of the

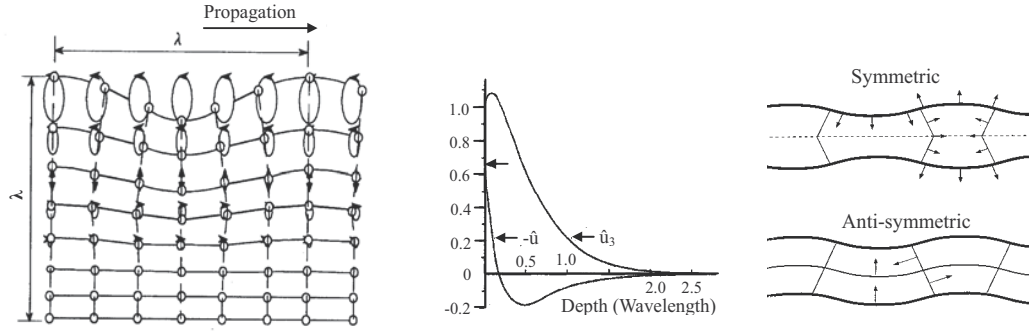


Fig. 2. Characteristics of SAWs propagating in elastic isotropic medium; (a) particle motion orbit of Rayleigh wave; (b) particle displacement of Rayleigh wave; (c) Lamb wave: symmetric and anti-symmetric modes.

substrate. For ultrasonic motors, the particle displacement of the surface is required to have a component perpendicular to the surface of the substrate, so, up to now, only Rayleigh and Lamb modes are used as the driving sources of the SAW motors.

3. Conventional SAW motors

Since a kind of ultrasonic micro-motors driven by Lamb waves with high frequencies excited by IDT was reported in 1989 (Moroney et al., 1989), several kinds of SAW motors driven by Rayleigh waves excited by IDTs have been developed. The first prototype of SAW linear motors was presented by Kurosawa et al., in which two pairs of IDTs with the central frequency about 10 MHz were prepared perpendicularly on a piezoelectric substrate, then two Rayleigh waves were excited in cross directions and a two-dimensional SAW motor was built-up (Kurosawa et al., 1994; 1996). Generally, for the conventional SAW linear motors, the sliders in contact with the stators are directly driven by the frictional forces between the sliders and stators. Based on the SAW linear motors, a kind of SAW rotary motor operated in similar conditions was also developed (Zhang et al., 2000). These SAW motors have similar operation principles, characteristics, and theories, which are described separately in this section.

3.1 Principle of SAW linear motors

SAW motors are composed of stators and sliders, where the stators are SAW devices (such as delay lines). A typical structure of SAW linear motors is shown in Fig.3(a) (Asai et al., 1999). The slider is in directly contact with the stator and driven by the frictional force between the slider and the stator induced by the SAW propagating in the stator. The acoustic wave mode used in the most of the conventional SAW motors is Rayleigh wave excited by IDT deposited on piezoelectric substrate. The driving force applied on the slider is induced by the particle motions and in the direction opposite to the SAW propagation. Since the amplitude of the Rayleigh wave decreases exponentially with the depth in the substrate, the acoustic energy is concentrated in a thin layer beneath the surface with a thickness of about one wavelength of the SAW. Therefore, the energy density is very high, which is beneficial for improving the utilization efficiency of the acoustic energy.

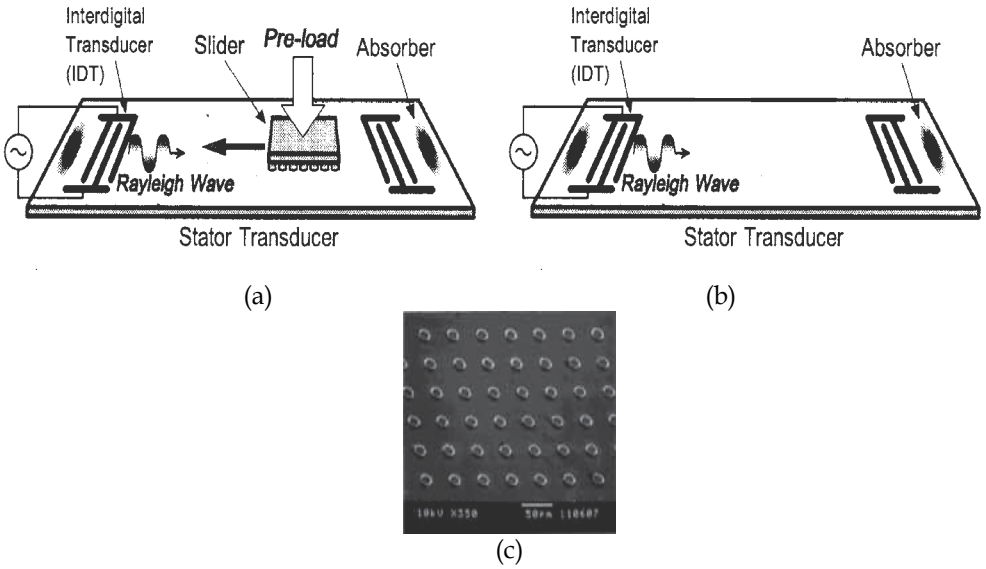


Fig. 3. SAW linear motor: (a) structure of motor; (b) stator with IDTs; (c) slider with projections on a Si wafer.

However, when the acoustic amplitude is large enough, the wave propagation may become nonlinear, such as generation of harmonic frequencies and frequency mixing. These effects may act as unwanted sources of wave attenuation. Therefore, in the practical applications of SAW motors, the operation conditions must be considered in compromise and optimization.

3.2 Structure of SAW linear motors

(a) SAW stators

For SAW motors, the ordinary SAW mode is Rayleigh waves, in which the substrates of 128^0 Y-cut X-propagation $LiNbO_3$ (128^0 YX-$LiNbO_3$) crystals are always used as the stators since the $LiNbO_3$ crystal substrates have a high electro-mechanical coupling coefficient, and the cut direction of $LiNbO_3$ crystals is the propagation direction of pure Rayleigh mode. Generally, one pair of IDTs are fabricated by the photolithographic technology on substrates as shown in Fig.3(b). One of the IDTs is applied by a RF voltage with the frequency consistent with the central frequency of the IDT, thus the SAW in Rayleigh mode is excited and propagating in two directions on the surface of the $LiNbO_3$ substrate. The other IDT can be used to receive the SAW for checking the wave propagation. In addition, some soft materials (absorbers) are applied on the areas between the IDTs and boundaries of the substrate to absorb the superfluous SAWs for eliminating the reflections of the boundaries. In the SAW motor studies, the SAW frequency is generally taken in the range of 1-100 MHz. Considering the vibration amplitude of the SAW is approximately proportional to the electrode number of the IDT, in order to increase the SAW energy, the electrode number should be large enough, such as more than 10 pairs. Meanwhile, to make the SAW field more homogeneous, the aperture of the IDTs should also be large, such as more than 20 wavelengths, thus the sizes of the IDT and the stator will be much larger. Therefore, to miniaturize the SAW motors, the frequency of the SAW should be increased to decrease the

size of the IDTs. However, the vibration amplitude should also be decreased because the SAW amplitude is approximately inversely proportional to the frequency of the IDTs. As the motors are fabricated to operate at very high frequencies, the vibration amplitudes of the SAWs are very small, so it is required that the surface of the stators should be very smooth, i.e., with very fine roughness (Takasaki et al., 1998; Cheng et al., 2002).
On the other hand, the Lamb wave is another kind of wave modes used in SAW motors, which is excited in thin piezoelectric plates (films) as shown in Fig.2(c) and is suitable for manufacturing micro-motors (micro-actuators) used in micro-electro-mechanical systems.

(b) Sliders

Generally, the sliders in SAW linear motors could be thin plates or small balls fabricated by various materials, such as silicon wafers or aluminum sheets, steel balls and/or ruby balls, etc. In order to control the contact pressure and contact area between the sliders and stators, spherical-shaped sliders may be preferably adopted. To increase the friction-driving force, it is better to manufacture the contact area of the slider with a multi-sphere shape, such as an array of small bumps at the contact surface of a silicon wafer slider, as shown in Fig. 3(c), especially for the motors operating in higher frequencies (Takasaki et al., 1998; 2000).

3.3 Characteristic and performance of SAW linear motors

The moving velocity and the output force of SAW linear motors driven by the frictional forces are dependent upon the driving voltage of the IDTs and the contact pressure between the stators and sliders. To get suitable velocity and output force of the motors, the contact pressure must be controlled by applying preload, such as applying leaf springs or magnets. For example, for a SAW motor with the frequency about 10 MHz, a Si wafer with projection array shown in Fig.3(c) was used as a slider under a leaf spring preload of about 30 N, the transient responses of the slider motion under different driving voltages were measured by a laser vibrometer as shown in Fig.4(a) (Kurosawa, 2000). Sequentially, a miniaturized SAW motor with the frequency of about 50 MHz was presented, in which the Si wafer was used as a slider and a magnetic force was used to control the preload. The moving speeds of the silicon slider under different driving voltages were measured as shown in Fig.4(b) (Takasaki et al., 1998), in which the maximum output force was calculated as 0.036 N, that is 28% of the preload,. To investigate the effect of the slider material on the transient response of the motor, three kinds of materials were used as the sliders and the results were shown as Fig.4(c) (Kurosawa et al., 1994).
Up to now, the SAW motors driven with much higher frequencies have been fabricated. For example, a motor operating at about 100 MHz has been accomplished, therefore the size of the stator was greatly reduced to $3\times12.5\times0.5$ mm^3. The results showed that the motor had a high speed of 0.3 m/s and a high output force of 13 mN (Shigematsu & Kurosawa, 2006).
To increase the efficiency of SAW motors, two kinds of power circulation methods were developed (Asai et al., 1999). The first power circulation method is shown in Fig.5(a), in which two driving IDTs and two unidirectional IDTs are required. The excited traveling wave is received by one unidirectional IDT and converted into electric energy. Another unidirectional IDT excites a circulated traveling wave using the electric energy. Each unidirectional IDT is located at a suitable position, then the excited wave and circulated wave can add up with each other. The second method is shown in Fig.5(b), in which two

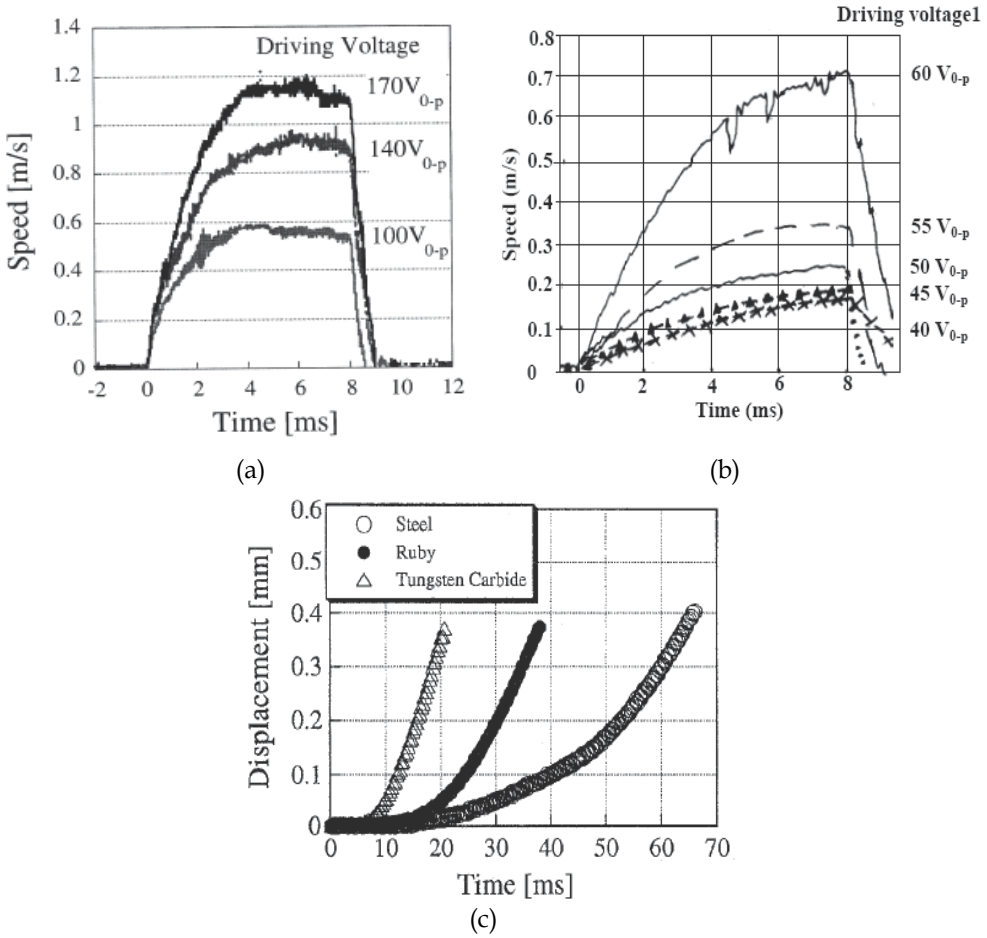


Fig. 4. Transient response of SAW motor: (a) at about 10 MHz; (b) at 50 MHz; (c) at about 10 MHz.

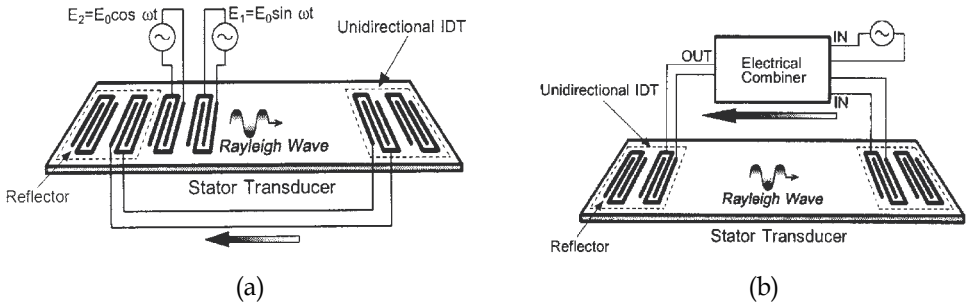


Fig. 5. Structure of stator using two power circulation methods

unidirectional IDTs and an electrical combiner are required. The output port of the combiner is connected to the exciting unidirectional IDT and the input ports are connected to the other IDT and a RF electric power source. The excited SAW is received by the other IDT and converted to electric energy, and further put into the combiner, which makes the energy to be recycled.
However, on the other hand, the abrasion is one problem that influences the lifetime of the motor driven by frictional force. To avoid the problem, segment-structured diamond-like carbon film coatings were proposed to be applied for the friction material (Fujii et al., 2007). The segment-structured film coatings could be deposited on the sliders and/or on the stators, which are available as the projections on the silicon sliders.

3.4 SAW rotary motors

On the basis of the SAW linear motors, SAW rotary motors were developed, in which two pairs of IDTs were deposited in parallel on the surface of 128^0 YX-$LiNbO_3$ substrate as the stator shown in Fig.6(a) (Zhang et al., 2000). In the stator, while IDT-1 and IDT-3 are applied a RF voltage with the central frequency of the IDTs, two SAWs with anti-parallel propagation directions are excited. As a rotor is located at the center of the stator and spans on both SAW waves, the anti-parallel propagations of both SAWs produce one pair of oppositely directed frictional forces on the rotor, which result in a torque driving the rotor to rotate in anti-clock-wise. It is easy to invert the rotation direction by using the other two IDTs (2 and 4) to excite other pair of anti-parallel SAW waves.
For the rotary motors initially built-up, the prototype using Rayleigh waves is shown in Fig.6(b), the IDTs are composed of 20 pairs interdigital electrodes with the aperture of 25 mm and the operation frequency of about 10 MHz. The rotors are circular disks, in which a series of small holes distributed axisymmetrically near the fringe, and the holes are filled with small balls as shown in Fig.6(c). The disks and balls can be made by different materials, such as plexiglass disks and steel balls, etc. The motor speeds were measured by a digital video camera and analyzed by an image processing program. As the IDTs are supplied with the driving voltage of 100 $V_{p\text{-}p}$ about 2 seconds, the angular displacement and rotation speed of the rotary motor gradually increase to about 2000 degrees (5.5 circles) and 20 rad/sec, respectively. From the transient response, the maximum torque and torque-speed relationships are estimated using the method proposed by Nakamura et al. (1991), by which it is found that the maximum output torque is 3.3 Ncm, and the steady rotation speed is 180 rpm at the driving voltage of 100 $V_{p\text{-}p}$.

4. Theoretical simulation of SAW motors

4.1 Theory and numerical simulation of linear SAW motors

The first attempt to describe the energy transfer from the acoustic wave to the slider in the linear SAW motors, perhaps, was presented by the mechanism that the slider is in alternative phase of levitation and contact with the stator, and then the slider motion is in sequential step-like behavior (Helin et al., 1998). Based on the mechanism, as the SAW motor driven by Rayleigh wave has a spherical slider, the theoretical model can be simplified as shown in Fig.7(a), in which the displacement (x_{ellip} ,y_{ellip}) of the surface particle of the stator can be expressed as (Morita et al., 1999)

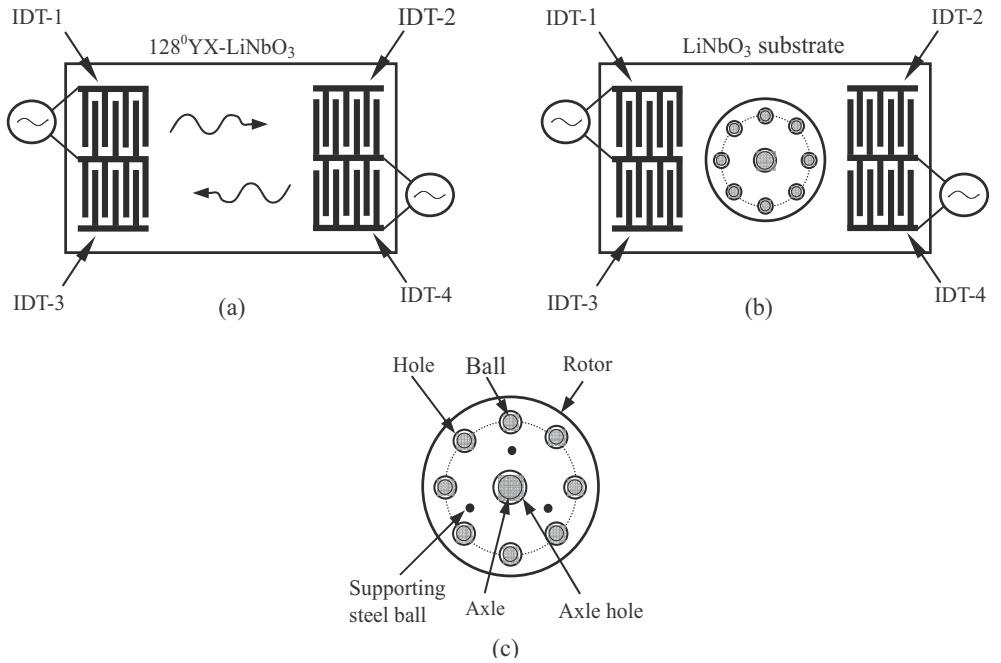


Fig. 6. Schematic diagram of the SAW rotary motor. (a) stator; (b) SAW rotary motor; (c) rotor

$$\begin{cases} x_{\text{ellip}} = A\alpha\cos(\omega t + kx_{\text{slider}}), \\ y_{\text{ellip}} = A\sin(\omega t + kx_{\text{slider}}), \end{cases} \tag{1}$$

where A is the amplitude, ω is the angular frequency, k is the wave number, α is the ratio of the tangential amplitude to the normal amplitude of the vibration, and x_{slider} and y_{slider} are the coordinate positions of the slider. During the operation, the slider experiences two situations: one is in levitation and the other is in contact with the stator as shown in Fig.7(b) and (c) respectively.

If $y_{ellip} < y_{slider}$, the slider is in levitation as shown in Fig.7(b), then the motion equations are

$$\begin{cases} Md^2y_{\text{slider}} / dt^2 = -P - Mg, \\ Md^2x_{\text{slider}} / dt^2 = 0, \end{cases} \tag{2}$$

where P is the preload to the slider, M is the mass of the slider, and g is the gravity.

If $y_{ellip} > y_{slider}$, the slider is in contact with the stator as shown in Fig.7(c), so it obtains a frictional force in the tangential direction. The motion equations are written as

$$\begin{cases} Md^2y_{\text{slider}} / dt^2 = N - P - Mg, \\ Md^2x_{\text{slider}} / dt^2 = Fv_{\text{slider}} / \left|v_{\text{slider}}\right| = \pm\mu N v_{\text{slider}} / \left|v_{\text{slider}}\right|, \end{cases} \tag{3}$$

where N is the normal supporting force in the normal direction, μ is the frictional coefficient and v_{slider} is the speed of the slider. The contact force is calculated using the Herzian contact theorem as follows (Timoshenko et al., 1970)

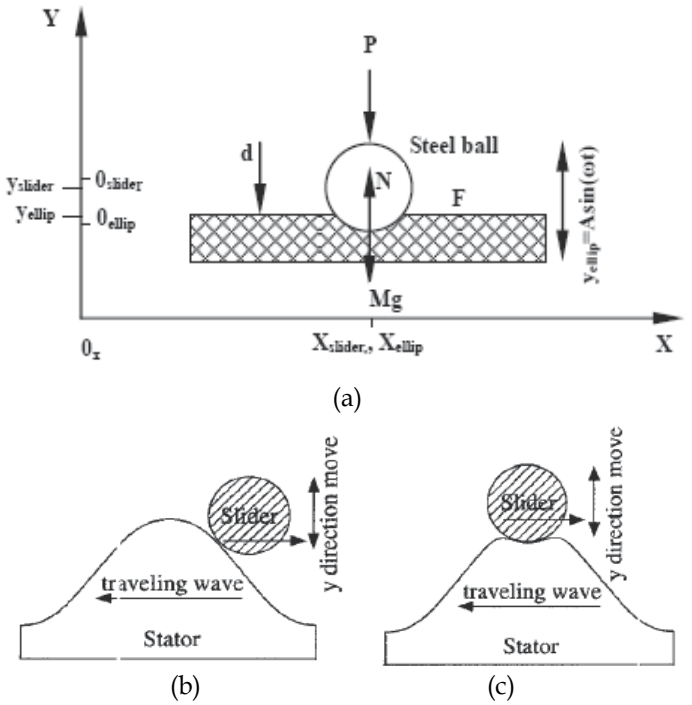


Fig. 7. Theoretical model of SAW linear motor: (a) position of slider; (b) levitation; (c) contact

$$N = \frac{4ER^{\frac{1}{2}}d^{\frac{1}{2}}}{3}, \qquad E = \left(\frac{1-\sigma_1^2}{E_1} + \frac{1-\sigma_2^2}{E_2} \right)^{-1}, \tag{4}$$

where R is the radius of the rotor, E_i and σ_i (i =1,2 represent the slider and the stator, respectively) are the Young's module and Poisson's ratio, respectively.
To calculate the displacement and speed of the slider, the Euler method is used as

$$\begin{cases} x_{slider}(t_n) = x_{slider}(t_{n-1}) + \dfrac{dx_{slider}}{dt}(t_{n-1})\Delta t, \\ v_{slider}(t_n) = v_{slider}(t_{n-1}) + \dfrac{dv_{slider}}{dt}(t_{n-1})\Delta t. \end{cases} \tag{5}$$

The simulation results for the motor driven by a pulse-modulated force with the frequency of 9.6 MHz are shown in Fig.8, where the left (Fig.8(a)) depicts the driving force, while the middle and right show the moving speed and the displacement respectively (Morita et al., 1999). The results are in agreement with those of the experiments.
Recently, further studies on the friction-driven SAW motors have been reported in details including the measurements, analyses, modeling, physics of contact, and design criteria, which were published in five papers separately (Shigematsu & Kurosawa, 2008a; 2008b; 2008c; 2008d; 2008e). It is clearly that these papers provide a systematical information, experiments and theories for optimizing the designs, manufactures and applications of the friction-driven SAW motors.

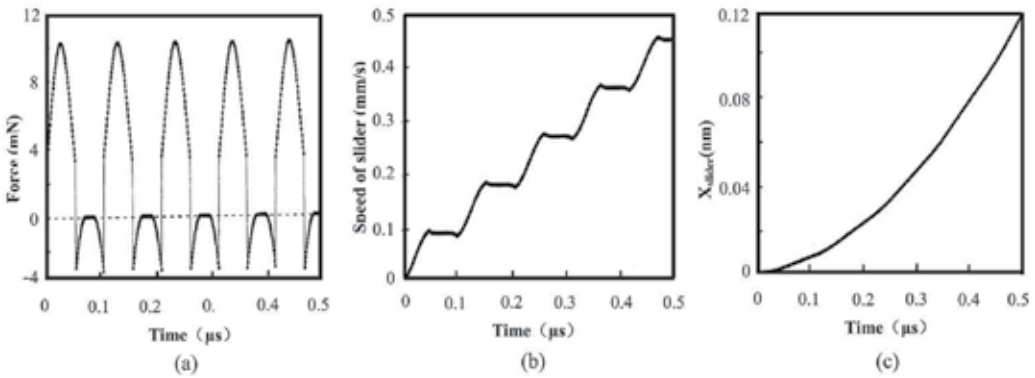


Fig. 8. Numerical simulations of transient response for linear motor: (a) driving force; (b) moving speed; (c) displacement

4.2 Rotary motor

On the basis of the theoretical model of the linear motors, the similar theoretical model of the rotary motors has been constructed (Cheng et al., 2002; 2003a; 2003b). Owing to the axisymmetry, only two balls at the fringe of the rotor are considered for simplicity as shown in Fig.9(a).

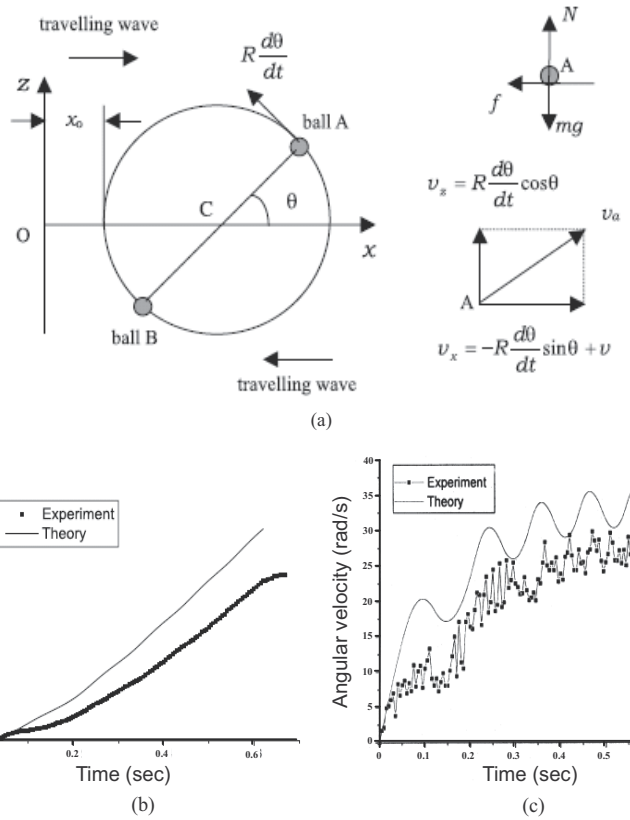


Fig. 9. Numerical simulations of rotary motors: (a) theoretical model; (b) angular displacement versus time; (c) anglular velocity versus time.

When the ball is levitating, the motion equations are as

$$\begin{cases} md^2y_{ball} / dt^2 = -mg, \\ Id^2\theta / dt^2 = -M, \end{cases} \tag{6}$$

where m is the mass of the ball, I is the moment of inertia of the rotor with respect to the axis. When the ball is contacting the stator, the motion equations are

$$\begin{cases} m\dfrac{d^2y_{ball}}{dt^2} = N - mg, \\ I\dfrac{d^2\theta}{dt^2} = 2\mu N \dfrac{(v - R\dfrac{d\theta}{dt}\sin\theta)R\sin\theta - R\dfrac{d\theta}{dt}\cos\theta \cdot R\cos\theta}{\sqrt{(R\dfrac{d\theta}{dt}\cos\theta)^2 + (v - R\dfrac{d\theta}{dt}\sin\theta)^2}} - M, \end{cases} \tag{7}$$

where R is the radius of the rotor. The supporting force on the ball is calculated as

$$N = \frac{4Er^{\frac{1}{2}}d^{\frac{1}{2}}}{3}, \qquad E = \left(\frac{1-\sigma_1^2}{E_1} + \frac{1-\sigma_2^2}{E_2}\right)^{-1}, \tag{8}$$

where r is the radius of the small ball.

For the rotary motor with the SAW frequency of 30 MHz, the experimental and theoretical results of the angular displacement and speed are shown in Fig.9(b) and (c) (Cheng et al., 2002). From the figure, it can be seen that the simulation results are roughly consistent with the experimental results with slight deviations, which might be induced by the following reasons: the friction between the axis and the rotor is neglected and, especially, only two balls' motions are considered in the theoretical calculations for simplicity. Further calculations showed that as the number of the balls increases, the rotary motion is more stable and closer to that of the experiments (Cheng et al., 2003b).

5. Non-contact SAW motors and actuators

The SAW motors described above are driven by the frictional forces generated by direct contact between the sliders (or rotors) and the stators, which restrict the motion velocity and working lifetime of the motors. On the other hand, as the amplitude of the SAW decreases to less than the surface roughness of the stators, it is not possible to drive the motors by the frictional force. In order to overcome the deficit of the friction-driven motors and expand the application fields of the motors, non-contact SAW motors were proposed, in which a fluid layer is introduced in the interface between the stator and slider (rotor). As a high frequency voltage is supplied to the IDT which is deposited on the stator, the SAW (Rayleigh wave) will be excited by the IDT, which further radiate longitudinal waves in the fluid media between the slider (rotor) and stator. As the second-order effect of the wave propagation, the acoustic streaming is induced, whose viscous force drives the slider (rotor) to move.

5.1 SAW streaming

When acoustic waves travel through a medium, if the acoustic intensity is not high, the acoustic wave propagation is a linear phenomenon, i.e., only the acoustic energies propagate, the medium does not move globally. However, as the wave amplitude increases, the nonlinear effect appears, and an interesting feature of the sound field, i.e., the medium presents steady motions, becomes evident. Such a nonlinear phenomenon is called "acoustic streaming" (Nyborg, 1958; 1965). These flows can be generated around the surfaces of obstacles immersed in intensive acoustic fields and/or vibrating elements and near bounding walls. Similarly, when SAWs (Rayleigh waves) with high intensity reach the boundary between fluids and solids, the transmission mode is changed from the Rayleigh mode to the leaky Rayleigh mode. Meanwhile, the leaky Rayleigh waves radiate the acoustic energy in the fluid, and then the fluid is driven to move, which is called "SAW streaming" because the phenomenon involves essentially the same physics as the "acoustic streaming" (Shiokawa et al., 1990).

(a) Experiment of SAW streaming

The Rayleigh wave can readily radiate a longitudinal wave into a fluid when the surface with the Rayleigh wave propagation is in contact with a fluid layer. When the fluid volume is small and the acoustic power increases above a certain threshold, the fluid begins to move in the direction of the SAW propagation, which provides an observable pattern of the SAW energy flow. Moreover, if the substrate surface is hydrophobic, a lot of droplets are expelled from the liquid surface as shown in Fig.10 (Shiokawa et al., 1989; 1995; Uchida et al., 1995).

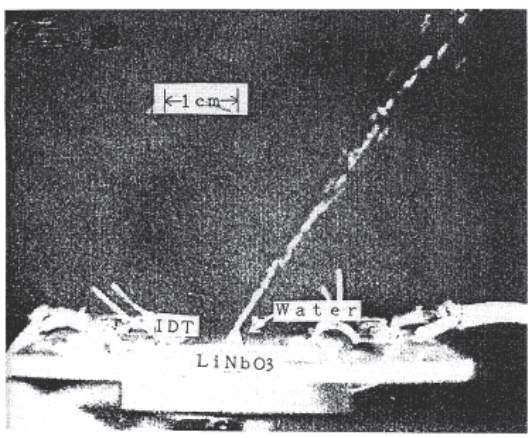


Fig. 10. Water jet streaming excited by SAW.

(b) Theory of SAW streaming

The general equation of hydrodynamics of viscous fluids is as follows (Shiokawa, et al., 1990):

$$F = -\nabla P + \eta \nabla^2 v + \frac{1}{3}\eta \nabla(\nabla \cdot v), \tag{9}$$

where F is the net force per unit volume due to stress, ρ is the density, η is the shear viscosity coefficient and v is the particle velocity.

Based on the continuity equation, one can obtain:

$$\frac{\partial(\rho v)}{\partial t} + \rho(v \cdot \nabla)v + v\nabla \cdot \rho v = F_0 - F_1, \tag{10}$$

where

$$F_0 = \frac{\partial(\rho v)}{\partial t}, \quad -F_1 = \rho(v \cdot \nabla)v + v\nabla \cdot \rho v. \tag{11}$$

For acoustic waves vary sinusoidally with time at the frequency ω, in order to obtain a streaming equation from Eqs. (10) and (11), it is need to retain the terms up to the second order and take the time average over a suitable number of cycles. Noting that the time average of F_0 should be zero in the steady state, then

$$-F_1 = \left\langle \rho_0(v \cdot \nabla)v + v\nabla \cdot \rho_0 v \right\rangle, \tag{12}$$

in which $\langle\ \rangle$ means "time average", and F_1 is the exact force of the acoustic streaming.

5.2 Non-contact SAW linear motors and actuators

(a) Initial non-contact SAW linear motors

The first non-contact linear SAW motor (manipulator) using Rayleigh wave was proposed as shown in Fig.11(a), in which a liquid droplet is put on the surface of the stator and a small slider is placed on the liquid droplet (Sano et al., 1997). The propagating SAW drives the liquid droplet to move in the SAW propagation direction by SAW streaming, and the slider moves with the droplet together. For the motors with different frequencies, the amplitude of the Rayleigh waves measured by a laser probe and normalized by the wavelength versus the driving voltage is shown in Fig. 11(b), and the dependence of moved distance of the slider on the voltage is shown in Fig. 11(c), which shows that the motors are suitable to be used as non-contact actuators, such as micromanipulators.

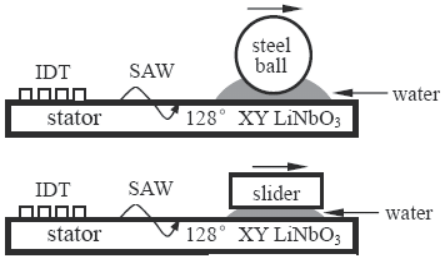
(b) New type of non-contact SAW linear motors

Recently, a new type of non-contact SAW motors was proposed, in which a thin liquid layer surrounded by a thin glass wall was put on the surface of the stator as shown in Fig.12(a). In the experiments, two kinds of IDTs with different frequencies and sizes were deposited, and the effects of the SAWs with different frequencies on the performances of the non-contact motors were studied (Gu et al., 2009).

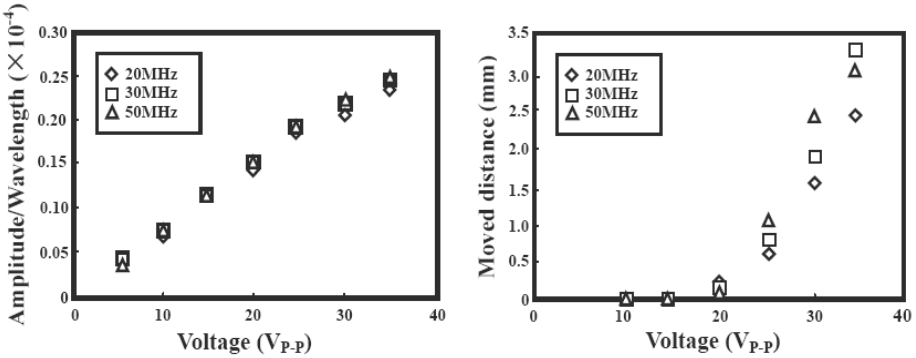
5.3 Theoretical model of non-contact SAW linear motors

If the SAW propagation surface is in contact with liquid, based on the ultrasonic wave radiation mechanism, an approximate theoretical model is developed as shown in Fig.12(b). In the liquids (z > 0), the particle displacement (u_x, u_z) can be written as (Shiokawa et al., 1989)

$$\begin{aligned} u_x &= A\exp j\omega t \cdot \exp(-jk_L x) \cdot \exp(-\alpha k_L z) \\ u_z &= -j\alpha A\exp j\omega t \cdot \exp(-jk_L x) \cdot \exp(-\alpha k_L z), \\ k_L &= \omega / V_L, \alpha^2 = 1 - (V_L / V_W)^2. \end{aligned} \tag{13}$$

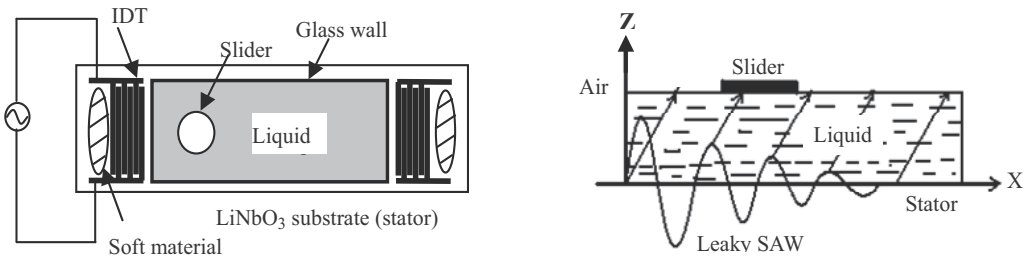


(a)



(b) (c)

Fig. 11. SAW manipulator: (a) schematic diagram; (b) amplitude/wavelength VS voltage; (c) distance VS voltage.



(a) (b)

Fig. 12. Non-contact SAW linear motors: (a) schematic diagram of structure; (b) theoretical model

where A is the horizontal vibration amplitude of the particle, ω is the angular frequency (for simplicity, the SAW harmonics are neglected), K_L is the wave number, and a is a complex propagation constant of the leaky SAW, V_L and V_W are the velocities of the leaky SAW and longitudinal wave of the liquid, respectively. V_L can be calculated by applying a perturbation method, which treats the leaky wave as a first-order perturbation on the non-leaky wave propagation, and assuming the boundary conditions satisfy that both the displacement and the stress are continuous at the interface between the liquid and solid substrate (z = 0) (Campbell & Jones, 1970).

According to the theory of near-boundary acoustic streaming, the basic equations used for the streaming are (Nyborg, 1958; Shiokawa et al., 1990)

$$\vec{F} = -\rho_0 \langle \vec{U}_1 \cdot \nabla \vec{U}_1 + \vec{U}_1 \nabla \cdot \vec{U}_1 \rangle , \tag{14}$$

$$\mu \nabla^2 \vec{U}_2 - \nabla \vec{P}_2 + \vec{F} = 0 , \tag{15}$$

where ρ_0 and μ are the density and viscosity of the fluid respectively, the bracket indicates a time average over a large number of cycles, $\vec{U}_1$ is the particle velocity of the stator surface, $\vec{U}_2$ is the streaming velocity, and $\vec{P}_2$ is the acoustic pressure of the liquid at the interface between the liquid and stator. Both $\vec{U}_2$ and $\vec{P}_2$ are time-independent second-order perturbation quantities as the effect of the viscosity of the liquid (such as water) is ignored, $\vec{F}$ is the acoustic streaming force, which is a nonlinear quantity.

Substituting the x and z components of $\vec{U}_1$, i.e., u_x and u_z of Eq. (13), into Eqs. (14), the x and z components of $\vec{F}$, i.e., F_x and F_z, are given by

$$\begin{aligned} F_x &= -\rho_0 (1+\alpha_1^2) A^2 \omega^2 k_i \exp 2(k_i x + \alpha_1 k_i z) \\ F_z &= -\rho_0 (1+\alpha_1^2) A^2 \omega^2 \alpha_1 k_i \exp 2(k_i x + \alpha_1 k_i z) \end{aligned} , \tag{16}$$

where α_1 and k_i are the imaginary parts of α and k_L, respectively. F_z does not contribute to the motion of the slider, so the F_x is just considered as the force acting on the levitated slider. Since k_i is always a minus value, F_x is attenuated with the distance x.

The following analyses are conducted with the model shown in Fig.12(b), in which the thickness of the liquid is z_0 and the acoustic streaming force acted on the slider in x direction is $F_x(z_0)$. As the slider is with the area L^2 and thickness h, considering the thickness of the slider is generally very small compared with the acoustic wavelength, the acoustic streaming force is constant in the thickness range of the slider. Thus the acoustic streaming force acted on the slider at x position, $F_{sum}(x)$, is given by:

$$F_{sum}(x) = Lh \int_x^{x+L} F_x(z_0) dx. \tag{17}$$

Therefore, the acceleration of the slider at x position is given by:

$$a = \frac{F_{sum}(x)}{\rho_s h L^2} , \tag{18}$$

where ρ_s is the density of the slider.

Considering the resistant force of the liquid, there is a threshold force $F_t(x)$ for the slider starts to move, i.e., only as the horizontal vibration amplitude $A > A_r$, the streaming force can drive the slider to move, then the acceleration can be rewritten as

$$a = \frac{F_{sum}(x) - F_t(x)}{\rho_s h L^2}. \tag{19}$$

According the acceleration, the transient velocity of the slider can be obtained, so as the displacement.

5.4 Experimental and theoretical results

In the experiments, the substrates of the SAW delay lines (stators) are 128^0 YX-$LiNbO_3$ piezoelectric crystals with the size of about 75 ×18 ×1 mm^3. On the substrates, two kinds of IDT pairs with different frequencies and sizes are fabricated for studying the effects of the operation frequency on the performance of the non-contact motors (Gu et al., 2009). In the first type of the motors, the IDT has 20 pairs of electrodes with the aperture of 12 mm and the central frequency of about 10 MHz, where a rectangular cell (31.3 ×12.8 ×12 mm^3) surrounded by thin glass walls with the thickness of 0.8 mm is put on the surface of the stator to contain the liquid (water) layer. The other type has 45 pairs of electrodes with the aperture of also 12 mm and the central frequency of about 24 MHz, and a rectangular cell with the size of 31.2 ×12 ×12 mm^3 is used. The exciting voltage $V_{p\text{-}p}$ of the IDT can be adjusted in the range of 0-140 V. The slider is a circular aluminum slice with the diameter of 6 mm, the thickness of 0.24 mm, and the density of 2.7 g/cm^3. The thickness of the liquid layer between the stator and slider can be adjusted in the range of 1-8 mm.

The motion (displacement) of the slider is recorded by a high-speed digital video camera, and then the data of the motion are extracted and processed by a computer. Thus the transient velocity of the sliders can be measured and calculated. Since the transient velocity in fluctuating, the average of the transient velocities is taken as the moving velocity of the slider. According to the results, the first motor with the frequency about 10 MHz has faster transient velocity than the second one (about 24 MHz).

(a) Transient velocity versus time

For the first motor, the transient velocity of the slider is obtained as shown in Fig. 13(a), in which the thickness of the liquid (water) layer is about 3.2 mm and the driving voltage $V_{p\text{-}p}$ is 21.2 V. From Fig.13(a) it can be seen that the transient velocity increases very fast at first (about 0.1 second), and then approaches to saturation with slight oscillations. The oscillation may be induced by the vortical flows and the measurement errors. For the second motor, similar results can also be obtained, however, in order to obtain the same transient velocity at the same thickness of the water layer, the exciting voltage $V_{p\text{-}p}$ should be about 79.2 V.

(b) Velocity of slider versus exciting voltage

The velocity variations of the slider with the exciting voltages of the motors are investigated. For the first motor, as the thickness of the liquid (water) layer is about 4.0 mm, the result is shown in Fig.13(b), in which the threshold voltage V_t is about 12 V. When the exciting voltage $V_{p\text{-}p}$ increases from 12 V to 60 V, the velocity increases linearly essentially, but the velocity increases with a little slower in the lower voltage range (less than 30 V). Similarly, for the second motor, the threshold voltage V_t is about 20 V, and the velocity of the slider increases approximately linearly with the voltage in the range of 80 – 120 V, although in the lower voltage range (20 – 80 V), the increasing tendency of the velocity is slower. Since the experimental conditions are very complicated, such as the reflections of waves by the glass walls of the cell and so on, the experimental errors are always in the range about 10%.

Generally speaking, the velocity varying behavior against the exciting voltage of the SAWs is similar for both motors with different frequencies. There exists a limited threshold voltage, below which the sliders cannot move. As the voltages are larger than the threshold, the velocity of the sliders increases with the exciting voltage approximately in linear way, although the increase is more like exponential in the lower voltage range.

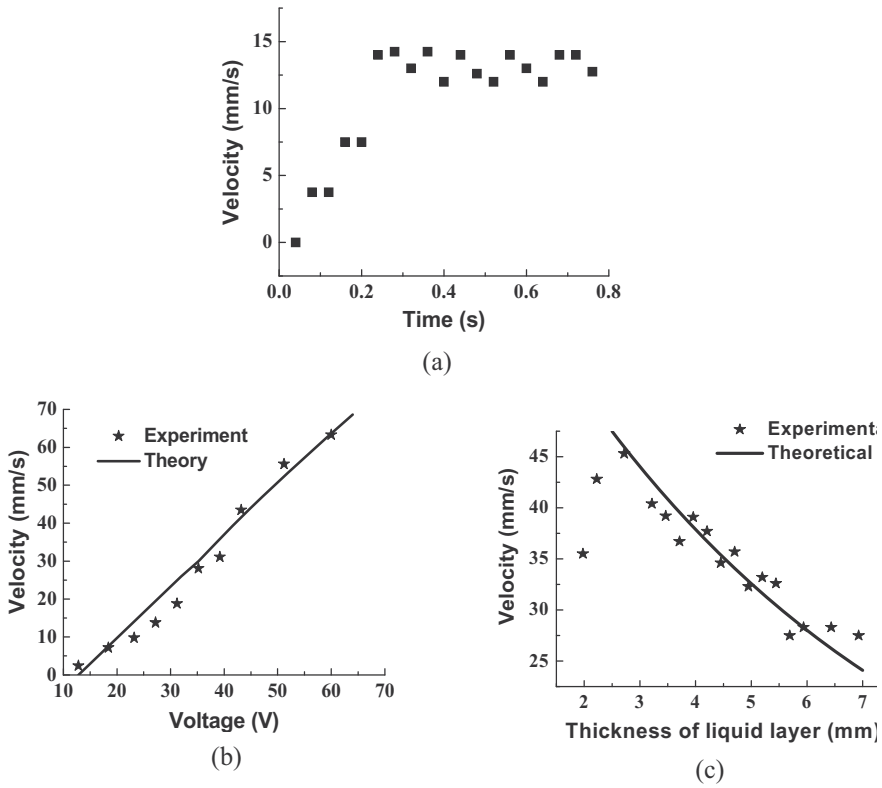


Fig. 13. Velocity of non-contact SAW linear motor: (a) transient velocity; (b) average velocity VS voltage; (c) average velocity VS thickness of liquid layer (f=9.959 MHz)

In the theoretical simulations, taking V_L = (3931 + j67.7) m/s and V_W = 1500 m/s (Campbell & Jones, 1970; Shiokawa et al., 1989), the relation between the slider velocity and the amplitude of the SAW for the motor can be obtained by Eqs.(16) - (19). Considering the amplitude of the SAW is proportional to the exciting voltage (Sano et al., 1997), the relation between the slider velocity and the exciting voltage for both motors can be obtained, one of which is shown in Fig.13(b), which shows that the slider velocity increases linearly with the exciting voltage as the voltage is beyond the threshold. It demonstrates that the theoretical result is essentially in agreement with that of the experiments.

Compared with the performance of the contact SAW motors, it can be found that the driving voltage for the non-contact motors is reduced distinctly. Therefore, the moving velocity and lifetime can be increased greatly.

(c) Slider velocity versus thickness of liquid layer

The relation between the slider velocity and the thickness of the liquid (water) layer for both motors is also studied experimentally and theoretically. In the experiments, it is found that the varying tendency of the slider velocity versus the thickness of the liquid layer for both motors is very similar. The result of the first motor is shown in Fig.13(c), which shows that, when the thickness of liquid (water) layer is larger than 3 mm, the velocity of the slider decreases exponentially with the increasing thickness of the water layer. However, as the thickness of the water layer is less than 3 mm, the slider velocity reduces with the

decreasing thickness of the water layer, which might be attributed to that the viscous drag effect increases when the thickness of the water layer gradually becomes small (Betchov & Criminale, 1967). As a result, while the thickness of the water layer increases gradually in the range of 0 – 3 mm until the liquid flow becomes stability, the velocity of the slider becomes faster. Besides, as the liquid layer thickness is less than 1 mm, the slider cannot be driven to move by the exciting voltage ($V_{p\text{-}p}$) of 50 V. In this case, the slider may be sometimes in contact with the stator since the liquid layer is too thin. It is also implied that the slider motion of the non-contact SAW motors is driven much easier than that of the contact SAW motors with the same SAW devices (stators).
Meanwhile, the theoretical simulations are also accomplished by suitably selecting the relevant parameters. The simulated result for the first motor is also plotted in Fig.13(c). It can be seen that the theoretical result is in good agreement with that of the experiment.

5.5 Non-contact SAW rotary motor

On the basis of the friction-driven rotary motors, non-contact rotary motors were developed, in which the same rotors and substrates are used as in contact motors. Similarly, the rotor is suspended in a thin liquid layer filled in a liquid cell located on the stator, and the angular displacement and angular speed are measured in different driving voltages as shown in Figs.14(a) and (b). In addition, to study the influence of the viscosity of the liquid layer, the mixed liquids composed of water and glycerine with different ratios are used, and the dynamic viscosity is measured by a viscometer. As the operating frequency is 9.845 MHz, the dynamic response is obtained at different times as shown in Fig. 14 (a), the angular speed changing with the driving voltage and the dynamic viscosity of the liquid are also obtained as shown in Figs.14(b) and (c), respectively. The related theoretical study is in progress.

6. Microactuator driven by different SAW modes

6.1 Microactuators driven by Lamb waves

An initial micromotor using a Lamb wave (flexural plate wave) has been presented, in which the stator is a ZnO film (1μm) deposited on Si_3O_4 film (2 μm) and both films are deposited on a Si wafer, and then the related area of Si wafer is etched as shown in Fig.15 (Moroney et al., 1989). A small polysilicon slice put on the Si_3O_4 film is driven by the Lamb wave excited by IDT fabricated on the ZnO film, the polysilicon slice can move linearly or rotationally as the Lamb waves are excited or reflected in different directions. This kind of motors can be used in MEMS technology.
Furthermore, since Lamb wave devices are able to operate in fluids (water), a fluid motion (SAW streaming) produced by Lamb waves excited by IDT at the composite of Si_3O_4 and ZnO membrane with the thickness of 4 μm was also observed (Moroney et al., 1991). The velocity of the flexural plate wave is much lower than the sound velocity in fluids (water), thus the structure acts as nearly lossless acoustic waveguides. Then the fluorescent polystyrene spheres with the diameter about 2.5 μm are put in the fluid to make the motion visible. The observed pumping speed is proportional to the square of the wave amplitude, the speed was 100 μm/s for a RF driving voltage of 8V and a wave amplitude of 6.5 nm. A nonlinear model based on acoustic streaming theory was presented to predict the velocities, which was in good agreement with the experiments.

In addition, some other micromachined actuators have also been presented by several groups, where ultrasonic flexural plate waves traveling along thin piezoelectric membranes were used to excite acoustic fields in the fluids contacted with the membranes (Luginbuhl et al., 1998; Meng et al., 2004). These kinds of microactuators have been used for transportation of particles and/or droplets in micro-biological research fields.

6.2 Micromanipulators driven by leaky waves

Practically, micromanipulators using leaky waves in thin liquid layers produced by SAW devices with very high frequency about 100 MHz were fabricated by several groups. One of them is shown in Fig. 16 (Takeuchi et al., 1994), where the tip of the substrate of the Rayleigh wave device was immersed with a Rayleigh angle in a liquid layer (or drop). At the interface of the substrate and liquid, the SAW is converted to leaky waves, which then become longitudinal waves in the liquid. As the wave intensity is high, the acoustic streaming (or radiated force) drives small glass particles to move in 1- or 2-dimensional way in the liquid if there are two pairs of SAW devices located perpendicular to each other. It can be used in bioengineering and micromachining (Renaudin et al., 2006).

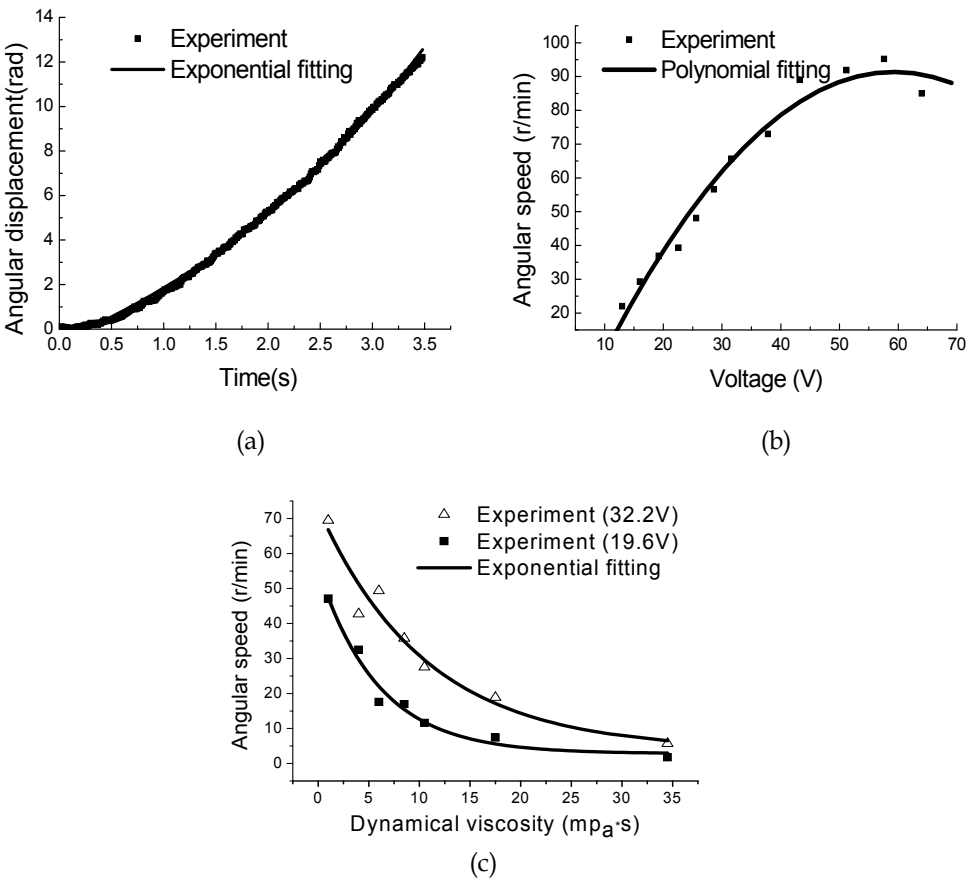


Fig. 14. Response of non-contact rotary motor: (a) angular displacement VS time; (b) angular speed VS voltage; (c) angular speed versus the dynamical viscosity of liquid layer

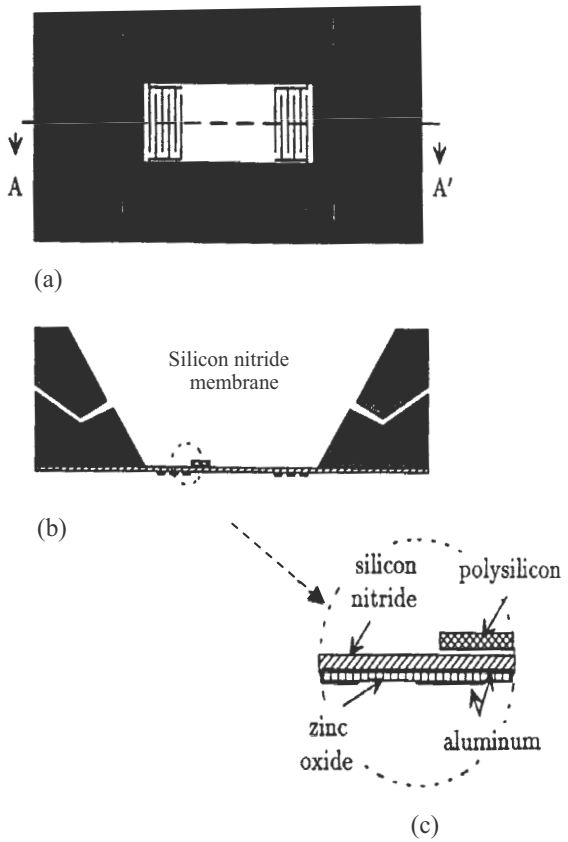


Fig. 15. Micromotor driven by Lamb wave: (a) back side; (b) cross section; (c) enlarged pattern of circle in (b)

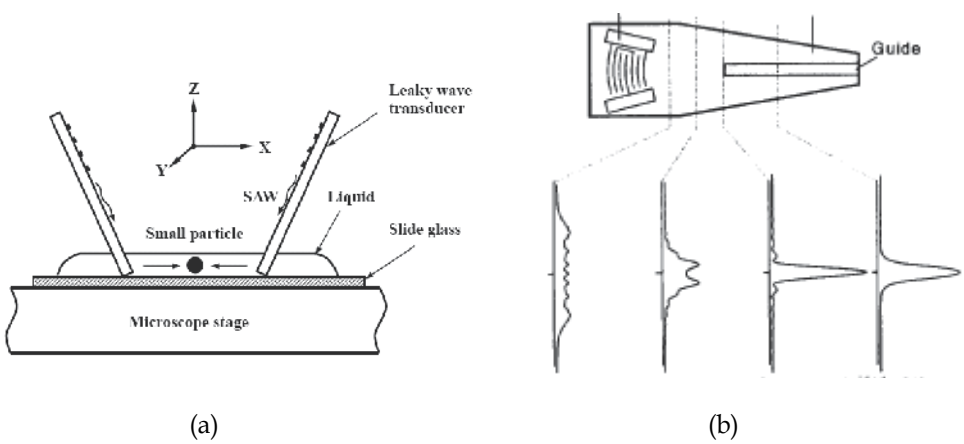


Fig. 16. Manipulator driven by leaky wave: (a) sketch of set up; (b) SAW device and corresponding acoustic field

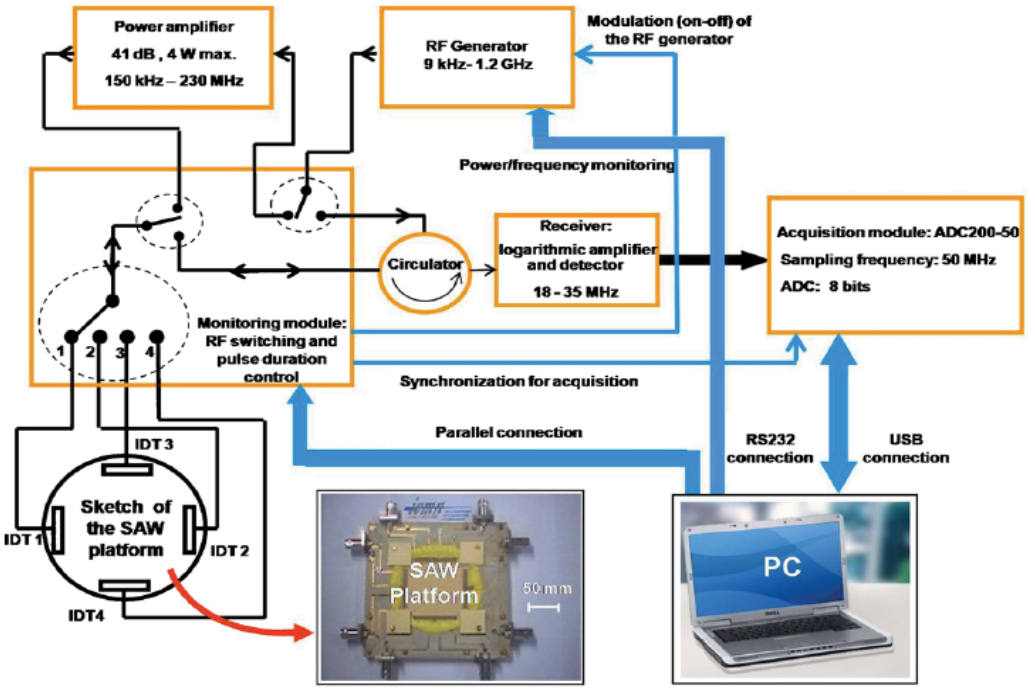


Fig. 17. Schematic diagram of SAW actuator system used for tracking microdroplet.

7. Potential applications

Apparently, the SAW motors and actuators have wide applications in micro-electro-mechanical systems and biological sciences and technologies. As examples, the motors would be capable of operating in multiple degrees of freedom, such as two dimensional translation stages and self-actuated spherical joints to be rotated about two orthogonal axes, which are amenable to integrate into diverse structures and machines, such as robots (NASA Tech Briefs, 2003). Practically, SAW-actuated devices have also been used for drop translation, microchannel pumping, drop manipulation, micromixing, microparticle collection and concentration, forming specific examples of the potential of technology, as lab-on-a-chip devices for microarray technology and rapid bioscaffold cell seeding, etc. (Qi et al., 2008). Meanwhile, the acoustic streaming has also been used in atomization systems.
As an example, a set up of SAW actuator used for tracking droplets is shown in Fig. 17 (Renaudin et al., 2009).

8. Conclusion

The conventional linear and rotary motors driven by IDT-excited SAWs in the frequency range of 1–100 MHz have been introduced briefly. For the conventional SAW motors, the sliders (or rotors) and the stators are in direct contact, and the driving mechanism is the frictional force between the sliders (rotors) and the stators. On the other hand, a new type of noncontact motors has been developed. Instead of the frictional forces, the acoustic streaming, which is excited by SAWs propagating in the fluids covered on the surfaces of the stators, is used to drive the sliders.

The SAW motors as micro-actuators will be very useful in advanced industrial technologies and/or biotechnical engineering due to their various advantages, such as very short response time, low driving voltage, large output force or torque, miniaturized size, easy control and batch production, etc.

On the other hand, several theoretical models have been presented and the numerical simulations for the motors and actuators in contact and noncontact types have also been performed. The theoretical results are in good agreement with those of the experiments, which demonstrate that the theoretical models can well be used to explain the working mechanism of the motors (actuators). Therefore, by combining the theoretical models and experimental results, the constructions and the performances of various SAW motors (actuators) can be optimized.

9. Acknowledgements

This work is supported by national natural science foundation of china (grant no. 10774074 and no. 10874083).

10. References

Asai, K.; Kurosawa, M. K. & Higuchi, T. (1999) Novel power circulation methods for a surface acoustic wave motor, *IEEE Ultrasonic Symposium Proceedings*, 667-670.

Auld, B. A. (1973) *Acoustic Fields and Waves in Solids*, A Wiley-Interscience Publication, John Wiley & Sons, New York, USA.

Betchov, R.& Criminale, W. O. J. (1967) *Stability of parallel flows*, Academic Press, New York, USA

Campbell, J. J. & Jones, W. R. (1070) Propagation of surface waves at the boundary between a piezoelectric crystal and a fluid medium, *IEEE Transaction on Sonics and Ultrasonics*, Vol. SU-17, 71-76.

Cheng L. P.; Zhang G. M.; Zhang S. Y.; Yu J. & Shui X. J. (2002), Miniaturization of surface acoustic waves rotary motor, *Ultrasonics*, Vol. 39, 591-594.

Cheng, L. P.; Zhang, G. M.; Zhang S. Y.; Zhang Z. N. & Shui X. J. (2003a) Theoretical simulation of a rotary motor driven by surface acoustic waves, *Acoustical Physics*, Vol. 49, 194-198.

Cheng, L. P.; Zhang, G. M.; Zhang S. Y.; Yu J. & Shui X. J. (2003b) Study of surface acoustic wave rotary motors, *Chinese Journal of Acoustics*, Vol.20, 323-328.

Cheng, L. P. & Zhang, S.,Y. (2007) Simulation of the flow induced by acoustic streaming in noncontact ultrasonic motors, *Applied Physics Letters*, Vol. 90, 244106.

Dong, S.; Wang, S.; Shen, W. & Li, L. (2000) A miniature piezoelectric ultrasonic motor based on circular bending vibration mode, *ASME Transactions on Mechanics*, Vol. 5, 325-330.

Fujii, Y.; Kotani, H.; Masaya, T.; Mizuno, T.; Aoki, Y.; Adachi, Y. & Ohtake, N. (2007) Surface acoustic wave linear motor using segment-structured diamond-like carbon films on contact surface, *IEEE Ultrasonic Symposium Proceedings*, 2543-2546.

Gu, H.H.; Cheng, L.P.; Zhang, S.Y.; Zhou, F. M. & Shui, X. J. (2008) Experimental Study on non-contact linear motors driven by surface acoustic waves, *IEEE Ultrasonic Symposium Proceedings*, 1835-1837.

Gu, H.H.; Zhang, S.Y.; Cheng, L.P.; Ma, D.; Zhou, F. M.; Chen, Z. J. & Shui, X. J. (2009) Study on non-contact linear motors driven by surface acoustic waves, *Sensors and Actuators*, A155, 163-167.

Helin, P.; Sadaune, V. & Dron, C. (1998) Theoretical and experimental study of linear motors using surface acoustic waves, *Sensors and Actuators* A, Vol. 70, 67-74.

Hu, J. H.; Yamazaki, T.; Nakamura & Ueha, S. (1995) An analysis of a noncontact ultrasonic motor with an ultrasonically levitated rotor, *Japanese Journal of Applied Physics*, Vol. 34, 2702-2706-

Koster, D. (2007) Numerical simulation of acoustic streaming on surface acoustic wave-driven biochips, *SLAM Journal on Scientific Computing*, Vol. 29, 2352-2380.

Kurosawa, M.; Takahashi, M. & Higuchi, T. (1994) An ultrasonic X-Y stage using 10 MHz surface acoustic wave, *IEEE Ultrasonic Symposium Proceedings*, 535-538.

Kurosawa, M., Takahashi, M. and Higuchi, T., (1996) Ultrasonic linear motor using surface acoustic waves, *IEEE Transaction on Ultrasonics, Ferroelectronics, and Frequency Control*, Vol. 43, 901-906.

Kurosawa, M. K., (2000) State-of-the-art surface acoustic wave linear motor and its future applications, *Ultrasonics*, Vol. 38, 15-19.

Luginbuhl, Ph.; Collins, S. D.; Racine, G. A.; Gretillat, M. A.; de Rooij, N. f.; Brooks, K. G. & Setter, N. (1998) Ultrasonic flexural Lamb-wave actuators based on PZT thin film, *Sensors and Actuators*, Vol. A 64, 41-49.

Meng A. H.; Nguyen N. T. & White R. M., (2004) Focused flow micropump using ultrasonic flexural plate waves, Biomedical Microdevices, Vol. 2, 169-174.

Morita, T., Kurosawa, M. K. and Higuchi, T., (1999) Simulation of surface acoustic wave motor with spherical slider, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. 46, 929-934.

Moroney, R. M.; White, R. M. & Howe, R. T. (1989) Ultrasonic micromotors: physics and application, *IEEE Ultrasonic Symposium Proceedings*, 745-748.

Moroney, R. M.; White, R. M. & Howe, R. T., (1991) Microtransport induced by ultrasonic Lamb waves, Applied Physics Letters, Vol. 59, 774 - 776

Nakamula, K.; Ito, T., Kurosawa, M. & Ueha, S. (1990) A trial construction of an ultrasonic motor with fluid coupling, *Japanese Journal of Applied Physics*, Vol. 29, L160-L161.

Nyborg, W. L. (1958) Acoustic streaming near a boundary, *The Journal of the Acoustical Society of America*, Vol. 30, 329-339.

Nyborg, W. L. (1965) Acoustic streaming. In: *Physical Acoustics*, Mason W. P. (Ed), Vol. 2B, 265–283, Academic Press, Englewood Cliffs, USA,.

Qi, A.; Yeo, L.Y. & Friend, J.R. (2008) Interfacial destabilization and atomization driven by surface acoustic waves, *Physics of Fluids*, Vol. 20, 074103-1-14.

Renaudin, A.; Tabourier, P.; Camart, J. C.& Druon, C. (2006) Surface acoustic wave two-dimensional transport and location of microdroplets using echo signal, *Journal of Applied Physics*, Vol. 100, 161101)

Renaudin, A.; Sozanski, J. P.; Verbeke, B.; Zhang, V.; Tabourier, P. & Bruon, C. (2009) Monitoring SAW-actuated microdroplets in view of biological applications, *Sensors and Actuators* B, Vol. 138, 374-382.

Sakano K, Kurosawa M. K., Shigematsu T., (2008) Surface acoustic wave motor with flat plane slider, *International Symposium on Micro-Nanomechatronics and Human Science*, 243-248.

Sano, A.; Matsui, Y. & Shiokawa, S. (1997) A new manipulator based on surface acoustic wave streaming, *IEEE Ultrasonic Symposium Proceedings*, 467-470.

Sashida, T. & Kenjo, T. (1993) *An introduction to ultrasonic motors*, Oxford University Press Inc., New York, USA.

Shigematsu, T. & Kurosawa, M. K. (2006) Miniaturized SAW motor with 100 MHz driving frequency, *IEE Japanese Transactions on Sensors and Micromachines*, Vol. 126, 166-167.

Shigematsu, T. & Kurosawa, M. K. (2008a) Friction drive of an SAW motor. Part I: Measurements, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol.55, 2005-2015.

Shigematsu, T. & Kurosawa, M. K. (2008b) Friction drive of an SAW motor. Part II: Analyses, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. 55, 2016-2024.

Shigematsu, T. & Kurosawa, M. K. (2008c) Friction drive of an SAW motor. Part III: Modeling, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol.55, 2066-2276.

Shigematsu, T. & Kurosawa, M. K. (2008d) Friction drive of an SAW motor. Part IV: Physics of contact, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol.55, 2277-2287.

Shigematsu, T. & Kurosawa, M. K. (2008e) Friction drive of an SAW motor. Part V: Design criteria, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol.55, 2288-2297.

Shiokawa, S., Matsui, Y. & Ueda, T. (1989) Liquid streaming and droplet formation caused by leaky Rayleigh waves, *IEEE Ultrasonic Symposium Proceedings*, 643-646.

Shiokawa, S., Matsui, Y. & Ueda, T. (1990) Study on SAW streaming and its application to fluid devices, *Japanese Journal of Applied Physics*, Vol. 29, 137-139.

Shiokawa, S. & Matsui, Y. (1995) The dynamics of SAW streaming and its application to fluid devices, *Material research Society Symposium Proceeding*, Vol. 360, 53-64

Takasaki, M., Osakabe, N., Kurosawa, M. & Higuchi T. (1998) Miniaturization of surface acoustic wave linear motor, *IEEE Ultrasonic Symposium Proceedings*, 679-682.

Takasaki, M., Kurosawa, M. K. & Higuchi T., (2000) Optimum contact conditions for miniaturized of surface acoustic wave linear motor, *Ultrasonics*, Vol. 38, 51-53.

Takeuchi, M., Abe, H. and Yamanouchi, K. (1994) Ultrasonic micromanipulation of small particles in liquid using VHF-range keaky wave transducer, *IEEE Ultrasonic Symposium Proceedings*, 607-610.

Takeuchi, M. and Nakano, K. (2005) Ultrasonic micromanipulation of liquid droplets for a lab-on-a-chip, *IEEE Ultrasonic Symposium Proceedings*, 1518-1521.

Timoshenko, S. & Goodier, J. (1970) *Theory of Elasticity*, Third Edition, McGraw-Hill, New York, USA.

Uchida, T.; Suzuki, T. & Shiokawa, S. (1995) Investigation of acoustic streaming excited by surface acoustic waves, *IEEE Ultrasonic Symposium Proceedings*, 1081-1084.

Ueha, S. & Tomikawa, Y. (1993) *Ultrasonic Motors: Theory and Application*, Oxford Science Publication, London, UK

Viktorov, I. A. (1967) *Rayleigh and Lamb Waves*, Plenum, New York, USA.

White, R. M. (1970) Surface elastic waves, *Proceedings IEEE Micro-electro-mechanical Systems*, Vol. 58, 1238-1276.

Yamayoshi, Y. & Hirose, S. (1992) Ultrasonic motor not using mechanical friction force, *International Journal of Applied Electro-magnetics in Materials*, Vol. 3, 179-182.

Yeo, L. Y. & Friend, J. R. (2009) Ultrafast microfluidics using surface acoustic waves, *Biomicrofluidics*, Vol. 3, 012002.

Zhang, G. M.; Cheng, L. P.; Zhang, S. Y.; Yu, J. & Shui, X. J. (2000) Surface acoustics wave rotation motor, *Electronics Letters*, Vol. 36, 1437-1438.

Zhang, H.; Dong, S. X.; Zhang, S. Y.; Wang, T. H.; Zhang, Z. N. & Fan, L. (2006) Ultrasonic micro-motor using miniature piezoelectric tube with diameter of 1.0 mm, *Ultrasonics*, Vol. 44, e603-e606.

11

Real Time Methods for Wideband Data Processing Based on Surface Acoustic Waves

N. V. Masalsky
Research Institute of System Researches, Russian Academy of Sciences,
Nakhimovskii pr. 36, korp. 1, Moscow, 117218
Russia

1. Introduction

The newly developed methods enjoy the advantages of an optical channel for data transfer (parallel data processing in real time, high operating speed, and noise immunity) and the advantages typical of integrated optics (compactness, low power consumption, high sensitivity in information signal processing, stability against external action, and the possibility of using the batch-fabrication technique). The aforesaid advantages account for the wide use of waveguide acoustooptic (AO) units (WAOU) in computing, fiber-optic, telecommunication, and other photonic systems [1-3]. Conventional WAOUs are constructed from 3D elements. However, size, power consumption, and sensitivity to environmental conditions (vibrations, temperature variations, etc.) set limits on the applications of these devices. The WAOU (typical functional scheme is shown on fig. 1) is based on the waveguide AO Bragg diffraction of the modulated light beam on a modulated surface acoustic wave (SAW) and the registration of this diffraction image. One of the main WAOU parameters is product $T_A\Delta f$, where $T_A = W_{opt}/V$ is the maximum delay time, which is known to be equivalent to the propagation time of the SAW leading edge through the optical beams, where W_{opt} is the optical beam width, V- the SAW velocity, and Δf is the working frequency band.

The x axis is directed along the propagation of the optical beam, the y axis is orthogonal to the waveguide plane and directed to its depth, and the z coordinate is perpendicular to the direction of propagation of the optical beam.

SAW operates as an optical driving transparency [4,5]. It executes following important operations:

- changing of deflection angle by means of change of SAW frequency (frequency of applied RF-signal) - this property is used for optical deflectors, scanners, switches;
- changing of intensity of diffracted beam by means of change of power of applied RF-signal - this property is used for amplitude modulation;
- changing of frequency of diffracted beam depended on changing of SAW frequency - frequency modulation.

SAW is generated with an electro-acoustical transducer [6]. The transducer type is usual inter digital transducer (IDT).

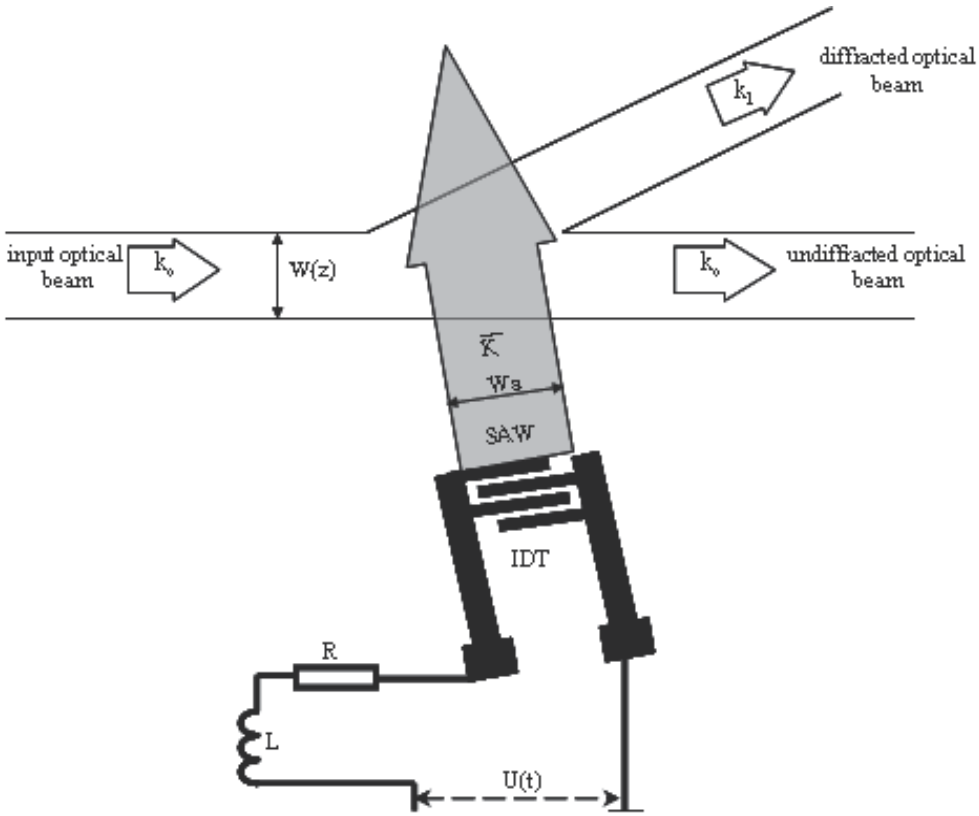


Fig. 1. Functional scheme WAOU, where k_0, k_1 – optical wave vectors, K – SAW wave vector, W(z) – optical aperture function, Wa – SAW aperture. Solid lines show optical beams. The source of optical radiation, lens system and registration unit are not shown.

The optical wave passes three areas. The optical wave aperture function W(z) is formed in first area. This means that the optical wave must be rather wide. There is AO diffraction in the second area. The SAW aperture W_a (is about hundreds μm) defines the AO diffraction region dimension along the optical beam. This area is situated in a diffused optical waveguide on the lithium niobate crystal surface for the device. There are two optical waves in this area: the undiffracted wave and the diffracted wave. The diffracted wave has the information product of the optical and acoustical signals. The third area realizes two functions: separation diffraction optical beam (from input optical beam) and the AO diffraction image transmission to a CCD array. The separation is obtained on the property that undiffracted and diffracted beams have the different propagation angles.

Main purpose. To increase the efficiency of the methods we must use and a proposed make it possible to process both a synthesized optical aperture and a synthesized acoustic aperture.

The problems of theoretical researches, mathematical models, simulation and experimental investigation of the based on Y-cut lithium niobate crystal AO units for real time data processing are discussed.

The propagation of plane optical beams is analyzed in the diffraction approximation with a negligible diffraction distortion (they are quasi parallel) and Fraungoffer approximation [7] is reasonable in this case. So, it is possible to use Fourier transformation for longitudinal optical fields. We suppose that only the TE_0 – optical mode exists in a planar optical waveguide. For such a waveguide, we can split the Maxwell equations and consider the time and coordinate components. The analysis of the coordinate components yields a one dimensional mode equation and a two-dimensional wave equation (with effective refractive index N_{eff}). For the *TE*-mode polarisation of guide optical waves the mode Maxwell equation become according to equation [8]:

$$\frac{d^2E}{d\rho^2} + (n^2(\rho) - N_{eff}^2)E = 0\,, \tag{1}$$

where $\rho=\kappa_0 y$, κ_0 is the optical wave number, E is optical field distribution normal to the boundary surface of the waveguide, $n(\rho)$ is function of refractive index profile (RIP) normal to the boundary surface of the waveguide. For titanium diffused waveguides the theoretical and experimental research leads us to a profile function for the refractive index given by [9]:

$$n^2(y) = n_s^2 + (n_0^2 - n_s^2)[(1-\alpha)e^{-\left(\frac{y}{a}\right)^2} + \alpha e^{-\frac{y}{b}}]\,, \tag{2}$$

where n_s is the substrate refractive index, n_0 - is the maximum the refractive index related to Ti-diffusion near the surface, a - is the effective depth of the Ti-profile, b - is the depth of the for out-diffusion, α - is the fraction of the exponential profile. This result was achieved by the optimization of the profile function comparing calculated and measured values of the effective refractive indices for multimode waveguides.

We study the propagation of light in the region of AO interaction assuming that the Bragg diffraction is realized, the properties of the waveguide mode remain unchanged, reflected waves are absent, the amplitude of the diffracted wave slowly increases (we can neglect the second derivation), and the perturbation of the waveguide permittivity caused by the SAW propagation is small [10,11]. The total diffraction losses are insignificant.

The diffraction efficiency of optical beam on SAW submits to the following parity [4,5]: $D = A^2(f)(\frac{\sin(qW_a)}{q})^2$, where $A^2(f) = \Gamma_{00}(f)P_{ac}$ - AO interaction coefficient, $\Gamma_{00}(f)$ - interaction integral for AO diffraction for TE_0- TE_0 mode regime (the interaction integral for Y-cut lithium niobate crystal was learned in [11, 12], and it's the frequency dependent is shown on fig. 2a), f – SAW frequency, $P_{ac} = U_{eff}^2 Y_{IDT}$, U_{eff} - IDT supply voltage, Y_{IDT} - IDT radiative transconductive, $q = (\frac{\Delta k}{2})^2 + A^2$, Δk - disagreement wave vector (see fig. 2b), W_a - ITD aperture.

With allowance for the phase mismatch Δk [3], the relation between the amplitude functions of the incident and undiffracted $T_0(x)$ and diffracted $T_1(x)$ optical waves is determined by the solution to the equations of bound modes [10]. Using these assumptions, we can analytically solve the wave equations and derive integral expressions [5].

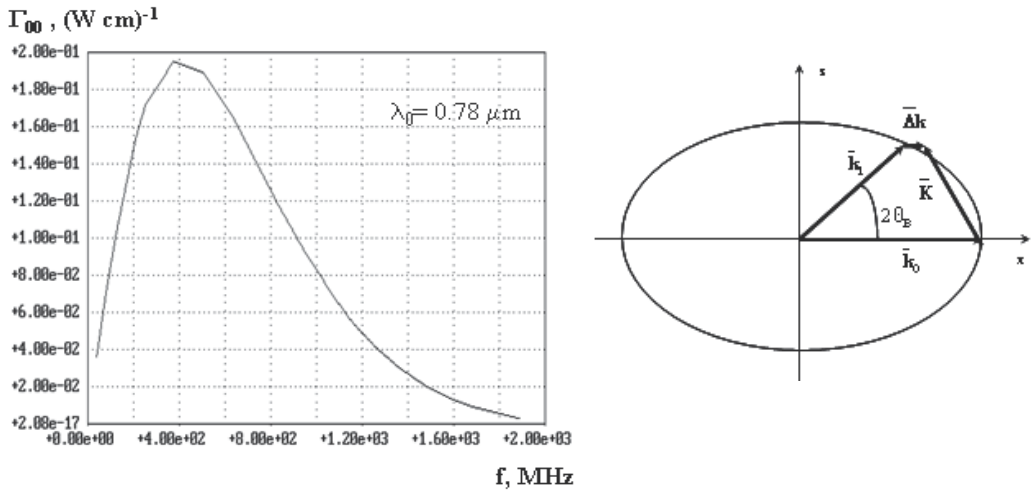


Fig. 2. (a) Plots of the frequency dependent of interaction integral for AO diffraction for TE_0-TE_0 mode regime, (b) The vector diagram of Bragg AO diffraction with disagreement optical wave vector

2. Real time multi-channel time-integrated correlation unit

2.1 Introduction

A classical time-integrated AO correlator (TIAOC) consists of a source of optical radiation (SOR), a collimating objective, an AO cell, an imaging objective, a spatial filter rejecting the null beam, and a CCD array [3, 13]. The main parameters of the AO correlator—the maximum delay time T_A. This parameter determines the range of delay times ΔT_A. The input signal $g_{in}(t)$ is used for time modulation of the light beam (for example, internal modulation of a laser or an LED). The second signal $h_{in}(t)$ modulates the amplitude of the acoustic wave with carrier frequency f_0. This wave is used for space–time modulation of the transmittance of the acousto-optic cell. The Bragg diffraction angle of the optical beam corresponds to the carrier frequency of the acoustic wave. Since the incident light intensity and the cell transmittance are proportional to $g_{in}(t)$ and $h_{in}(t - z/V)$, respectively, the intensity distribution of the diffracted beam depends on the product $g_{in}(t)h_{in}(t - z/V)$. The cell image is projected on the CCD array, which integrates the intensity distribution with respect to time. The signal from the CCD array yields the spatial distribution of the correlation functions of signals $g_{in}(t)$ and $h_{in}(t)$.

2.2 Mathematical model of a mutli-channal WAOU for time-integrated correlation data processing in real time

One method for increasing the integration of the correlation channels is multicolor data processing [14]. In this case, a single device contains N independent correlators. The correlation channels employ optical beams with different wavelengths. This approach facilitates parallel data processing in real time, diminishes the operating costs of an individual acousto-optic correlation channel, and makes it possible to avoid crosstalk and intermode losses.

For the multi-channel TIAOC whose scheme is shown in fig. 3, the input signal $g_{\text{in}}(t)$ represents a superposition of N independent optical signals corresponding to different wavelengths λ_m, where $m = 1...N$. Each optical beam is amplitude modulated. The second input signal $h_{\text{in}}(t)$ represents a superposition of N independent electric signals, each of which generates a SAW at the corresponding frequency f_{0m} and modulates its amplitude. The value of the carrier frequency is determined from the condition for the maximum in the overlap integral for the AO interaction of the SAW and the corresponding optical beam with a certain wavelength. Each optical beam is diffracted by the corresponding SAW in a waveguide AO Bragg cell (WBAOC). All cells are commoning the own chip. In the far-field region, the total diffracted optical field is a superposition of diffracted optical fields with different wavelengths. A prism is used in the focal plane of the imaging objective to spatially separate the optical beams. Thus, N independent correlation signals are simultaneously detected in the image plane in real time, which means that N independent correlation channels are realized in a single device.

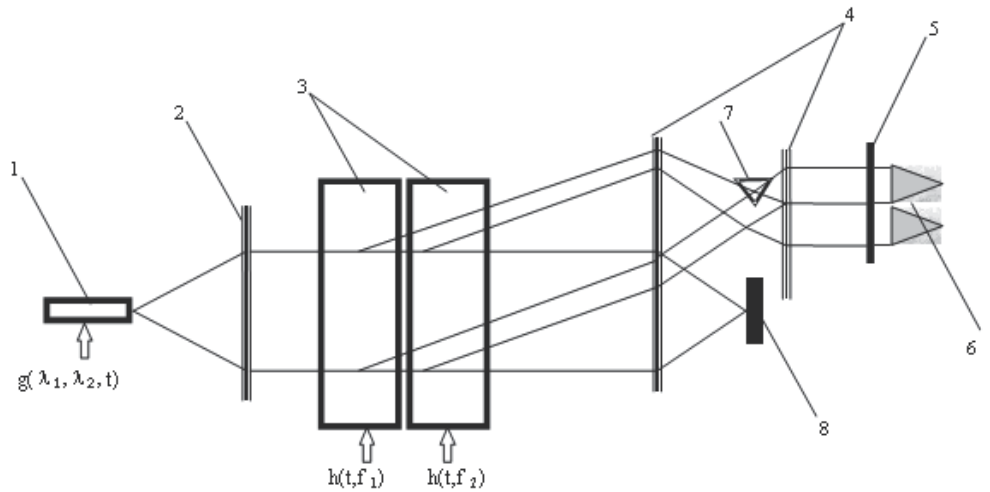


Fig. 3. Functional optical scheme of the multi-channel real time integrated TIAOC: 1 SOR, 2 collimating tens, 3 WBAOC, 4 image lens objective, 5 CCD array, 6 total correlation function 7 prism, 8 spatial filter. It's show two independent channel with deferent wavelengths). Solid lines show typical paths of optical beams in the device from the SOR to the CCD-array.

With regard to the above conditions, the diffracted field of the mth optical wave obeys the following equation:

$$\Delta T_{1_m}(x,z) + (\frac{2\pi}{\lambda_m} N_{eff_m})^2 T_{1_m}(x,z) = -\left[\int_{-\infty}^{\infty} dy k_{1_m}^2 E_{0_m}(y) E_{0_m}^*(y)\right] T_{0_m}(x,z) \tag{3}$$

where T_0 and T_1 are the longitudinal components of the incident and diffracted waves, respectively. Then, the distribution of the longitudinal component of the m-th diffracted optical wave is represented as

$$T_{1_m}(x,z,t) = g_{in_m}(t) \int_0^t d\tau \, R(z,t,\tau) h_{in_m}(\tau) ,$$

$$R(z,t,\tau) = ie^{-i2\pi f_{o_m} t} \int df e^{-i2\pi(f-fo_m)(t-\tau)} A(\lambda_m, f_{0_m} + f) r(z) e^{i\frac{2\pi}{\lambda_m} N_{eff_m} (Wa_m + Sin\frac{\Theta_m}{2} z)}, \quad (4)$$

$$r(z) = \int d\eta W_{opt}(\eta) \int d\rho \, e^{-i2\pi\rho(z-\eta)} Sinc \frac{\Delta k_m(\rho) W_{a_m}}{2} e^{i\frac{\Delta k_m(\rho) Wa_m}{2}},$$

where f is the modulation frequency.

The optical field outside the waveguide is a 2D field, whose transverse size is mainly determined by the quantity W_{opt} and the diffraction broadening. In the first approximation, we can neglect the diffraction broadening of the beam, since its angular divergence is about 0,1x10^{-2} rad. The vertical size of the field depends on the ratio λ / n_{sub}. For the TE_0 mode, the field distribution along the vertical axis determined by (1) is close to the Gaussian distribution. Therefore, all effects are concentrated in the transverse cross section of the optical field. This makes it possible to analyze the propagation of light in the correlator under study in the one-dimensional approximation.

For known parameters of the imaging objective, the intensity distribution for the mth correlation channel in the image plane is given by

$$J_m(z) = \frac{1}{\sqrt{2} T_{Int} R_{mag} \sigma} \int_0^{T_{\text{int}}} dt \int_{-\infty}^{\infty} d\xi I_{WBAOC_m} t_{in} t_{out} t_s e^{-(\frac{\xi + \Delta x - \frac{z}{R_{mag}}}{\sqrt{2}\sigma})^2}, \quad (5)$$

where R_{mag} is the magnification of the imaging objective, T_{Int} is time integrated, σ is the resolution of the imaging objective, and I_{WBAOC_m} is the integral intensity of the diffracted light for a m-th correlation channel, t_{in}, t_{out} are the translate coefficients of the prisms providing the chip incoupling/outcoupling of optical radiation, t_s is the translate coefficient of the prism providing the separation of optical radiation.

2.3 Computer simulation results

Using the model proposed, we theoretically analyze the values of the physical parameters of a hybrid five channel TIAOC. The WAOC consists of five WAOBCs that serve as a time optical transparency. The lens system of the correlator is made up of volume objective lenses [15]. For the numerical calculations, we use the following parameters. The effective depths of the refractive index profile are 2.24 and 6.52 μm and Δn^2 = 0.005. The SAW velocity is 3488 m/s. Using these parameters, we perform computer simulation to optimize the IDT structure for each correlation channel. The value of the rectangular optical aperture function is W_{opt} = 7 mm. The resulting field is a superposition of independent optical fields corresponding to individual correlation channels. The spatial separation depends on the optical properties and configuration of the prism. The aperture and the focal length of the imaging objective are 120 and 16 mm, respectively. The magnification is R_{mag} = 1/16. Both input modulating signals are rectangular signals of equal duration. The computer simulation results predict complete separation of the diffracted and undiffracted optical beams in the focal plane of the imaging objective.

At the given values of the physical parameters, we numerically study the characteristics of a hybrid five channel TIAOC. The table 1 shows the main parameters of this device. Figure 4

demonstrates the results of the computer simulation of the diffraction patterns in the correlator image plane for various durations of the input signals. In the case under consideration, the total correlation function consists of five parts.

Parameters/correlation channel	1	2	3	4	5
1. Optical radiation wavelength in vacuum, μm	1,06	0,92	0,88	0,83	0,78
2. Maximum delay time Ta, μs	2	2	2	2	2
3. IDT parameters					
3.1. Central carrier frequency f_0, MHz	251	277	321	380	436
3.2. Principal Bragg angle, deg.	1,0	1,1	1,2	1,3	1,5
3.3. Frequency bandwidth, MHz	14	15	16	16	18
3.4. Aperture, μm	380	400	400	420	450
3.5. Number of the fingers	53	51	49	47	41
3.5. Voltage standing-wave ratio in the working frequency range	1,2	1,2	1,2	1,2	1,2
4. Dynamic range, dB	25	25	25	25	25
5. Minimum power of the RF signal, mW	90	70	60	50	50
6. Minimum power of the optical signal, mkW	0,22	0,19	0,17	0,15	0,15

Table 1.

The maximum number of independent TIAOCs employing WAOCs based on the *Y*-cut lithium niobate substrate is 100.

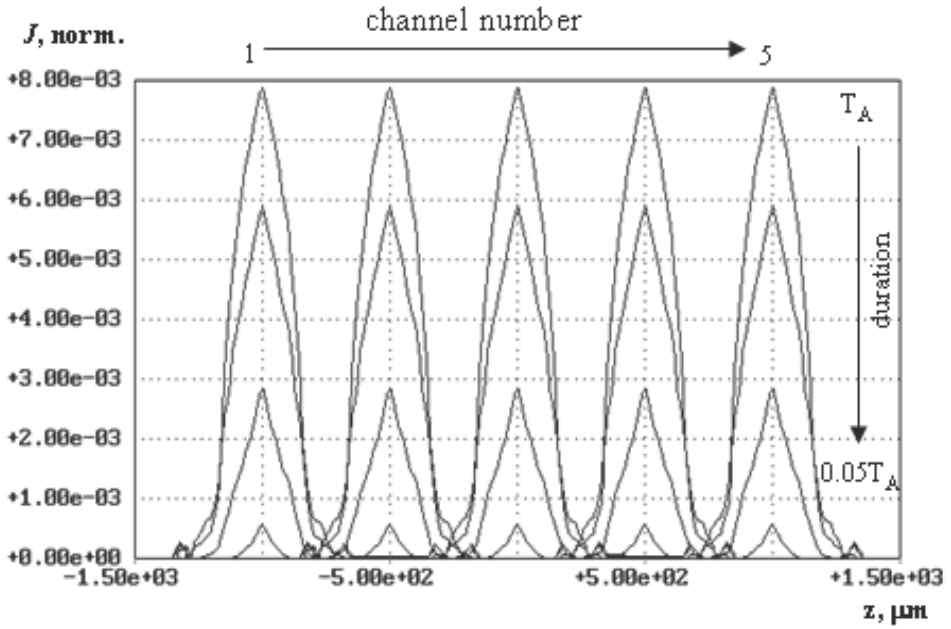


Fig. 4. Total correlation output signals for various durations of the input signals: from T_A to $0.05T_A$ (the delay time of $g_{in}(t)$ relative to $h_{in}(t)$ is $0.5T_A$). The families of curves correspond to the first (left-hand) to fives (right-hand) correlation channels, respectively

Using a mathematical model of the multi-channel TIAOC, one can also numerically analyze the characteristics of a hybrid monochrome TIAOC, in particular, at relatively small SAW variations. The validity of such an approach follows from the comparative analysis of the theoretical and experimental data. We employ an experimental prototype whose structure is described in detail in [16]. For the numerical calculations, we assume that, as in [16], WAOC contains a single WBAOC and two prisms providing the incoupling and outcoupling of optical radiation. The numerical experiments are performed for the wavelength in a vacuum λ_0 = 0.78 μm, the SAW velocity V = 3488 m/s, and an optical aperture function whose value is 7 mm. The remaining initial data correspond to the experimental sample from [16].
Figure 5a demonstrates the results of the computer simulation and experimental data on the dependences of the normalized correlation peak height on the duration of the input signals at their relative duration of 0.5 T_A. The experimental peak value is determined relative to the optical signal that is generated at the output of the device in the steady-state mode. The numerically calculated results can be presented in a similar way. Note also that the contributions related to additional biases applied to the optical source and WBAOC (pedestal contributions) are not taken into account in the results of the numerical calculations and experimental data.

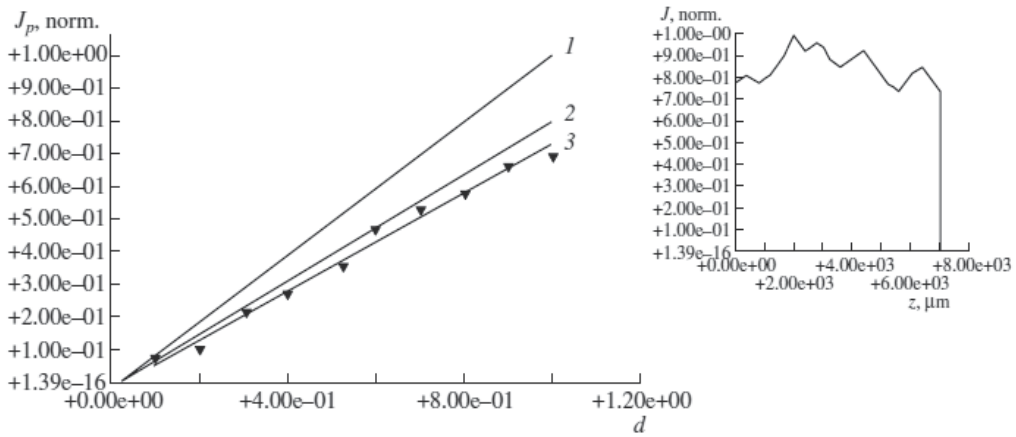


Fig. 5. (a) Plots of the normalized correlation peak height (Jp, norm) vs. duration (in fractions of T_A) of input signals at a relative delay of $0.5T_A$: (1) rectangular optical aperture function, (2) experimental optical aperture function, (3) experimental data from [16], and (triangles) experimental points. (b) Experimental optical aperture function.

Based on the comparative analysis of the numerical and experimental data, we draw the following conclusion. The proposed mathematical model can be used to analyze the affect of variations in the SAW velocity on the correlation signal, since both dependences are linear. However, the slope of the calculated curve differs from the slope of the experimental curve. This is due to the fact that the shape of the experimental aperture function significantly differs from the rectangular shape. The real shape is important in the calculation of the parameters of the given TIAOC. For comparison, Fig. 5a shows similar results calculated with allowance for the shape of the aperture function used in the experiments (Fig. 5b). In this case, the slopes are virtually identical.

2.4 Conclusion

We propose an AO method for time-integrated multi-channel correlation data processing in real time. This method enables one to significantly increase the number of independent correlation channels. We develop a mathematical model of a multicolor AO device for real-time correlation analysis with time integration. We theoretically and numerically analyze the characteristics of a hybrid five channel correlator, whose AO chip is based on the *Y*-cut lithium niobate substrate. Multi-channel AO correlators with time integration make it possible to simultaneously detect up to 100 correlation functions provided that *Y*-cut lithium niobate serves as the substrate.

3. Bandwidth AO unit for the real-time spectrum analyzis of optical signals

3.1 Introduction

The design and development original AO method for the real-time spectral analysis of broadband optical signals are discussed. The unit implemented this mean allows for the real-time processing of an optical signal that exhibits a complicated spectrum with hundreds of components. This unit may also serve as a high-accuracy real-time optical-frequency detector working in a wide spectral range.

Figure 6 shows that the classical AO unit consists of three components: optical port (*1*), waveguide AO chip (*2*), and photodetector unit (*3*). The optical signal from the first unit is fed to the WAOC, where it is collimated with an aplanatic lens (*4*) to a wide-aperture plane wave (*5*). Then, the wave propagates through the region of the perturbed planar waveguide (*6*), which is induced by the propagating surface acoustic wave (*7*) generated by an electroacoustic IDT (*8*). In this case, the Rayleigh SAW with strictly fixed frequency f_0 serves as an optical transparency. Then, a component of the optical signal is deflected due to Bragg diffraction by the angle [10]

$$\Theta_m = \frac{\lambda_m}{V} f_0, \tag{6}$$

where λ_m is the radiation wavelength of the m-th component in the waveguide.

The number of the diffracted optical beams corresponds to the number of the frequency components of the original optical signal. The resulting diffraction field is focused with an aplanatic lens (*9*) on the end of the waveguide and is detected with a photodetector array (*3*).

The mathematical model of the waveguide AO spectrum analyzer of the optical radiation is similar to the well-developed and experimentally tested model of the waveguide AO spectrum analyzer of radio signals [16]. The differences are as follows: first, the parameter under study is the radiation wavelength rather than the radio frequency (which is fixed); second, the chromatic dependence of the overlap integral of the Bragg AO interaction needs to be taken into account.

The optical wave diffracts on SAW. Thus the diffraction efficiency of an each optical component submits to the following parity:

$$D = A^2(\lambda_m, f_0)\left(\frac{\sin(q_m W_a)}{q_m}\right)^2, \tag{7}$$

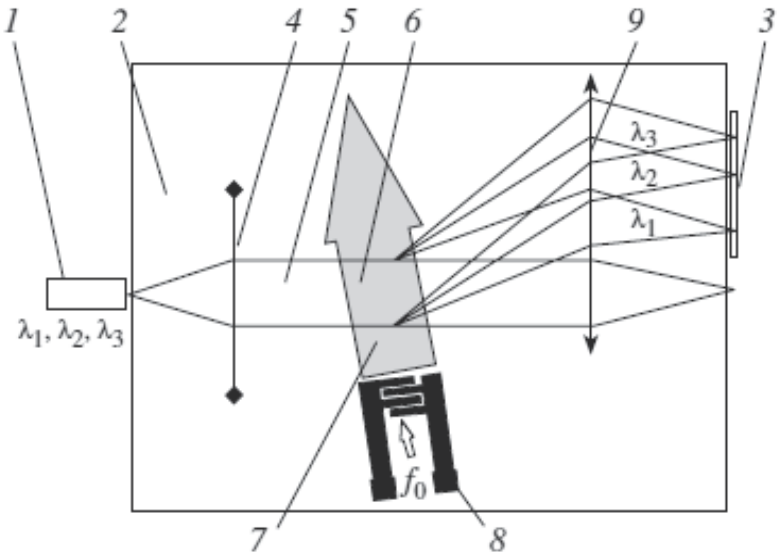


Fig. 6. Scheme of the waveguide Bragg AO device for the real-time spectral analysis of the broadband optical signals: solid lines show optical beams.

where $A^2(\lambda_m, f_0) = \Gamma_{00}(\lambda_m) P_{ac}$ - AO interaction coefficient for m-th optical component, $\Gamma_{00}(\lambda_m)$ - chromatic interaction integral for AO diffraction for TE_0- TE_0 mode regime, $q_m = \left(\frac{\Delta k_m}{2}\right)^2 + A^2(\lambda_m, f_0)$, Δk_m - disagreement wave vector for corresponding optical component.

The frequency of a everyone diffracted optical beam will be moved on frequency f_o. And the quantity of diffracted optical beams corresponds to quantity of frequency components of an initial signal.

In general, the band of the working optical frequencies is bounded by the SAW excitation band, the AO interaction band, the frequency band of the overlap integral, which characterizes the AO properties of the material, and the band of the optical frequencies in which the planar waveguide supports a single TE_0 mode. The last two parameters are significantly more important than the first two parameters. A single IDT is insufficient for the analysis of the broadband optical signals, since the band of the analyzed optical frequencies is relatively narrow (about 100 nm) [4, 5]. In contrast to the monochromatic case, the characteristic is strongly asymmetric, since the overlap integral decreases with increasing wavelength at relatively large $\Delta\lambda$. For this reason, a multiple-IDT system is used to increase the width of the frequency band. In this system, the crossed structure of acoustic beams enables one to maximize the diffraction efficiency. Several IDTs that excite SAW are placed at the Bragg angle relative to the propagation direction of the channeled radiation. The complexity of this scheme lies in the manifestation of the interference effects for the optical beams diffracted by different SAW beams. This leads to a significant nonuniformity of the frequency characteristic. To solve this problem, we need to employ computer simulation for the amplitude and phase tuning of the device (both the topological parameters of the cell and the electric circuit for the matching of the IDT and RF systems are tuned). It is demonstrated in [5] that the amplitude and phase matching is stable only for three IDT's.

3.2 Mathematical model of AO units for the real-time spectrum aanalysis of bandwidth optical signals

In case greatest possible value (~300 nm) of a bandwidth of an optical signal (such signals can propagated in waveguide at only TE_0-mode regime) use 3 ITD obviously insufficiently. One of ways of increase of a optical frequency bandwidth is entering an additional mismatch of wave vectors with the purpose of increase of a total optical frequency bandwidth, but as the consequence is observed decrease of diffraction efficiency. This approach is realised by change of value of a angle between an optical beam and SAW front, that is achieved additional turn of the ITD. Thus should be satisfied condition:

$$\frac{\sin(q_i W_{a_i})}{q_i} = \left[\frac{\pi}{2}(4n+1)\right]^{1/2}, \tag{8}$$

where n=0,1,2,3.

Thus the diffraction efficiency of a each optical component submits to a parity (7). The regime of low diffraction efficiency allows to simplify parities describing optical fields:

$$\begin{aligned} T_0 &= const \\ T_1 &= -i\frac{A(\lambda_m)T_0}{\Delta k}e^{-i(\Delta k x+\Phi(z))}, \end{aligned} \tag{9}$$

where T_0 - longitudinal initial optical field component, T_1- longitudinal diffracted optical field component, $\Phi(z)$ - phase of longitudinal diffracted optical field component [4,5].

The optical field each wave components is the sum an optical field generated appropriate SAW, extending under different angles to an initial optical beam. Presence several coherent the optical field results in them interference. Then the optical field separate wave components under condition of low efficiency diffraction submits to the following parity:

$$T_{1_m}(x) = -i\frac{A(\lambda_m)T(x)e^{-i(\Delta k_1 x+\Phi_1(z))}}{\Delta k_1}\left(1+\sum_{j=2}^{N}\frac{\Delta k_1}{\Delta k_j}e^{-i((\Delta k_j-\Delta k_1)x+(\Phi_j(z)-\Phi_1(z)+\Phi_{opt})}\right). \tag{10}$$

The given field by means of an integrating lens is transferred on a ruler of photoreceivers placed in a lens focal plane. The intensity distribution of each wave component in a focal plane of an integrating lens is described by parity [17]:

$$\begin{aligned} W_{F_m}(x_0,z) &= \int_{-\infty}^{\infty} d\xi \frac{(1+\phi)}{\sqrt{s}} T_{1_m}(\xi,\eta)e^{-i(\Phi_2(z)+k_o N_{eff}(\psi)s\cos(\psi))} \\ s &= \sqrt{(x_0-\eta)^2+(z-\xi)^2} \end{aligned}, \tag{11}$$

where s - the distance from the point on last lens contour to the observation point, ϕ - the angle between the ray coming to the point ξ to point (x_0,z), ψ - the angle between the phase and group velocity for the ray directed towards the observation point, $\Phi_2(z)$ - the phase of optical ray at the point ξ.

The total optical field is superposition of fields a each optical component.

An increase in the aperture function leads to an additional decrease in the sensitivity of the frequency characteristic to the interference effects. This is due to the averaging of the phases

of the diffracted beams. It follows from the analysis of the phase dependences of the interfering beams that the interference effects are completely eliminated provided that the optical aperture satisfies the following condition:

$$W_{opt} >> \frac{\Delta\lambda}{4\Delta\Theta^2} \tag{12}$$

where $\Delta\lambda$ is the wavelength band of the analyzed radiation and $\Delta\Theta$ is the difference between the Bragg angles of the neighboring IDT's. For the typical parameters $\Delta\lambda$ = 100 nm and $\Delta\Theta$ =0.0017, condition (12) is satisfied for an optical aperture of about 100 mm. However, such large apertures cannot be technically realized in the integrated optical devices, and the aforementioned property provides for only partial compensation of the interference effects.

3.3 Computer simulation results

We develop two modifications of the AO devices for the spectral analysis of the red and IR broadband optical signals. In these devices are based on *Y*-cut lithium niobate. The refractive index profile of diffusion planar optical waveguide is described the following parameters. For the first modification, the effective depths are 2.24 and 6.52 mkm and Δn^2= 0.005. For the second modification, the effective depths are 2.42 and 6.12 mkm and Δn^2 = 0.008. TIPE technology [18] is used to fabricate aplanatic lenses. We assume that the optical aperture function is nearly Gaussian and its values for the first and second modifications are *W*opt = 5 and 7 mm, respectively. The SAW velocity is 3488 m/s. Based on these parameters, we employ computer simulation to optimize (i) the IDT-system design with respect to the maximization of the band of the analyzed optical frequencies and (ii) the topology of the lens system with respect to the maximization of the optical resolution using the method proposed in [19]. The table 2 demonstrates the main parameters of the two modifications obtained with the numerical and experimental study.

Parameters	Mod. 1		Mod. 2
	simul.	exper. [16]	simul.
1 Optical radiation waveguide range			
- low wavelength, nm,	600	605	765
- high wavelength, nm,	935	915	1125
2 Working wavelength band, nm	335	315	360
3 Resolution of two optical wavelengths, nm,	2.9	3.5	3.1
4 Measuring fixed frequency for RF signal, MHz	394	396	326
5 Diffraction efficiency, %/W	1.5	0.8	0.3
6 Dynamic range, dB	25	23	25
7 Low level of measured optical signal, μW	10	10	10
8 Input RF signal power level, W	0.4	0.6	1.2
9 Voltage standing wave ratio for driving RF signal	1,8	2	1.8

Table 2.

Figure 7a shows the numerically calculated dependence of the optical transmittance of the first modification on the wavelength. At a level of –3 dB, the working wavelength band is slightly greater than 335 nm in the range from 935 to 600 nm. For comparison, we also present the experimental characteristic of this AO device from [16]. Figure 7b shows the wavelength dependence of the calculated optical transmittance for the second modification. At a level of –3 dB, the working wavelength band is 360 nm in the range from 1125 to 765 nm. Note that the parameters of the device are chosen with allowance for the technological requirements for its fabrication. A working band of 360 nm is the ultimate band for the single mode diffusion waveguides under study [14]. The optical signals with such maximum bandwidth can be supported with the waveguides under study in the IR range

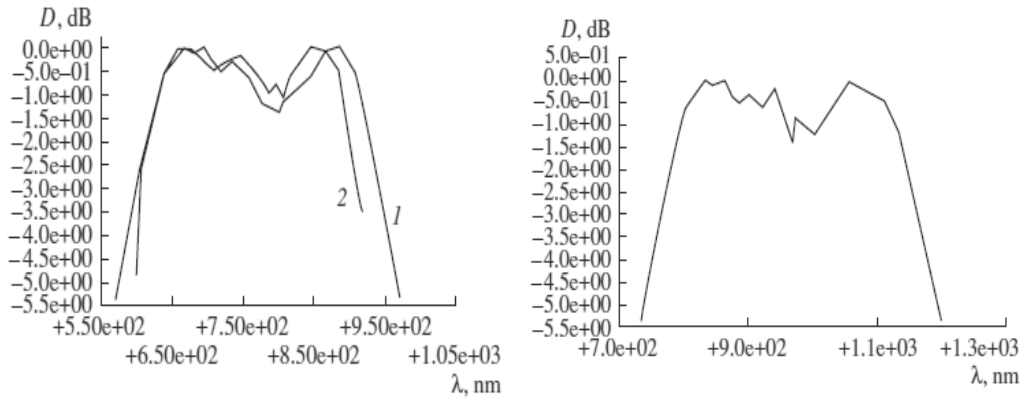


Fig. 7. Plots of optical transmittance *D* vs. wavelength for (a) the first modification of the spectrum analyzer, (b) for the second modification of the spectrum analyzer.

The comparative analysis yields a good agreement between the theoretical and experimental data on the first modification. In particular, the difference between the RF signals is less than 1%, the difference between the central optical frequencies is almost 10%, and the difference between the working transmission bands is less than 10%.

Figure 8 demonstrates the experimentally determined accuracy of the measurement of the optical signal wavelength with the first modification in which a laser diode serves as the source of the optical signal.

The left panel shows the typical response of the device. At a fixed frequency of the RF signal and a stable radiation of the laser diode, the response is symmetric (Fig. 8a). A variation in the input current of the laser diode leads to a variation in the radiation wavelength, which causes a variation in the response (Fig. 8b). In this case, the voltage difference between the *N*th and (*N* + 1)th pixels is 0.2 V. The response can be symmetrized by tuning the oscillator frequency. The difference between the oscillator frequencies is 38 kHz. Thus, the wavelength of the laser diode is varied by 0.054 nm. In practice, we can observe and measure a voltage difference of about 10 mV between the (*N* – 1)th and (*N* +1)th pixels. Hence, the ultimate accuracy of the wavelength measurement for the optical signal in the device under study is no worse than 0.01 nm.

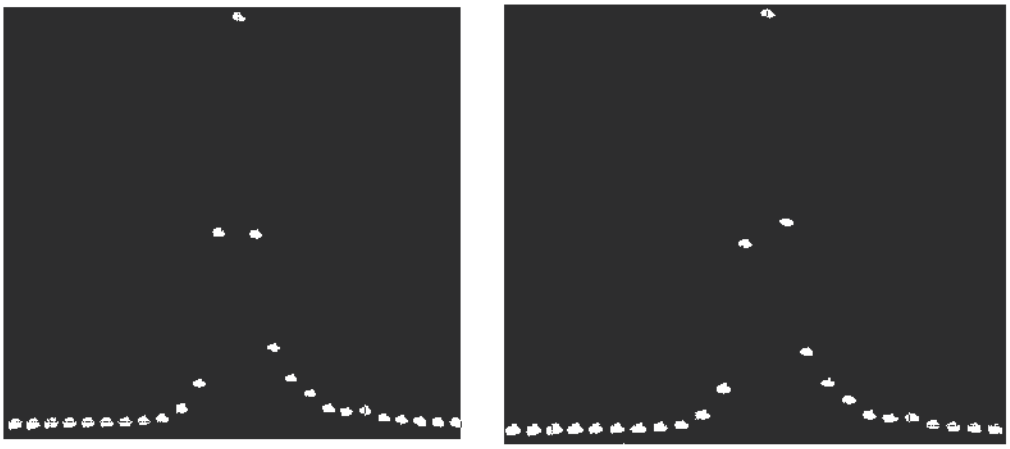


Fig. 8. Photographs of the responses for the first modification of the AO device.

3.4 Conclusion

We developed and numerically studied two modifications of the AO device for the spectral analysis of broadband red and IR optical signals. We employ a mathematical model which is similar to the experimentally tested mathematical model of the waveguide AO analyzer of radio signals. The comparative analysis yields a good agreement between the theoretical and experimental results. The band of the working optical frequencies is 360 nm in the range 765–1125 nm. The wavelength resolution of the two optical signals determined with the Rayleigh criterion is 3 nm. The optical wavelength is measured with an accuracy of 0.01 nm. The minimum level of the analyzed optical signal is 10 mkW. In particular, the differences between the RF signals, the central optical frequencies, and the working transmission bands are less than 1%, almost 10%, and less than 10%, respectively. The parameters of the device are chosen based on the real technological and topological requirements for its elements.

4. Methods for real-time optical scanning of larde data arrays

4.1 Introduction

A distinctive feature of the methods proposed is a significant increase in the number of scanning points (information bits) in comparison to the classical method involving spectral analysis of broadband RF signals based on waveguide AO Bragg interaction [4, 16]. In a conventional scanning system, the frequency of the controlling RF signal fed to a scanning unit is discretely varied within the frequency band of the device. A variation in frequency leads to a variation in the diffraction angle of the optical beam [10]:

$$\Theta_{dif} = \Theta_o + \arcsin\left(\frac{\lambda_o f}{N_{eff} V} - \sin\Theta_a\right) \tag{13}$$

where Θ_o is the angle between the fronts of the optical beam and surface acoustic wave, and Θ_a is the angle between the SAW propagation direction and the z axis.

A variation in the diffraction angle causes a shift of the focused optical beam to the neighboring point in the focal plane coinciding with the side surface of the waveguide. The

scanning is thus realized. The discrete step in the frequency variation depends on the frequency resolution of the optical focusing system [4]. The number of scanning points is given by

$$N_T = \frac{\Delta f}{R_f} \tag{14}$$

where Δf is the frequency band and R_f is the resolution of the focusing optical system.

The hardware for realizing the above method employs a source of optical radiation and a waveguide acoustooptic chip. The chip contains a planar optical waveguide, optical input/output elements, an optical system consisting of collimating and focusing lenses, and a waveguide AO Bragg cell. Note that for scanner elements based on *Y*-cut lithium niobate, the number of scanning points is no greater than 1000 for the ultimate values of Δf and R_f [4]. However, the technical realization of such a device is difficult and expensive [16].

4.2 Method for one-dimensional scanning

A conventional method for 1D scanning can be improved to achieve a significant increase in the amount of data processed in real time using an approach that employs a synthesized (with a few optical beams) aperture of the optical field. The total information field is divided into subareas, each of which is irradiated with an individual optical beam controlled by the corresponding WAOBC. In the case under consideration, the first optical beam is incident on the corresponding cell at an angle $\Theta_{\max}$ equal to a double Bragg angle corresponding to the upper bound limit of the cell working frequency range. The next optical beam is incident on the second WAOBC at an angle $\Theta_{\max} + \Delta\Theta_c$, where $\Delta\Theta_c$ is the working range of angles of the previous cell. In the general case, the angle of incidence of the current optical beam differs from the previous one by $\Delta\Theta_c$ (Fig. 9). Thus, the scanning range increases by a factor of *M*, where *M* is the number of cells.

A similar WAOBC system is placed symmetrically with respect to the optical axis. Note the matching of the working range of angles of the focusing system (Θ_L) and the total working range of the WAOBC system. Hence, the number *M* satisfies the following condition:

$$M < \frac{\Theta_L}{2\Delta\Theta_c} \tag{15}$$

Thus, the total number of scanning points is given by

$$N_{T_{total}} = 2MN_T \tag{16}$$

We investigate the propagation of a monochromatic coherent quasi-plane optical wave with a predetermined aperture function *W*(*z*). We assume that the AO Bragg phase matching condition [5] is satisfied, the diffraction efficiency is about a few percent. In the general case, the aperture of the surface acoustic field is synthesized with a few electroacoustic IDTs.

In the far-field zone, the electrical component of optical field diffracted by a each WAOBC is represented as

$$T_1(x,z,t) = T_0(x)\int_{-\infty}^{\infty} dk_z \sum_{j=1}^{N} \frac{A_B(f)}{\Delta k(x,z)} \left(e^{i\Delta k W_{a_j}} - 1\right) e^{-i\left((\omega\pm\Omega)(t-T_a) - N_{eff}\left(\sqrt{k_o^2 - k_z^2}\right)x + \Phi_B(x,z,f) + Kz\right)} e^{-\alpha(f)z} \tag{17}$$

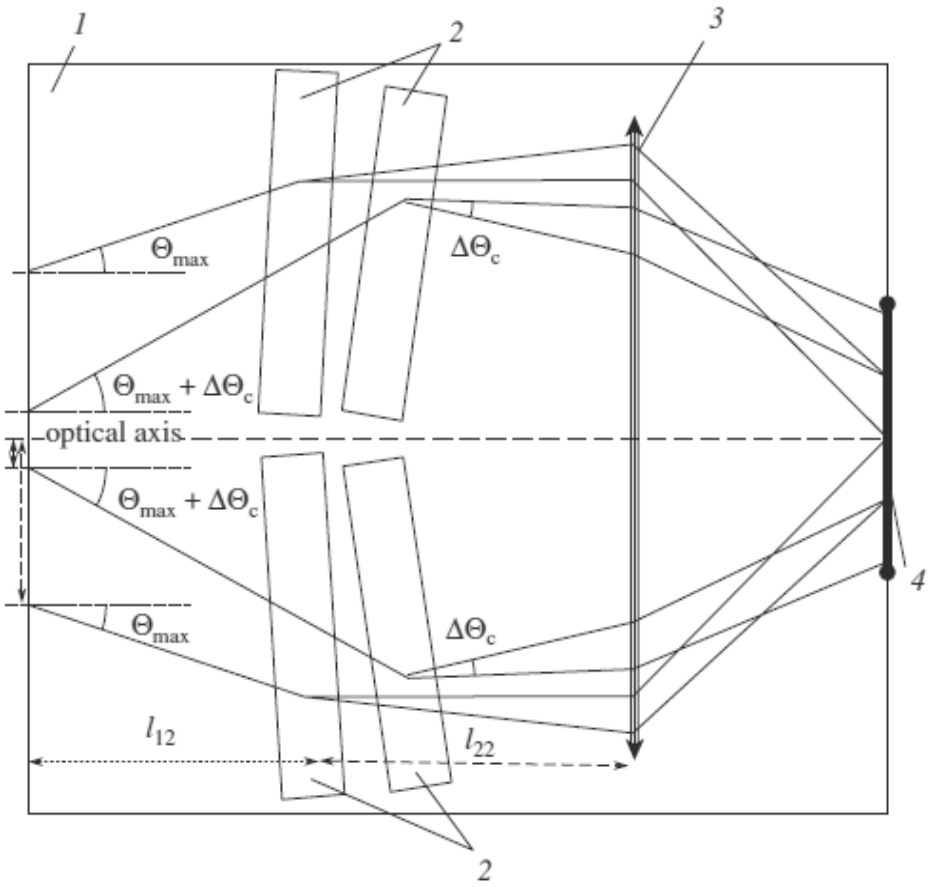


Fig. 9. Functional scheme of the waveguide AO scanning device: *1* WAOC, *2* WAOBC, *3* focusing lens objective, and *4*total information field. Solid lines show typical paths of optical beams in the device from the SOR to the scanning region. The SOR and collimating lens system are not shown.

where ϖ and k_0 are the frequency and wave number of the TE_0 mode; Ω and K are the frequency and wave number of the SAW; $A_B(f)$ is the Bragg AO interaction coefficient corresponding to parameter $A(f)$, determined in sec. 1., N is the number of IDTs; Wa is the j-th IDT aperture; $\Phi_B(x,z,f)$ is the phase related to the RF electric signal passing through the matching electric circuit, SAW propagation from the electroacoustic transducer to the region of acoustooptic interaction, and interference depending on the acoustic field topology [5, 11]; and $\alpha(f)$ is the SAW decay coefficient, which was analyzed in detail in [6]. This field is focused using a waveguide lens objective with focal length F. In the back focal plane of the objective, the field distribution is represented as

$$W_F(x_F,z)=\int_0^{T_{Sc}} d\tau \int_{-\infty}^{\infty} d\xi \frac{(1+\phi)}{\sqrt{s}} T_1(\xi,\eta,\tau) e^{-i(\Phi_2(s)+k_o N_{eff}(\psi)s\cos(\psi))} \tag{18}$$

where T_{Sc} is the scanning time.

4.3 Computer simulation results

Using the above approach, we developed a AO planar scanning device based on *Y*-cut lithium niobate. This device makes it possible to selectively process more than 3000 information bits in real time. The device consists of four WAOBCs (two cells on each side of the optical axis). It is possible to arbitrarily control the cells. Using the model proposed, we numerically study the characteristics of this scanner working in real time. In numerical experiments, we employ the following typical initial data presented in the table 3. The aperture function of the optical beam is nearly Gaussian, and the corresponding value is *W*opt = 5 mm.

Parameter	
Planar waveguide	
Effective refractive index for TE_0 mode	2.19176
Difference between maximum refractive indices of the waveguide and substrate	0.8×10^{-2}
Depth of the Gaussian profile, μm	2.42
WAOBC parameters	
Total working frequency range at a level of 3 dB, MHz	1090
Range of diffraction angles, rad	0.111
Maximum diffraction efficiency, %/W	0.97
Number of sections in the first IDT	17
Aperture of the first IDT, μm	450
Number of sections in the second IDT	23
Aperture of the second IDT, μm	280
Electric power of the RF signal, W	0.6
Voltage standing-wave ratio in the working frequency range	less than 2
Parameters of the optical lens system	
Difference between maximum refractive indices of the TIPE lens and substrate	
Effective depth of the TIPE lens, μm	0,0328
Focal length of the collimating system, mm	4.24
Number of lenses of the focusing objective	2
Focal length of the focusing objective, mm	35
Aperture of the focusing objective, mm	15.8
Size of the scanning optical spot, mkm	0.22
Main characteristics of the WAOC	
Number of WAOBCs	4
Aperture of the optical beam, mm	5
Angle between the optical axis and the optical beam for the first WAOBC, rad	0.130
Angle between the optical axis and the optical beam for the second WAOBC, rad	0.240
Total number of scanning points	4360

Table 3.

Based on these parameters, we employ computer simulation to optimize the IDT system of the WAOBC. The system consists of two fan IDTs [20] spanning the following frequency ranges: 230–720 and 720–1320 MHz. Figure 10 shows the frequency characteristic of the WAOBC. The elements transforming the aperture function of the optical beam are manufactured using TIPE technology [19]. The topology and parameters of the lens system are determined by (i) the ratio of effective refractive indices outside and inside the lens with regard to the aberration of optical beams with large angular aperture, (ii) condition (15), (iii) the optical beam aperture, and (iv) the size of the substrate. With allowance for all these characteristics and a substrate length of 80 mm, we use computer simulation to optimize the topological parameters of the scanning device. The table lists the results of optimization. The maximum allowed frequency resolution R_f of Gaussian optical beams with a 5-mm aperture equals 1 MHz [19]. Then, the number of points scanned in real time using the device under consideration is 4360. The size of the scanning optical spot (0.22 μm) is close to the diffraction limit.

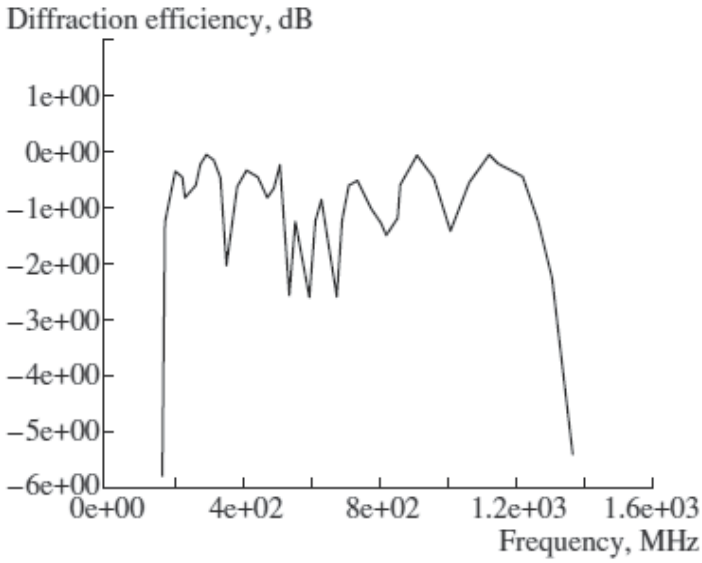


Fig. 10. Frequency characteristic of the WAOBC

4.4 Method for 2D scanning

We can realize 2D optical scanning simultaneously using two types of waveguide AO diffraction (Bragg (coplanar) and radiative (collinear)) [21]. Figure 11 gives a functional scheme of the 2D scanning device taking into account the aforementioned results.

In the case under consideration, the waveguide mode of TE_0 polarization propagating in the WAOC from the corresponding SOR to the scanning plane is transformed in the following manner. First, the collimating system (not shown in Fig. 11) forms a wide-aperture optical beam. Then, the beam propagating through the WAOBC is diffracted by a SAW under Bragg conditions. The diffracted optical beam is deflected from the initial direction by the double Bragg angle (see expression (13)). In the case of radiative diffraction, this beam is deflected into the substrate owing to the interaction with a collinearly propagating SAW. The angle at which the optical beam is re-emitted into the substrate is given by [10]

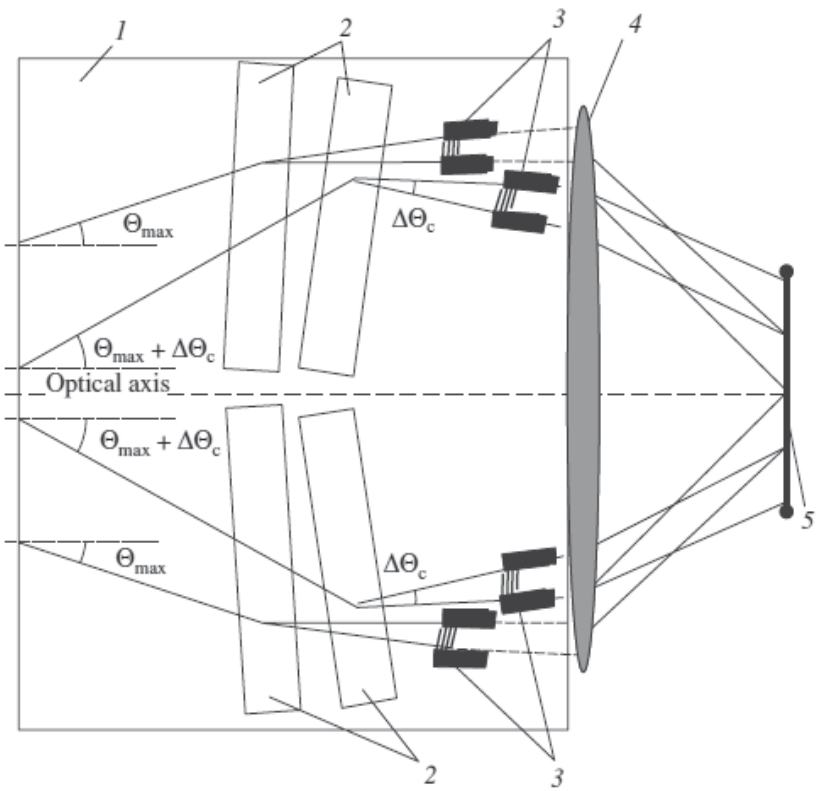


Fig. 11. Functional scheme of the AO scanning device for processing of 2D data arrays: (*1*) WAOC, (*2*) WAOBC, (*3*) IDT, (*4*) volume focusing lens objective, and (*5*) total information field. Solid lines show typical paths of optical beams in the waveguide and to the scanning plane. Dashed lines show the trajectories of optical beams re-emitted into the substrate.

$$\Theta_{\mathrm{rd}} = \arccos \frac{N_{eff} - \dfrac{\lambda_0 f}{V}}{n_{sub}}, \tag{19}$$

where Θ_{rd} is the angle between the guided mode and the mode of radiation and n_{sub} is the refractive index of the substrate. Then, separation of the optical beams takes place. A doubly diffracted optical beam is projected onto the scanning plane using a volume lens objective. Non diffracted and singly diffracted optical beams are rejected with a spatial filter. A variation in the working frequencies of the IDT system placed at a Bragg angle relative to the *z* axis and IDT placed at a mean Bragg angle relative to the *x* axis leads to 2D motion of the doubly diffracted optical beam on the screen. We analyze the propagation of light in a single mode planar waveguide with simultaneous application of two types of AO interaction using approximations similar to those mentioned above. Then, the diffraction field of a forward optical wave at a certain cross section is given by the following system of equations:

$$\frac{dT_0(x)}{dx} = iA_B T_1(x) e^{-i(\Delta k_B x - \Phi_B)},$$

$$\frac{dT_1(x)}{dx} = iA_B T_0(x) e^{i(\Delta k_b x - \Phi_B)}$$

$$\frac{dT_1(x)}{dx} = i\frac{1}{(1-2\beta)x} A_{rd} T_2(x) e^{-(i\Delta k_{rd} x + \alpha)} \tag{20}$$

$$\frac{dT_2(x)}{dx} = i\frac{1}{(1-2\beta)x} A_{rd} T_1(x) e^{(i\Delta k_{rd} x - \alpha)}$$

where $T_2(x)$ is the amplitude function of a doubly diffracted optical wave, α is the sum of decay coefficients of the optical beam and the collinearly propagating SAW, β is the diffraction divergence coefficient of the collinearly propagating SAW, A_{rd} is the efficiency of radiative AO interaction, which similar to coefficient $A(f)$, where the parameter Γ_{00} is replace of the factor Γ_{rd}, and Γ_{rd} is the overlap integral for TE_0–TMsub diffraction.
For an arbitrary RIP of a waveguide, the solution to this system of equations can be found only using numerical methods under the following initial conditions: $T_0(0) = 1$ and $T_1(0) = T_2(0) = 0$. Note that, in contrast to the Bragg case, the frequency characteristic of radiative diffraction exhibits a strong nonuniformity owing to interference effects related to the finiteness of the spectrum of the radiation mode.

4.5 Computer simulation of characteristic of the AO unit for 2D scanning

In the part of the approach proposed, we developed an AO unit for 2D scanning whose WAOC is based on *Y*-cut lithium niobate. For numerical experiments, we choose similar initial data. The optical aperture function is W_{opt} = 3 mm. We employ the same WAOBC. Based on these parameters, we use computer simulation to optimize an IDT generating a SAW that propagates collinearly to the optical beam. The parameters of the SAW are as follows: central frequency, 571 MHz; frequency range, 523–635 MHz; IDT aperture, 3100 μm; metallization coefficient, 0.51; and voltage standing-wave ratio, no greater than 2 in the entire frequency range of the IDT. Based on the deflection of each diffracted optical beam, the angles of IDTs relative to the *x* axis are -0.24, -0.13, 0.13, and 0.24 rad. Figure 12 shows the results of computer simulation of the resulting frequency characteristic for a single 2D AO scanning channel at a radiative AO interaction length L_{a-o} = 2.5 cm. In the case under consideration, for an electric power of the control RF signals of 0.5 W, the maximum diffraction efficiency is 0.3%. In accordance with the Rayleigh criterion, the number of resolved states is 650 points on the vertical axis and 650 points on the horizontal axis. The total information field is then 2600 x650 points (information bits).

4.6 Conclusion

We analyze original methods for real-time optical scanning of large data arrays based on waveguide AO Bragg interaction with a synthesized optical aperture using a few optical beams. The methods proposed make it possible to process both one- and 2D data arrays. A distinctive feature of the methods is an increase in the number of scanning points in comparison to the classical method involving spectral analysis of broadband RF signals based on waveguide AO Bragg interaction. We developed a mathematical model of a waveguide AO scanning device with a large number of based on *Y*-cut lithium niobate, the

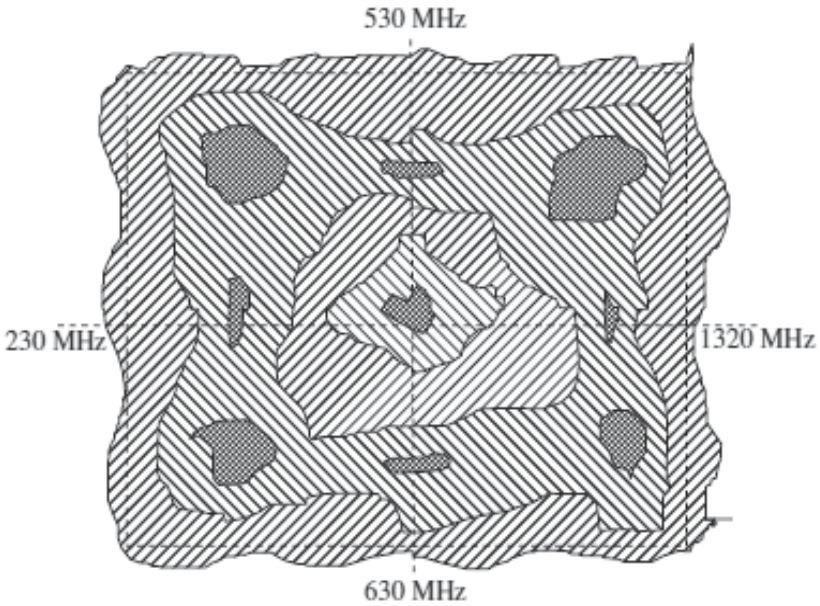


Fig. 12. Frequency characteristic of a single 2D scanning channel: right, left, and cross hatching correspond to the ranges - 2-3, - 1-2, - 0–1 dB, respectively.

number of points for real-time 1D scanning is 4360. There exist two factors accounting for the practical importance of the mathematical model proposed in this work. First, the mathematical model is based on the experimentally tested mathematical model of a waveguide AO spectrum analyzer. Second, we choose the parameters of the device based on real technological and topological requirements for its elements. Note that all the elements of the AO scanning device (except for SOR) can be realized for a single substrate.

Using the method for 1D scanning, we developed a mathematical model of a WAOU for real-time 2D scanning. We perform computer simulation of its characteristics for *Y*-cut lithium niobate. The number of points for real-time scanning is 650 on the vertical axis and 2600 on the horizontal axis. The parameters of the IDT generating a collinear SAW are calculated under the same conditions as the parameters of the IDT in a WAOBC. The aforesaid facts make it possible to conclude that the model proposed can be applied in practice.

5. Saw velocity fluctuation effects

5.1 Introduction

In the development of waveguide AO devices for data processing, one must analyze the accuracy level in the reproduction of the predetermined characteristics. For example, if an AO device is designed based on a certain value of the surface acoustic wave velocity V_0 and a real value of velocity V_1 differs from this value owing to several factors, the characteristics are changed. In practice, multiple mechanisms lead to a variation in the physical properties of the substrate material, which, in turn, causes variations in the SAW velocity [6]. The most typical mechanisms are related to the ambient temperature, the ionization effect, and the spread of velocities due to errors in the crystal orientation. Note that we consider the case in

which the external action does not result in mechanical variations in the components of devices.

We numerically study the characteristics of two basic AO devices for real-time information processing: time-integrating correlator and broadband spectrum analyzer (AOSA) of radio signals. Both devices are based on a waveguide AO chip that is fabricated on a *Y*-cut lithium niobate substrate.

5.2 Mathematic model

We analyze the scenarios in which the deviation of the SAW velocity from the calculated value is small, so that the relative variation in the velocity is several percent. In this approximation, we assume that, in the range under study, the SAW velocity exhibits a linear dependence on a certain factor. This assumption is in agreement with the experimental data from [1]. Then, such a dependence can be represented as [22]

$$V_1 = \chi V_0 = (1 - \frac{\Delta V}{V_0})V_0, \tag{21}$$

where $\chi = 1 - \frac{\Delta V}{V_0}$, and $\Delta V = V_0 - V_1$ is a variation in the SAW velocity. Parameter χ can be physically interpreted as a coefficient of the SAW velocity fluctuations. Note that, in the case under study, this coefficient is constant.

Morgan proposes studying the effect of a minor degradation of the SAW velocity on the characteristics of the device using the model of a linear source and a linear receiver located at a free crystal surface [22]. He also assumes that the diffraction, propagation loss, and SAW dispersion can be neglected in the analysis. For distance l between the source and the receiver, the calculated time delay is $t_0 = l/V_0$, whereas $t_1 = l/V_1$ is the time delay that corresponds to a certain level of the effect of one of the degradation mechanisms. The relationship between these delays is given by

$$t_1 = \frac{1}{\chi} t_0, \tag{22}$$

This causes a variation in the time scale of the pulse response of the electroacoustic IDT, which is described with functions $h_0(t)$ and $h_1(t)$ at the SAW velocities V_0 and V_1, respectively. The relationship between the pulse responses is represented as $h_1(t) = h_0(\chi t)$. We assume that the pulse responses correspond to the short-circuiting mode, so that the effect of the external action on the matching circuits is eliminated. It is demonstrated in [22] with neglect of minor variations in the amplitude that the frequency response functions of the device exhibit frequency rescaling with a factor of $1/\chi$. Thus, parameter $1/\chi$ can be physically interpreted as the scaling factor. In the range under study, coefficient remains unchanged, so that parameter $1/\chi$ is also constant.

A variation in the SAW velocity results in a variation in the characteristics of the AO diffraction. For the isotropic Bragg diffraction, a relatively small diffraction angle Θ_1 is given by

$$\Theta_1 = \frac{\lambda_0 f_0}{\chi N_{eff} V_0} = \frac{1}{\chi}\Theta_0, \tag{23}$$

Hence, the diffraction angle is scaled as $1/\chi$ in the range under study. Therefore, first, the spectrum of spatial frequency is corrected in the same way and, second, an additional mismatch of the wavevectors of the incident and diffracted waves emerges. This leads to a shift of the SAW frequency f_0, at which the AO diffraction is realized with the maximum efficiency. Then, a new value of the frequency is written as $f_1 = \chi f_0$. Note that the maximum diffraction efficiency remains unchanged provided that a variation in the overlap integral is neglected.

A minor variation in the diffraction angle affects separation Δs of the incident and diffracted beams in the focal plane of a lens. In the approximation of geometrical optics, the following expression is valid: $\Delta s_1 = \frac{1}{\chi}\Delta s_0$, where $\Delta s_1, \Delta s_0$ are the separations with regard to and with neglect of the degradation, respectively.

The working frequency band Δf of the AO device is determined by the following main factors: the SAW excitation band and the AO interaction band [5]. For the isotropic Bragg diffraction, the working frequency band is given by the following expression taking into account the spread of the SAW velocities:

$$\Delta f_1 = \frac{2N_{eff}\chi^2 V_0^2 (1 - \Theta_1/4)}{\lambda_0 f_c W_a} \approx \chi^2 \Delta f_0, \tag{24}$$

where f_c is the central frequency. Neglecting the second term in parentheses in the numerator, we conclude that Δf quadratically depends on the SAW velocity fluctuation coefficient.

5.3 Analysis of the TIAOC characteristics

A variation in the time scale directly affects one of the main parameters of the AO correlator—the maximum delay time T_A. In the presence of the SAW velocity fluctuations, this parameter is given by

$$T_{A_1} = \frac{W_{opt}}{\chi V_0} = \frac{1}{\chi} T_{A_0}, \tag{25}$$

where W_{opt} is the optical beam aperture. Note that the scaling laws for ΔT_A and T_A are identical.

A relative delay τ of the AO correlator is scaled similarly to T_A. One must take into account the time of the SAW propagation from IDT to the AO interaction region, which is similarly scaled:

$$\tau_1 = \frac{\gamma T_{A_0}}{\chi} + \frac{L}{\chi V_0} = \frac{1}{\chi}\tau_0, \tag{26}$$

where γ is a fraction of T_A and L is the distance between IDT and the AO interaction region.

Based on the analysis of expressions (25) and (26), we draw the following conclusion. In spite of the scaling of parameters T_A, ΔT_A and τ, the range of the relative delays $\Delta\tau$ of the correlator is the same as in the absence of the SAW velocity spread, so that quantity γ is not scaled.

The scaling of the working frequency band leads to a variation in the AOTIC base. At a constant integration time, it can be estimated as

$$N_1 = \chi^2 N_0, \tag{27}$$

where N_1 and N_0 are the AO correlator bases with regard to and with neglect of the SAW velocity spread, respectively.

Consider the effect of a variation in the SAW velocity on the transformation of the correlation signal. For numerical experiments, we employ a mathematical model of the multi-channel TIAOC. Using a mathematical model of the multi-channel TIAOC, one can also numerically analyze the characteristics of a hybrid monochrome TIAOC, in particular, at relatively small SAW variations. The validity of such an approach follows from the comparative analysis of the theoretical and experimental data. We employ an experimental prototype whose structure is described in detail in [16]. The numerical experiments are performed for the wavelength in a vacuum λ_0 = 0.78 μm, the initial SAW velocity V_0 = 3488 m/s, and a rectangular optical aperture function whose value is 7 mm.

To estimate the variation in the shape of the correlation signal in the presence of the SAW velocity fluctuations, we analyze a model scenario in the absence of propagation loss. Such an approach makes it possible to account for the transformations of the correlation signal. Let information signals with equal durations (T_A) have a relative delay of 0.5 T_A. In this case, the SAW velocity can differ from the calculated value. For the numerical experiments, we choose the same initial data as in the previous case. The optical aperture function and all of the signals applied to the device have rectangular shapes.

Figure 13 demonstrates the results of the computer simulation. At the first stage, when the SAW velocity coincides with the calculated velocity, a classical correlation signal in the image plane represents an almost equilateral triangle. When the SAW velocity decreases, the shape of this signal is varied. In the case under study, a simultaneous shrinkage of the left- and right hand wings of the correlation signals leads to its narrowing. This is due to the spatial compression of the acoustic signal, which is proportional to a variation in the SAW velocity. However, in all of the transformations, the correlation peak amplitude remains unchanged. The reason for this lies in the dependence of this amplitude on the integration time, which remains constant (T_A). The following conclusion can be drawn from the simulation results. The expression that can be used to estimate the narrowing of the correlation signal in the above range of a decrease in the SAW velocity is written as

$$W_{cor_1} = \chi W_{cor_0}, \tag{28}$$

where W_{cor_1} and W_{cor_0} are the apertures (widths) of the correlation signal with regard to and neglect of a decrease in the SAW velocity, respectively.

An increase in the SAW velocity leads to a substantially different transformation of the correlation signal: it becomes asymmetric. In the case under study, a decrease in the left-hand wing is accompanied by an increase in the right-hand wing. This is due to the spatial extension of the acoustic signal, which is proportional to the variation in the SAW velocity.

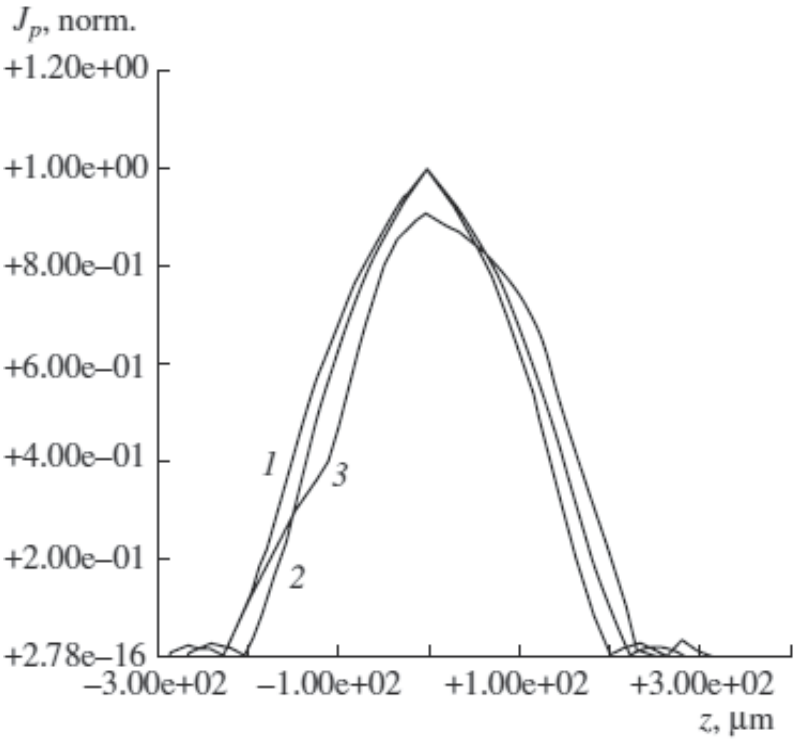


Fig. 13. Normalized correlation signal (*J*, norm) in the TIAOC image plane with regard to the SAW velocity fluctuations for an input signal duration of T_A and a relative delay of 0.5*T*A at χ = (*1*) 1.2, (*2*) 0.9, and (*3*) 1.1

The maximum amplitude of the correlation peak linearly decreases with an increasing parameter χ.

A similar analysis can be performed for other pulse durations and the corresponding integration times. It follows from the calculated results that a decrease in the pulse duration causes transformations of the correlation signal that are similar to the above transformations.

In a real device, we must take into account the variation in the delay in the acoustic path. As mentioned, this leads to a variation in the relative delay of the signals. To eliminate the effect of the acoustic delay, we consider a general case in which the relative delay is varied in the entire allowed range 0– T_A. It follows from [23] that the correlation peak height exhibits an approximately parabolic decrease in the absence of the SAW velocity degradation. Using a similar approach, we generalize the simulated results on the case of the spread of the SAW velocities. Figure 14 shows the results of the analysis. It is seen that a decrease in the SAW velocity causes a symmetric decrease in the correlation peak. Opposite results are obtained when the SAW velocity increases. A sharper decrease in the correlation

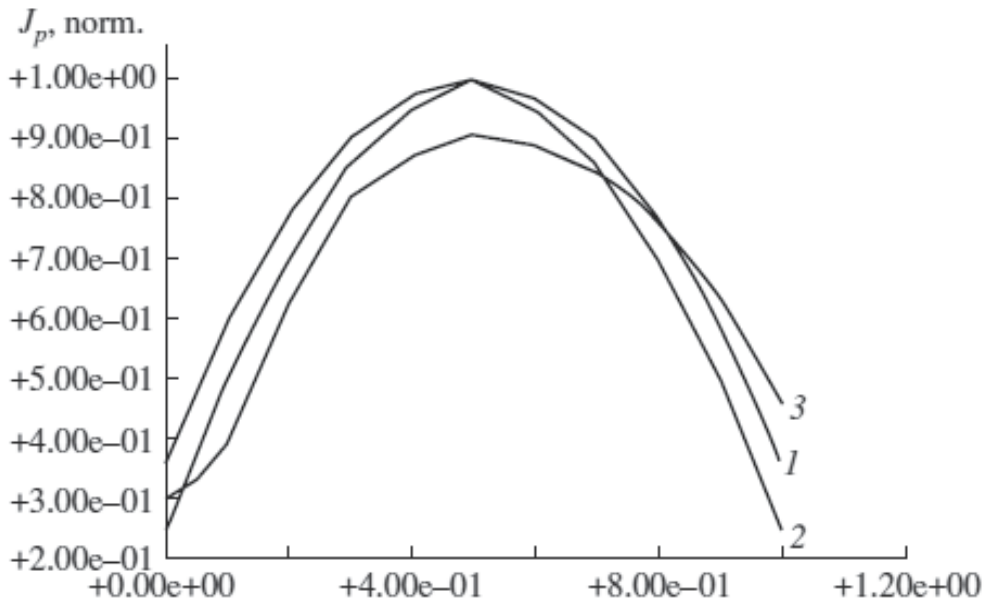


Fig. 14. Plots of the normalized correlation peak height (Jp, norm) vs. relative delay γ (in fractions of T_A) for input pulses with a duration of T_A with regard to the SAW velocity fluctuations at χ = (*1*) 1.2, (*2*) 0.9, and (*3*) 1.1

peak height is observed at $\gamma < 0.5\ T_A$, whereas the peak height starts increasing (in comparison with the ideal scenario) at $\gamma > 0.5\ T_A$. Note that, at $\gamma > 0.8\ T_A$, the peak height becomes higher than that in the ideal case. Such a character of this dependence is realized at any variations in the SAW velocity in the range under study.

5.4 Analysis of the AOSA characteristics

We analyze the manifestation of a minor variation in the SAW velocity in the output characteristics of the broadband AOSA. For this purpose, we employ the experimentally verified mathematical model of this device from [5]. By analogy with the above analysis, we choose the experimental device whose structure is described in detail in [16]. For the numerical calculations, we assume that the WAOC contains WBAOC with a complicated ITD system, integrating lens, and incoupling and outcoupling optical prisms. The remaining initial data correspond to the experimental device from [16]. The aplanatic lenses for the WAOC are manufactured using the TIPE technology [19]. We assume that the optical aperture function exhibits a nearly Gaussian shape with the width W_{opt} = 5 mm. Using these parameters, we perform computer simulation to optimize the interdigitated structure, which is needed to realize the maximum band of the analyzed radio frequencies, and to optimize the topology of the WAOC lens system, which is needed for the maximum optical resolution with the technique from [5].

Note several features of the realization of the broadband AO diffraction by the synthesized acoustic field. These are related to the differences of the overlap integral, interference phenomena, and acoustic loss at low and high frequencies. To estimate the variation in the working frequency band related to the SAW velocity fluctuations, we analyze two WBAOC modifications. Each cell contains three IDT structures that generate an acoustic field with crossed beams. The first cell is constructed for the frequency range 0.3–0.7 GHz. The second cell works in the frequency range 0.7–1.3 GHz. The experimental and theoretical parameters and frequency characteristics of the cells can be found in [23].

Figure 15 shows the simulated diffraction characteristics of both cells in the presence of the SAW velocity fluctuations: plots of the normalized working frequency band (at a level of –3 dB) and the diffraction efficiency at the critical frequency (nonuniformity of the frequency response), where the highest sensitivity to the SAW velocity variations is realized, vs. parameter χ. For the first and second cells, the critical frequencies are 567 and 1010 MHz, respectively. The analysis of the simulated results makes it possible to reach the following conclusions. The scaling of the working frequency band in the range of relatively small variations in the SAW velocity linearly depends on parameter χ. The slope of this curve for the high-frequency cell *2* is higher than that for the low-frequency cell *1*. An increase in the nonuniformity of the frequency response is related to a violation of the phase relationships [5]. Note its developed frequency dependence. To determine the level of the effect of the mechanism under study at which the allowed level (e.g., –3 dB) is exceeded, a detailed simulation is needed in each specific case.

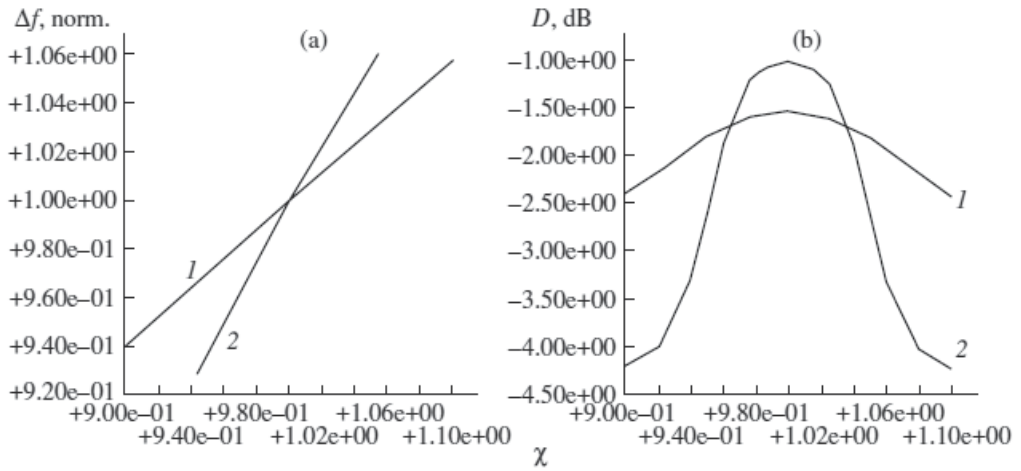


Fig. 15. Plots of (a) the normalized working frequency band (Δ*f*, norm) and (b) diffraction efficiency *D* at the critical frequency vs. SAW velocity fluctuation coefficient for (*1*) the first and (2) second cells.

The SAW velocity fluctuation also causes a variation in the WAOC frequency resolution. It follows from the numerically calculated distributions of the optical intensity at the focal plane of the integrating lens that, for the processing of two RF signals with allowance for a 10% variation in the SAW velocity, the resolution is virtually inversely proportional to parameter χ (Fig. 16). Note that, in the ideal case, the frequency resolution is 2 MHz.

Based on the simulated results for the working frequency band and the resolution, we conclude that the number of points of the broadband AOSA remains almost unchanged in the low-frequency range in the presence of SAW velocity fluctuations.

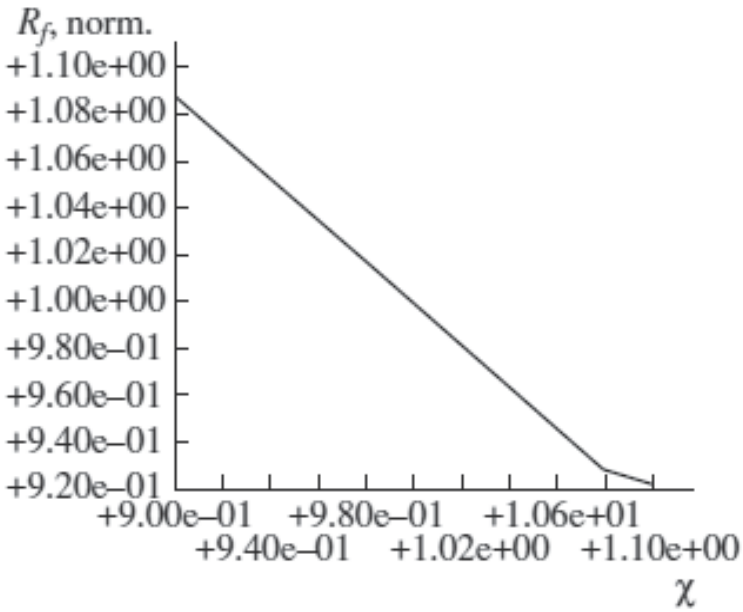


Fig. 16. Plot of the AOSA normalized frequency resolution (*Rf*, norm) vs. SAW velocity fluctuation coefficient

5.5 Conclusion

The characteristics of two basic AO devices (time-integrating correlator and broadband spectrum analyzer of radio signals) are theoretically analyzed and numerically studied in the presence of SAW velocity fluctuations. The devices are based on a waveguide AO chip that is fabricated at a *Y*-cut lithium niobate substrate. In the analyzed range of the SAW velocity fluctuations, the SAW velocity linearly depends on the level of the degradation factor. This causes the scaling of the characteristics of the device. Several parameters (Bragg diffraction angle, separation of optical beams, maximum delay time, range of delay times, relative delay, and frequency resolution) are scaled with a coefficient that is equal to the inverse relative variation in the SAW velocity. For the narrowband and broadband scenarios, the working frequency band exhibits quadratic and linear scalings, respectively.

The correlator base is quadratically scaled, and the number of points of the spectrum analyzer remains almost unchanged. The correlation peak height exhibits complicated variations in the presence of the SAW velocity fluctuations. The range of the correlator relative delays is not scaled.

6. Conclusions

The newly developed methods for real time data processing are proposed. To increase the efficiency of the methods we must use and a proposed make it possible to process both a synthesized optical aperture and a synthesized acoustic aperture. The mathematical models of the units whose acousto-optic chip is based on the *Y*-cut lithium niobate substrate are developed. We theoretically and numerically analyzed the unit characteristics under design and technology conditions. The results obtained can be used for the practical elaboration of the high efficiency AO devices.

7. Acknowledgments

I am grateful to M.Yu. Kvasha for creation of the experimental sample and S.M. Zakharov and M.Yu. Kvasha for helpful discussion.

8. References

[1] A. Korpel, Proc. IEEE 69, 48 (1981)
[2] *Integrated Optics*, Ed. by T. Tamir (Springer-Verlag, New York, 1975)
[3] A. VanderLugt, *Optical Signal Processing* (Wiley, New York, 1992).
[4] C. S. Tsai, J. Mod. Opt. 35, 965 (1988)
[5] N. V. Masalsky, Proc. SPIE 3737, 125 (1999)
[6] *Acoustic Surface Waves*, Ed. by A. A. Oliner (Springer-Verlag, New York, 1978)
[7] Born M., Wolf E., *Principles of optics* (Pergamon press, Oxford - London - Edinburgh - New York - Paris – Frankfurt,1968)
[8] Epikhin E.N., Jendges R., Masalsky N.V. Arbeitspapiere der GMD 888, p. 1-14, 1994.
[9] E.N. Epikhin, M.Y. Kvasha, N.V. Masalsky, N.V. Praschikin, V.A. Volkov, K.-L. Paap, Photonics and Optoelectronics, vol.2, pp. 137 - 148, 1994
[10] R. V. Schmidt, IEEE Trans. Sonics Ultrason. 23, 22 (1976)
[11] M. A. Alhaider, L. T. Nguyen, B. Kim, and C. S. Tsai, Proc. IEEE 64, 318 (1976)
[12] E.G. Lean, , J.M. White, C.D.W. Wilkinson, Proc. IEEE 64, p.775 (1976)
[13] W. T. Rhodes, Proc. IEEE 69, 65 (1981)
[14] N. V. Masalsky, Laser Phys. 14, 882 (2004)
[15] A. S. Jensen, K.-L. Paap, B. Klaassen, *et al.*, Proc. SPIE 2969, 507 (1996)
[16] A. G. Sobolev, V. A. Volkov, E. N. Epikhin, et al., Proc. SPIE 3714, 170 (1999)
[17] N. V. Masalsky, Laser Phys. 16, 1352 (2006)
[18] C.S. Tsai, D.Y. Zang, Appl. Opt., 25, 2264, (1986)
[19] A. L. Belostotsky, A. S. Leonov, and D. V. Petrov, Opt. Comm. 2, 9 (1992)
[20] D. Gregoris and V. Ristic, J. Mod. Opt. 35, 979 (1988)
[21] E. M. Korablev and V. V. Proklov, Photonics Optoelectronics. 1, 7 (1993)

[22] D. P. Morgan, *Surface-Wave Devices for Signal Processing* (Elsevier, Amsterdam, 1985)
[23] E. N. Epikhin, M. Yu. Kvasha, N. V. Masalsky, et al., Proc. SPIE 3900, 242 (1999)

12

Aluminium Nitride thin Film Acoustic Wave Device for Microfluidic and Biosensing Applications

Y.Q. Fu[1], J. S. Cherng[2], J. K. Luo[3], M.P.Y. Desmulliez[1], Y. Li[4], A. J. Walton[4] and F. Placido[5]

[1]*School of Engineering and Physical Sciences, Institute of Integrated Systems, Heriot-Watt University, Edinburgh, EH14 4AS,*

[2]*Department of Materials Engineering, Mingchi University of Technology, Taishan, Taipei,*

[3]*Centre for Material Research and Innovation, University of Bolton, Deane Road, Bolton, BL3 5AB,*

[4] *Scottish Microelectronics Centre, School of Engineering, Institute of Integrated Systems, University of Edinburgh, Edinburgh, EH10 7AT,*

[5]*Thin Film Centre, University of the West of Scotland, Paisley, PA1 2BE,*

[1,3,4,5]*UK*

[2]*Taiwan*

1. Introduction

When an alternating electric field is applied to an interdigitated transducer (IDT) on a piezoelectric material, an acoustic wave is generated. The wave can propagate in a direction perpendicular to the surface of the material into the bulk (bulk acoustic wave, BAW) or along the surface of the material (surface acoustic wave, SAW). This piezoelectric effect is manifested in either a Rayleigh mode (vertical and surface normal) or as a shear horizontal wave (in-plane) [Galipeau et al 1997]. The most commonly used bulk acoustic wave device is the Quartz Crystal Microbalance (QCM), which is generally made of quartz sandwiched between two electrodes. In contrast a surface acoustic wave propagating within a thin surface layer, which has a lower acoustic velocity than that of the piezoelectric substrate, is called a Love wave and such devices are typically operated in the Shear Horizontal (SH) wave mode. Waves propagating in a thin plate with a thickness much less than the acoustic wavelength are called a flexural plate or Lamb waves [Luginbuhl et al 1997]. These acoustic wave technologies and devices have been commercially exploited for more than 60 years in industrial applications [Ballantine et al 1996. Hoummady et al., 1997] and currently the telecommunications industry is one of the largest consumers, primarily in mobile phones and base stations, which account for ~3 billion acoustic wave filters annually. Other promising and growing applications include automotive applications (pressure acceleration, or shock sensors), medical applications (chemical sensors), and other industrial applications (including temperature, mass, viscosity, vapour and humidity sensors).

Most acoustic wave devices can be used as sensors because they are sensitive to mechanical, chemical, or electrical perturbations on the surface of the device [Lucklum & P. Hauptmann 2003, Grate et al 2003]. Acoustic wave sensors have the advantage that they are versatile, sensitive and reliable, being able to detect not only mass/density changes, but also viscosity, wave functions, elastic modulus, conductivity and dielectric properties. They have many applications in monitoring a large number of parameters which include pressure, moisture, temperature, force, acceleration, shock, viscosity, flow, pH, ionic contaminants, odour, radiation and electric fields [Shiokawa & Kondoh 2004, Wohltjen et al. 1997]. Recently, there has been an increasing interest in acoustic wave based biosensors to detect traces of biomolecules through specific bioreactions with biomarkers. These include DNA, proteins (enzymes, antibodies, and receptors), cells (microorganisms, animal and plant cells, cancer cells etc.), tissues, viruses, as well as the detection of chemical substances through specific chemical absorption layers [Cote et al 2003, Kuznestsova, and Coakley 2007, Teles & Fonseca 2003]. By detecting traces of associated molecules, it is possible to diagnose diseases and genetic disorders, prevent potential bioattachment, and monitor the spread of viruses and pandemics [Vellekoop 1998, Shiokawa & Kondoh 2004, Gizeli 1997]. Compared with other common bio-sensing technologies, such as surface plasmon resonance (SPR), optical fibres, and sensors based on field effect transistors or cantilever-based detectors, acoustic wave based technologies have the combined advantages of simple operation, high sensitivity, small size and low cost, with no need for bulky optical detection systems [Lange et al 2008].

By far the most commonly reported acoustic wave based biosensor is QCM [Markx, 2003], which can be operated in a liquid environment using a thickness shear-mode. The advantages of QCM include: (1) simplicity in design and (2) a high Q factor. However, less attractive features of QCM biosensors are a low detection resolution due to the low operating frequency in the range of 5~20 MHz and a large base mass; a thick substrate (0.5~1 mm) and large surface area (>1 cm^2) which cannot easily be scaled down. In contract SAW based biosensors have their acoustic energy confined within a region about one wave length from the surface, and so the basemass of the active layer is roughly one order of magnitude smaller than that of the QCM. Therefore, the sensitivity of the SAW devices is dramatically larger than that of the QCM. The longitudinal or Rayleigh mode SAW device has a substantial surface-normal displacement that rapidly dissipates the acoustic wave energy into the liquid, leading to excessive damping, and hence poor sensitivity and noise. However, waves in a SH-SAW device propagate in a shear horizontal mode, and therefore do not easily radiate acoustic energy into the liquid [Barie & Rapp 2001, Kovacs & Venema 1992] and hence the device maintains a high sensitivity in liquids. Consequently SH-SAW devices are particularly well suitable for bio-detection, especially for "real-time" monitoring. In most cases, Love wave devices operate in the SH wave mode with the acoustic energy trapped within a thin waveguide layer (typically sub-micron). This enhances the detection sensitivity by more than two orders of magnitude compared with a conventional SAW device owing to their much reduced base mass [Josse et al 2001, Mchale 2003]. They are therefore frequently employed to perform biosensing in liquid conditions [Lindner 2008, Kovacs et al 1992, Jacoby & Vellekoop 1997].

Acoustic wave technologies are also particularly well suited to mixing and pumping and as a result are an attractive option for microfluidics applications [Luo et al 2009]. Taking the SAW device as one example, Rayleigh-based SAW waves have a longitudinal component that can be coupled with a medium in contact with the surface of the device. When liquid

(either in bulk or droplet form) exists on the surface of a SAW device, the energy and momentum of the acoustic wave are coupled into the fluid with a Rayleigh angle, following Snell's law of refraction (see Fig. 1) [Wixforth 2004, Shiokawa et al 1989]. The Rayleigh angle, θ, is defined by

$$\theta = \sin^{-1}\left(\frac{v_l}{v_S}\right) \tag{2}$$

where v_l and v_s are the velocities of the longitudinal wave in solid and liquid. The generated acoustic pressure can create significant acoustic streaming in a liquid which can be used to enable liquid mixing, pumping, ejection and atomization [Newton et al 1999]. This pressure facilitates rapid liquid movement and also internal agitation, which can be used to speed up biochemical reactions, minimize non-specific bio-binding, and accelerate hybridization reactions in protein and DNA analysis which are routinely used in proteomics and genomics [Toegl et al 2003, Wixforth et al 2004]. Surface acoustic wave based liquid pumps and mixers [Tseng et al 2006, Sritharan et al 2006], droplet positioning and manipulation [Sano et al 1998], droplet ejection and atomization systems [Chono et al 2004, Murochi et al 2007], and fluidic dispenser arrays [Strobl et al 2004] have been proposed and developed. They have distinct advantages, such as a simple device structure, no moving-parts, electronic control, high speed, programmability, manufacturability, remote control, compactness and high frequency response [Renaudin et al 2006, Togle et al 2004, Franke & Wixforth 2008].

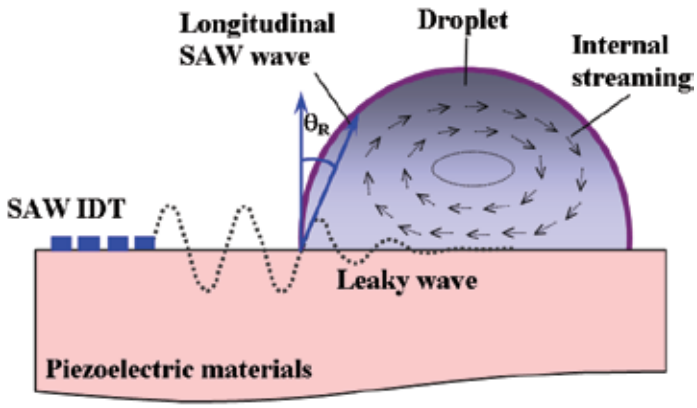


Fig. 1. Principle of surface acoustic wave streaming effect: interaction between propagating surface acoustic wave and a liquid droplet causing acoustic streaming inside droplet

Acoustic wave devices can be used for both biosensing and microfluidics applications, which are two of the major components for lab-on-a-chip systems. Therefore, it is attractive to develop lab-on-chip bio-detection platforms using acoustic wave devices as this integrates the functions of microdroplet transportation, mixing and bio-detection. To date, most of the acoustic devices have been made from bulk piezoelectric materials, such as quartz (SiO_2), lithium tantalate ($LiTaO_3$), lithium niobate ($LiNbO_3$) and sapphire (Al_2O_3). These bulk materials are expensive, and are less easily integrated with electronics for control and signal processing. Piezoelectric thin films such as PZT, ZnO and AlN have good

piezoelectric properties, high electro-mechanical coupling coefficient, high sensitivity and reliability [Pearton et al 2005]. They can be grown in thin film form on a variety of substrates, which include silicon, making these materials promising for integration with electronic circuitry, particularly for devices aimed for one-time use, low-price and mass production [Muralt 2008] (see Table 1). Amongst these, PZT has the highest piezoelectric constant and electromechanical coupling coefficient. However, for biosensing applications, PZT films have disadvantages such as higher acoustic wave attenuation, lower sound wave velocities, poor biocompatibility and worst of all, the requirement for extremely high temperature sintering and high electric field polarization, which make them largely unsuitable for integration with electronics (see Table 1). ZnO shows a high piezoelectric coupling, and it is easy to control the film stoichiometry, texture and other properties compared with that for AlN film [Jagadish & Pearton 2006]. Zinc oxide is considered

Materials	ZnO	AlN	PZT	Quartz	128° cut $LiNbO_3$	36° cut $LiTaO_3$	PVDF
Density (g/cm^3)	5.61	3.3	7.8	2.64	4.64	7.45	1.79
Moulus (GPa)	110-140	300-350	61	71.7		225	0.16
Hardness	4-5 GPa	15 GPa	7-18 GPa	Moh's 7	Moh's 5 Knoop 800-1000	70-110 Knoop 700-1200	Shore D75-85
refractive index	1.9 to 2.0	1.96	2.40	1.46	2.29	2.18	1.42
Piezo-constant d33 (pC/N)	12	4.5, 6.4	289-380, 117	2.3(d11)	19-27	-21	-35
Coupling coefficient, k	0.15-0.33	0.17-0.5	0.49	0.0014	0.23	0.2	0.12-0.2
Effective coupling coefficient, k^2 (%)	1.5-1.7	3.1-8	20-35	8.8-16	2-11.3	0.66-0.77	2.9
Acoustic velocity by transverse (m/s)	6336 (2650)	11050 (6090)	4500 (2200)	5960 (3310)	3970	3230-3295	2600
Dielectric constant	8.66	8.5-10	380	4.3	85 (29)	54 (43)	6-8
Coefficient of thermal expansion (CTE, $x10^{-6}$)	4	5.2	1.75	5.5	15	-16.5	42-75

Table 1. Comparison of common piezoelectric materials [Fu et al 2010]

biosafe and therefore suitable for biomedical applications that immobilize and modify biomolecules [Kumar & Shen 2008]. A summary of the recent development on ZnO film based microfluidics and sensing have been reported by Fu et al 2010. Currently, there is some concern that ZnO film is reactive, and unstable even in air or moisture and the stability and reliability is potentially a major problem.

AlN has a very large volume resistivity and is a hard material with a bulk hardness similar to quartz, and is also chemically stable to attack by atmospheric gases at temperatures less than 700°C. Compared with ZnO, AlN also shows a slightly lower piezoelectric coupling. However, the Rayleigh wave phase velocity in AlN is much higher than that in ZnO, which suggests that AlN is better for high frequency and high sensitivity applications [Lee et al 2004]. The combination of its physical and chemical properties is consequently promising for practical applications of AlN both in bulk and thin-film forms. Using AlN potentially enables the development of acoustic devices operating at higher frequencies, with improved sensitivity and performance (insertion loss and resistance) in harsh environments [Wingqvist et al 2007a]. AlN thin films have other attractive properties such as high thermal conductivity, good electrical isolation and a wide band gap (6.2 eV). Therefore, AlN thin films have been used, not only for the surface passivation of semiconductors and insulators, but also for both optical devices in the ultraviolet spectral region and acousto-optic devices.

This chapter will focus on reviewing recent progress covering the issues related to AlN film preparation, its microstructure, piezoelectric properties and device fabrication as well as applications related to microfluidcis and biosensing.

2. AlN film processing and characterization

The AlN crystal belongs to a hexagonal class or a distorted tetrahedron (see Fig. 2), with each Al atom surrounded by four N atoms [Chiu et al 2007]. The four Al–N bonds can be categorized into two types: three are equivalent Al–$N_{(x)}$ (x = 1, 2, 3) bonds, B_1, and one is a unique Al–N bond, B_2, in the *c*-axis direction or the (002) orientation. Since the B_2 is more ionic, it has a lower bonding energy than the other bonds [Chiu et al 2007]. The highest value of K_t^2 and the piezoelectric constant are in the *c*-axis direction, thus the AlN film growing with *c*-axis orientation has much better piezoelectricity when an acoustic wave device is excited in the film thickness direction.

2.1 AlN deposition methods

Many different methods have been used to prepare AlN films. These include chemical vapour deposition (CVD) or plasma enhanced CVD (PECVD) [Sanchez et al 2008, Tanosch et al 2006, Ishihara et al 2000, Liu et al 2003], filtered arc vacuum arc (FAVC) [Ji et al 2004], molecular beam deposition (MBE) [Kern et al 1998], hydride vapour phase epitaxy (HVPE) [Kumagai et al 2005], pulsed laser deposition (PLD) [Lu et al, 2000, Liu et al 2003, Baek et al 2007], and sputtering [Mortet et al 2003 and 2004, Auger et al 2005, Clement et al 2003]. Of these technologies, MBE can grow a single-crystal epitaxial AlN film with other advantages which include precise control over the deposition parameters, atomic scale control of film thickness and *in situ* diagnostic capabilities. However, it has limitations of low growth rate, expensive instrument setup and a high process temperature from 800 to 1000°C. Unfortunately this results in thermal damage of the AlN layers during deposition, as well as the substrate depending on the material. CVD technology including metal organic CVD

(MOCVD) and PECVD is also of great interest for AlN film growth because it not only gives rise to high-quality films but also is applicable to large-scale production. However, its high process temperature (about 500 to 1000 °C) may be inappropriate for CMOS-compatible processes and this causes large thermal stresses in the films, which potentially restricts the choice of substrate. The main advantages of PLD are its ability to create high-energy source particles, permitting high-quality film growth at potentially low substrate temperatures (typically ranging from 200 to 800 °C) in high ambient gas pressures in the 10^{-5}–10^{-1} Torr range. One disadvantages of PLD is its limited deposition size and uniformity.

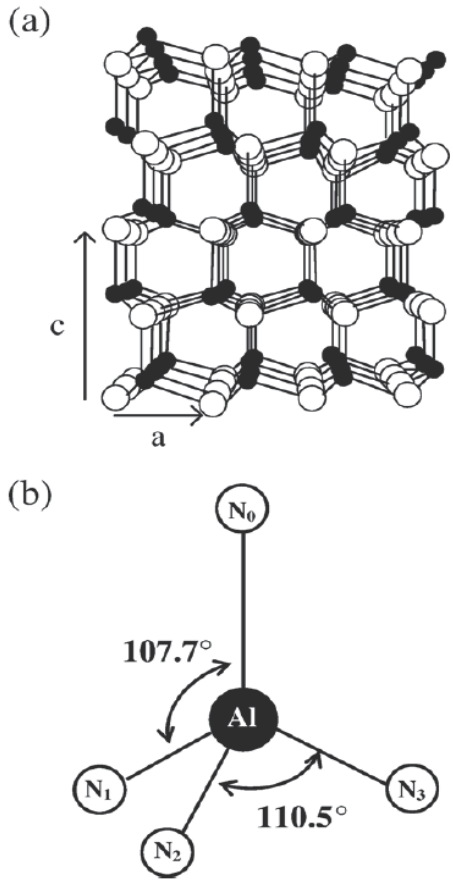


Fig. 2. (a) Hexagonal structure of AlN and (b) tetrahedral structure, with one Al atom surrounded by four N atoms [Chiu et al 2007].

One of the most popular thin film deposition techniques for AlN films is sputtering (DC, radio-frequency magnetron and reactive sputtering). They can be deposited in an N_2/Ar reactive atmosphere by DC reactive sputtering pure Al, or by RF sputtering using an AlN target. Sputtering methods can deposit a good crystalline AlN thin film at a relatively low temperature (between 25 °C and 500 °C) and the sputtered films normally exhibit good epitaxial film structure [Engelmark et al 2000]. DC Sputtering using an Al target can result in "target poisoning" caused by the accumulation of charging on the target, which causes arcing or a decrease in the sputtering rate. Switching the choice of power supply from DC to

RF addresses this problem, but at the cost of lower deposition rate and more expensive and complex equipment. Pulsed-DC reactive sputtering provides a solution to this limitation and also brings other advantages, which include higher film uniformity and higher plasma activity [Cherng et al. 2007, 2008].
From a MEMS fabrication point of view, reactive sputtering is one of the best methods, with good reproducibility and compatibility with planar device fabrication technology. In this section, we will focus on the processing, texture and acoustic wave properties of the sputtered AlN films.

2.2 Influence of process parameters

The quality of the sputtered AlN thin films depends on plasma power, working pressure, substrate temperature, RF power and substrate materials. Increasing the RF power causes higher kinetic energy of adatoms when they arrive on the substrate, which provides enough energy for the formation of the (0 0 0 2) preferred orientation of AlN layers. On the other hand, increased RF power also raises the number of ejected species from the target, which results in an increased growth rate as a function of RF power.
Gas pressure potentially also has a significant influence on AlN film deposition with increasing the sputtering pressure up to 1.33 Pa being reported to improve the crystalline quality of the (0 0 0 2)-oriented AlN layers. However, it was also noted that further increases in the sputtering pressure degraded the crystalline quality [Gao et al 2007]. Increasing in the sputtering pressure will raise the probability of collisions between sputtered particles and nitrogen atoms simply because of more gas atoms are available for ionization. Therefore, the average energy of the sputtered particles is increased which improves the crystalline quality. However, further increase in sputtering pressure results in the reduction of mean free path of N or Ar ions, which leads to a reduction of the energy of sputtered and deposited atoms, thus degrading the crystalline quality [Gao et al 2007].
Okamoto et al 2000 observed a change of the preferred crystallographic orientation by increasing the N_2 partial pressure, and Baek et al. 2007 detected the same effect when the substrate temperature and N_2 gas fluence were changed. Sudhir et al. 1998 demonstrated that the surface morphology and structure of the AlN films can be actively controlled by adjusting the nitrogen partial pressure during the film deposition. They attributed the observed dependence of the structural quality to the change in the surface diffusion of adatoms, given by $L \sim (D\tau)^{1/2}$, where D is the diffusion coefficient and τ is the residence time of adatoms. Larger values of diffusion length imply more time for the adatoms to find energetically favourable lattice positions, thus reducing the density of surface defects and improving the crystal quality [Sudhir et al 1998].
Leong and Ong 2004 prepared reactive magnetron sputtered AlN films by varying parameters such as substrate temperature *Ts*, radio frequency power *Pw*, and substrate materials (including silicon, platinum coated silicon and sapphire). The effects of these parameters on film microstructure as a function of deposition temperature are shown in Fig. 3. This identifies the regions of nearly amorphous (na-) AlN, polycrystalline (p-) AlN, texture (t-) AlN and epitaxial (e-) AlN on three substrate materials, i.e. Si(100), Pt(111)/Si(100) and Al_2O_3(001), respectively. The 'na-AlN" means that the microstructure of AlN has a highly disordered matrix containing small randomly orientated crystals, which normally forms at a lower rf power, and low temperature [Leong & Ong 2004]. At higher temperature and power, the thermal energy gained by the depositing species is larger, and the atoms are more mobile. Hence, the species more readily aggregate and crystallize,

resulting in the formation of larger grains compared with those present in the na-AlN structure. Increases in *Ts* and *Pw* have the effects of increasing the thermal energy of the species on the substrate surface, and enhancing the crystallization of the deposits and preferential orientation of grains. It should be noted that sapphire substrate have better lattice matching with the AlN, which facilitates the epitaxial growth of the AlN structure [Leong & Ong 2004].

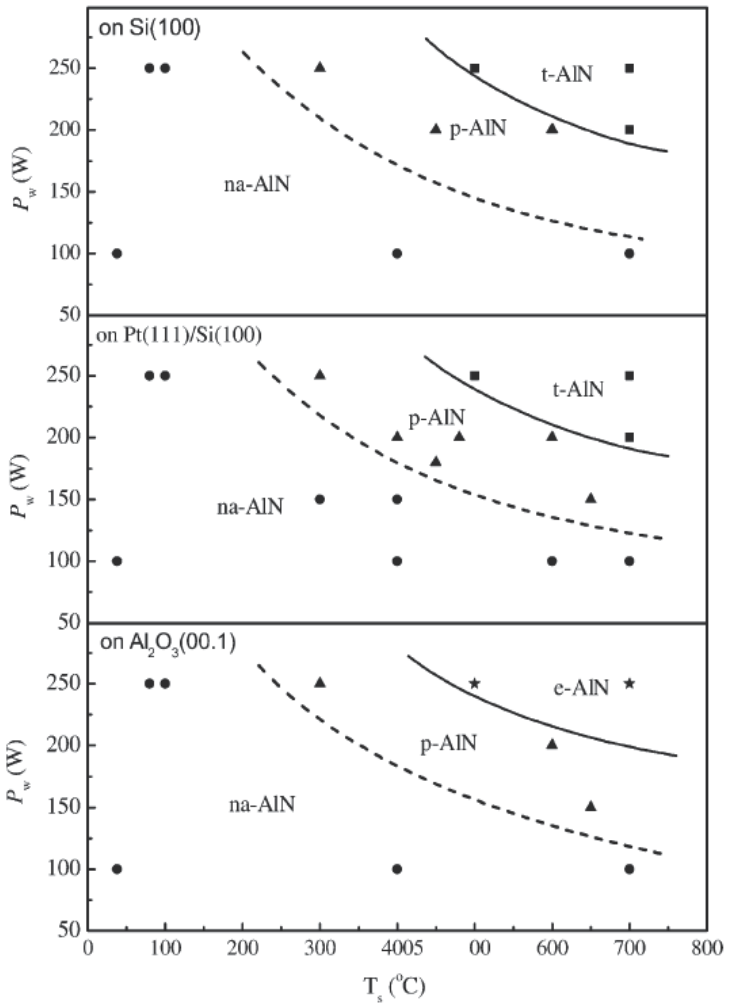


Fig. 3. Effects of the process parameters on film microstructure on three substrate materials, i.e. Si(100), Pt(111)/Si(100) and Al_2O_3(001) [Leung & Ong 2004]

Because of the reactivity of Al, a high-purity source Al material and an oxygen-free environment are required to grow high-quality AlN film [Vashaei et al 2009]. Hence, oxygen has a significant influence on AlN film growth during sputtering, and contamination due to residual oxygen or water can seriously interfere with the formation of the AlN film structure. Growth rate of the AlN film decrease with increased oxygen in the sputtering gas and their predominant polarity also changes from Al polarity to N polarity with increase in

the oxygen concentration [Vergara et al 2004, Cherng et al 2008 a and b]. Increased oxygen concentration in sputtering gas increases Al-O bonding, as the bonding energy of Al-O (511 kJ/mol) is higher than that of Al-N (230 kCal/mol) [Akiyama et al 2008], and formation of Al-O bond significantly deteriorates the piezoelectric response of the AlN films.

The quality of AlN films is affected by any contamination during sputtering [Cheung & Ong 2004], resulting from target impurity, gas impurity, and residual oxygen/moisture from both inside (adsorption) and outside (leakage) the working chamber. Out-gassing is a critical parameter that must be controlled for quality of AlN crystals, and effect of the out-gassing rate has been evaluated by observing the pressure increase with time after the designated base pressure has been reached and the pump was shut down (as shown in Fig. 4). The FWHM (full width of half maximum) from an X-ray diffraction rocking curve and the residual stress of the films has been obtained in order to compare the film quality [Cherng 2008 and 2009].

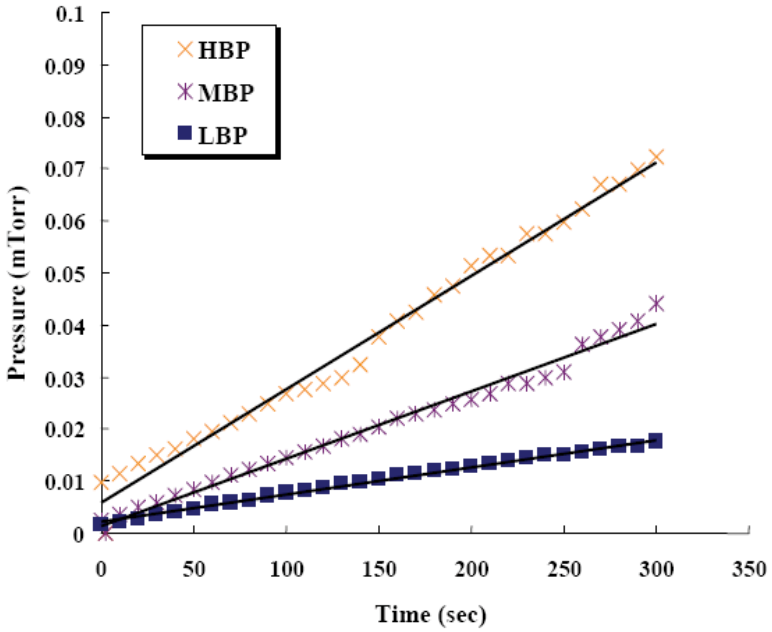


Fig. 4. Outgassing rate evaluated by observing the pressure increase with time after the designated base pressure was reached and the pump was shut down where the slope of each curve indicates its outgassing rate respectively. The sputtering system was either pumped down to a base pressure of 3×10^{-6} Torr (thus termed HBP, high base pressure) or 1×10^{-6} Torr (thus termed MBP, medium base pressure) or 4×10^{-7} Torr (thus termed LBP, low base pressure) before admitting the gas mixture in, in order to examine the effects of outgassing [Cherng & Chang, 2008]

Figures 5(a) and (b) show the effect of working pressure on FWMH and film stress at different outgasing levels. The FWHM decreases and residual stress becomes more compressive with decreasing working pressure. As the pressure is decreased, the mean free path of the sputtered atoms becomes comparable with the target-to-substrate distance (λ_{mfp}=5 / P, where λ_{mfp} is in cm and P in mTorr) [O'Hanlon 1989], and hence less gas phase scattering is observed. The result is that sputtered Al atoms arrive on the surface of the

growing film with most of their energy retained. They transfer a substantial amount of energy to the growing film, and thus increase the mobility of the adatoms and can then move to the lattice sites which form a closest-packed (0002) plane with the lowest surface energy. In fact, the energy delivered to the growing film is sufficiently high so that fully (0002)-textured (texture coefficient=1) AlN films with FWHM of the rocking curve lower than 2° are readily obtainable without substrate heating. In addition to the aforementioned "atom-assisted deposition" [Iriarte et al 2002], a second mechanism, namely, "atomic peening" [Windischmann 1992] is also at work. Since N atoms are lighter than Al, the reflection coefficient of N ions is high sufficient for a large fraction of them bombarding the Al target to be neutralized and reflected off the target surface upon impact. This results in additional bombardment of the growing film by energetic N neutrals. On the other hand, Ar ions are effectively not reflected since they are heavier than Al. Both the atom assisted deposition and atomic peening mechanisms require a sufficiently low working pressure so the energetic particles do not lose much of their energy while travelling through the gas phase. This explains why as the working pressure decreases, the FWHM of the rocking curve decreases and the residual stress becomes more compressive [Cherng & Chang, 2008].
Lower outgassing levels show a better figure-of-merit that not only the FWHM of the rocking curve is lower, but also the change of residual stress with pressure occurs in a much smoother manner and with much smaller magnitude. X-ray Photoelectron Spectroscopy (XPS) analyses for four selected samples circled in Fig. 5(a), reveal higher oxygen contents for samples with higher outgassing. SEM observations show thinner and slanter columnar structure in the AlN film when outgassing is higher upon sputtering. Both of the lower residual stress levels and the lower FWHM values at lower outgassing can be attributed to oxygen-related extended defects [Cherng & Chang, 2008].
Figure 5 © shows the relationship between FWHM and pressure at different target-to-substrate distances. At a longer target-to-substrate distance, the insensitive region shrinks and the threshold value shifts to a lower pressure [Cherng & Chang, 2008]. This is due to the decreasing ratio of mean free path to target-to-substrate distance, indicating more gas phase scattering and thus worse film quality.
With increasing nitrogen concentration, atomic peening is favoured while atom-assisted deposition basically remains unaffected. The former explains the decreasing FWHM values and more compressive stress with increasing N_2 %, as shown in Figs. 6(a) and (b). At a lower base pressure, the influence of atmospheric composition diminishes to such an extent that the FWHM of the rocking curve practically stays the same between 20 and 90 % N_2. This finding, together with the insensitive FWHM vs. pressure regions (see Fig. 6) reveal that oxygen contamination is the most dominant factor for the film properties. In the other hand the residual stress at lower outgassing rates varies little with nitrogen content. The oxygen related extended defects are deductive to compressive stress, instead of tensile stress, which is normally caused by re-sputtering type of defects. As seen in Fig. 6(c), the FWHM of the rocking curve decreases with increasing substrate temperature. This is consistent with the higher mobility of adatoms at higher substrate temperatures. Once again, the behaviour at lower outgassing becomes insensitive with substrate temperature. At this point, it is worth noting that at low outgassing, a somewhat "insensitive" region and/or a so-called "threshold" behaviour exists with all process-related parameters, e.g., working pressure, atmosphere composition, and substrate temperature. This emphasizes the crucial role oxygen contamination plays in pulsed-DC reactive sputtering of AlN thin films.

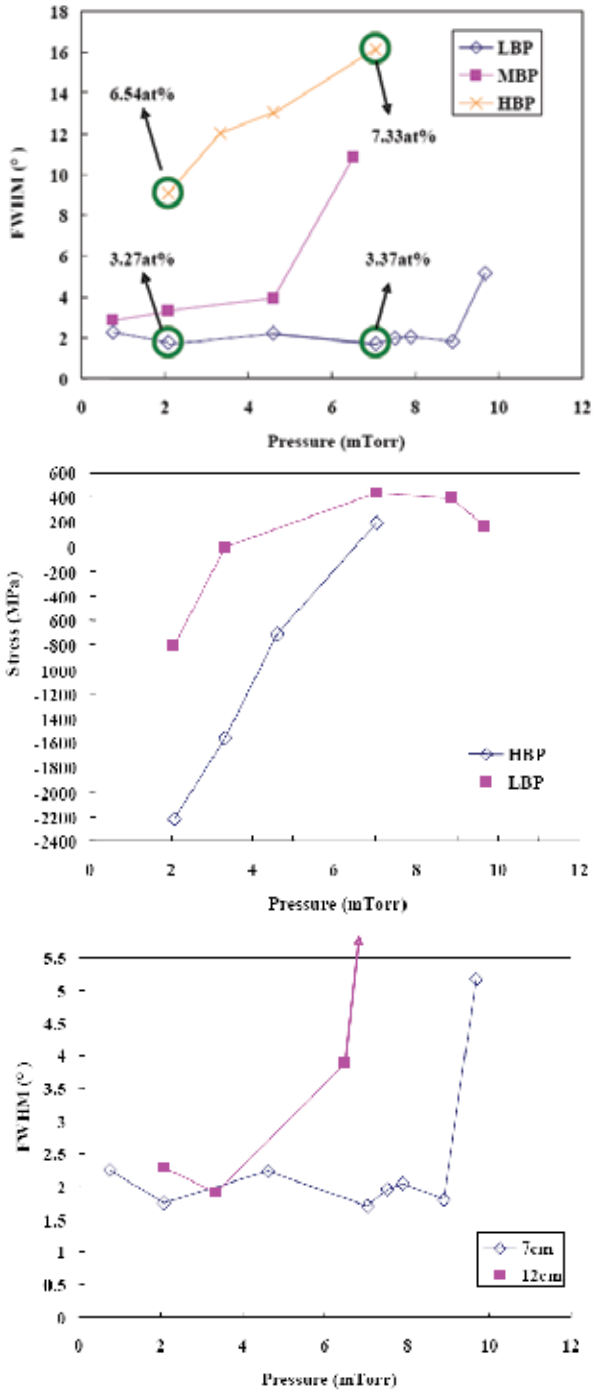


Fig. 5. Effect of working pressure on (a) XRD FWHM; and (b) film stress at various outgassing levels; and (c) on XRD FWHM at various target-to-substrate distances [Cherng et al 2008].

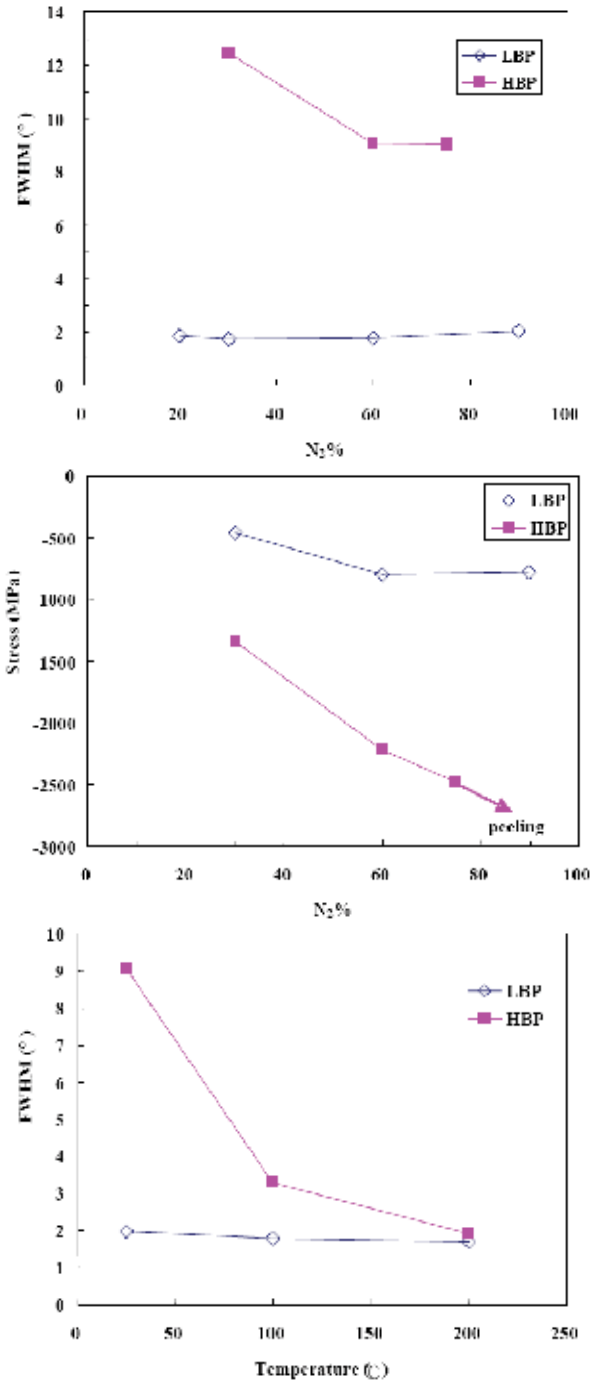


Fig. 6. Effect of atmospheric composition on (a) XRD FWHM (b) residual stress at various outgassing levels; (c) effect of substrate temperature on XRD FWHM at various outgassing levels [Cherng et al 2008].

2.3 Two-step deposition

The growth dynamic or surface kinetic roughening of the sputtered AlN films grown on Si (100) substrates has been thoroughly studied, and a two-stage growth regime identified [Auger et al 2005]. In the first step, the growth dynamic is unstable with significant sticking probabilities of the impinging particles. The films have a mixture of well textured and randomly oriented crystals. In the second regime, the films are homogeneous and well textured, and the growth is dominated by the shadowing effect induced by the bombardment of impinging particles [Auger et al 2005]. Based on this effect, a two-step pulsed-DC reactive sputtering model has been proposed with various process parameters including working pressure, discharge power, and reactive atmosphere during two stage sputtering [Cherng et al 2008, 2009]. Two-step sputtering for an AlN piezoelectric layer generally consists of a 10-min deposition for the base layer and a subsequent 50-min sputter for the top layer. As a comparison, one-step sputtering (60 min) has also been conducted with the same sputter parameters as those used for the base layer in two-step sputtering.

2.3.1 Two-step working pressure method

Figure 7 shows the effects of working pressure on (a) XRD FWHM, and (b) residual stress of AlN piezoelectric layer for both one-step and two-step sputtering, respectively. The data for two-step sputtering, when compared to their one-step counterparts, show a better figure-of-merit in that not only the FWHM of the rocking curve is smaller, but also the magnitude of the residual stress is smaller and its variation with pressure is smoother [Cherng et al 2008]. If we attribute the first step sputtering to initial nucleation and the second step to the subsequent growth of the AlN film, then the better film quality for two-step sputtering (when compared to its one-step counterpart with the same process parameters used for the base layer) has to be due to the sputtering conditions for the growth of the top layer [Cherng et al 2008]. Therefore, as far as the rocking curve width and residual stress are concerned, it is fair to say that growth, instead of nucleation, dominates the performance of two-step working pressure method.

For the AlN film deposited on Mo substrates, the FWHM values for both the 1-step and 2-step methods do not vary with working pressure and remain at the same low value of about 1.3^{o} as shown in Fig. 7 (a) [Cherng et al 2008]. This is further confirmed by Fig. 7(b), where both the 1-step and 2-step methods using Mo substrates show low residual stress, regardless of the working pressure.

2.3.2 Two-step power method

For one-step sputtering on Si, the FWHM of the rocking curve decreases with increasing discharge power as shown in Fig. 8. This is due to the enhanced atom-assisted deposition and atomic peening mechanisms at a higher power. The sputter yield (at higher discharge voltage) and plasma concentration (more ionized species at higher discharge current) have been increased. Growth, instead of nucleation, dominates the performance of two-step power method on Si, because the data of two-step sputtering are much better than that of its one-step counterpart as also shown in Fig. 8 [Cherng et al 2009].

2.3.3 Two-step nitrogen concentration method

With increasing nitrogen concentration, atomic peening is favoured while atom-assisted deposition remains unaffected. For one-step sputtering on Si, this enhanced atomic peening

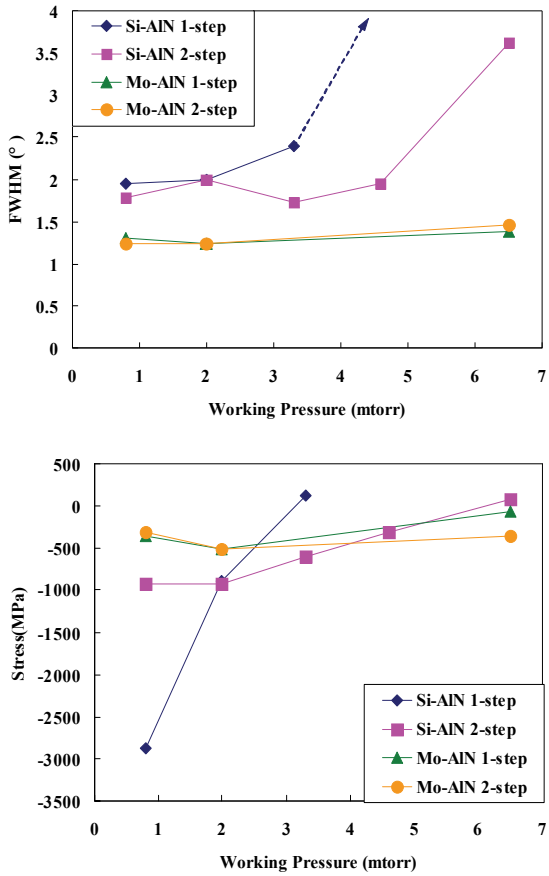


Fig. 7. Effects of working pressure on (a) XRD FWHM, and (b) residual stress of AlN piezoelectric layer for both one-step and two-step sputtering, respectively [Cherng et al 2008]

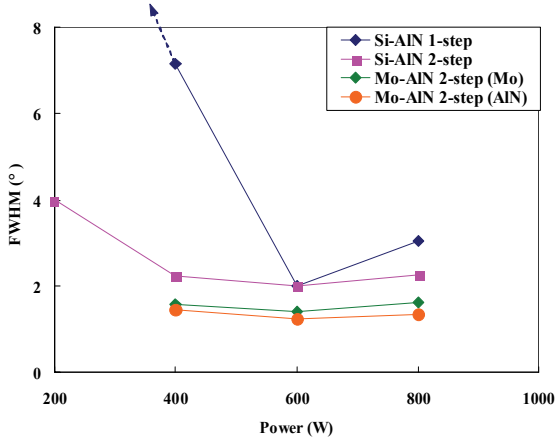


Fig. 8. Effects of discharge power on XRD FWHM for both one-step and two-step sputtering [Cherng et al 2009].

is thought to be responsible for the decreasing FWHM of the rocking curve from 40 to 60 % of N_2, as shown in Fig. 9. The higher FWHM value at 100 % N_2 is probably due both to the excess atomic peening (causing re-sputtering) and to the worsened sputter yield (since N has a lower sputter yield than Ar). It is worth noting that the target does not exhibit any hysteresis-related phenomenon even under pure nitrogen. The employment of pulsed power is believed to be able to clean up the surface of the Al target effectively [Cherng et al 2009]. On the other hand, the FWHM behaviour for the two-step atmosphere method on Si seems to be mostly determined by initial nucleation rather than subsequent growth. The data for this is much closer to those of the one-step counterparts which employ the same sputtering conditions for the base layer. This phenomenon is just the opposite to the one observed for the other two-step methods described above, and has to be closely related to the atomic peening mechanism. It is thought that in the case of lighter bombarding particles (N atoms for atomic peening vs. Al atoms for atom assisted growth), the sputtering conditions for subsequent growth are not appropriate to alter the effects of the initial nucleation [Cherng et al 2009]. For deposition on Mo, once again, the quality of the AlN piezoelectric film is dominated by the underlying Mo film, regardless of the reactive atmosphere, as evidenced by the two-step sputtering data of the AlN and Mo films.

In conclusion, a methodology of two-step pulsed-DC reactive sputtering has been systematically evaluated for making (0002)-textured AlN thin films with independent control of rocking curve width and residual stress. This methodology was best demonstrated by the two-step working pressure method on Si, which was capable of reactively sputtering AlN thin films with almost constant rocking curve widths of about 2°, with a constant deposition rate of about 36 nm/min, and a continuously adjustable residual stress between −926 and −317 MPa [Cherng et al 2008 and 2009]. In addition, it was noted that growth dominated the performances of both the two step working pressure method and the two-step power method, while nucleation dominated the two-step atmosphere method.

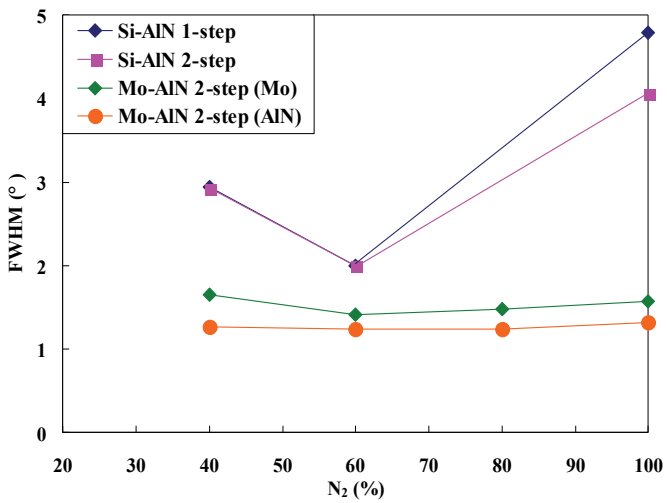


Fig. 9. Effects of reactive atmosphere on XRD FWHM for both one-step and two-step sputtering [Cherng et al 2009].

3. Piezoelectric properties of sputtered AlN films

3.1 Film thickness effect

For a SAW device made on a very thin AlN film (less than a few hundreds of nanometers), the acoustic wave can penetrate much deeper into the substrate as the film thickness is normally much less than one wavelength. In this case, the energy of a SAW device is largely dissipated in the substrate where the wave predominantly propagates. Therefore, the wave velocity of the SAW approaches the Rayleigh velocity of the substrate material as shown in Fig. 10(a) [Clement et al 2003, 2004]. When the AlN film thickness is increased, the acoustic velocity gradually changes to that of AlN film. However, there is normally a cut-off thickness, below which no wave mode can be detected, due to the low electromechanical coupling coefficient for a very thin AlN film. A Rayleigh-type wave (called the fundamental mode or mode 0) can be generated when the film is thin. With increasing film thickness, a higher order acoustic wave mode known as the Sezawa wave (mode 1) can be obtained. A Sezawa mode is realized from a layered structure in which the substrate has a higher acoustic velocity than the overlying film. This wave exhibits a larger phase velocity (higher resonant frequency) than the Rayleigh wave for a fixed thickness, and is thus desirable for high frequency applications. In a similar manner to that of Rayleigh wave, the resonant frequency and the phase velocity of the Sezawa wave decreases with film thickness. There are other higher order acoustic wave modes (modes 2 and 3, etc.) as shown in Fig. 10(a) [Clement et al 2003].

There are two key issues for the piezoelectric properties of the AlN acoustic wave device: the electro-mechanical coupling coefficient k_{eff}^2 and the quality factor Q. The effective coupling coefficient (k_{eff}^2) is related to the relative spacing between the resonant frequency and the parallel resonant frequency, and it determines the bandwidth for a band-pass filters. Fig 10 (b) shows the effective coupling coefficients of different wave modes as a function of the thickness ratio of the electrode-to-piezoelectric layers for AlN thin-film resonators [Clement et al 2003]. The quality factor Q is determined by the energy conversion efficiency from electrical into mechanical energy. However, improving one of those two parameters can cause a decrease of the other one, therefore, it would be necessary to optimize both parameters using one figure of merit (FOM), defined by the product of $k_{eff}^2 \times Q_D$ [Grate 2000].

3.2 Effects of electrodes

For an AlN based acoustic wave device, parameters such as the Q factor, resonant frequency and effective coupling constant are determined by the film and electrode material quality, as well as the electrode thickness and film roughness [Lee et al 2002]. Common used electrode materials include (111) oriented face centered cubic (fcc) metals such as Al, Pt and Ni, (110) oriented body centered (bcc) materials like Mo and W, and hexagonal metals with a (002) orientation including Ti and Ru [Lee et al 2004]. Some commonly used electrode materials for AlN SAW devices include Mo, W, Ti, Al, Au, Pt, Ni and TiN, and Ag, Co, Cr, Cu Fe, Nb, Ni, Zn, Zr have also been reported as electrodes for these acoustic wave devices [Lee et al 2004, Akiyama et al 2004]. Metallic electrodes can help promote the growth of highly *c*-axis oriented AlN films., and they can also contribute to the confinement of the mechanical energy in the piezoelectric layer at the resonant frequency. A high acoustic impedance mismatch between the piezoelectric layer and the electrodes is normally preferred, thus for this purpose, the heavy and stiff metals are the candidates of choice.

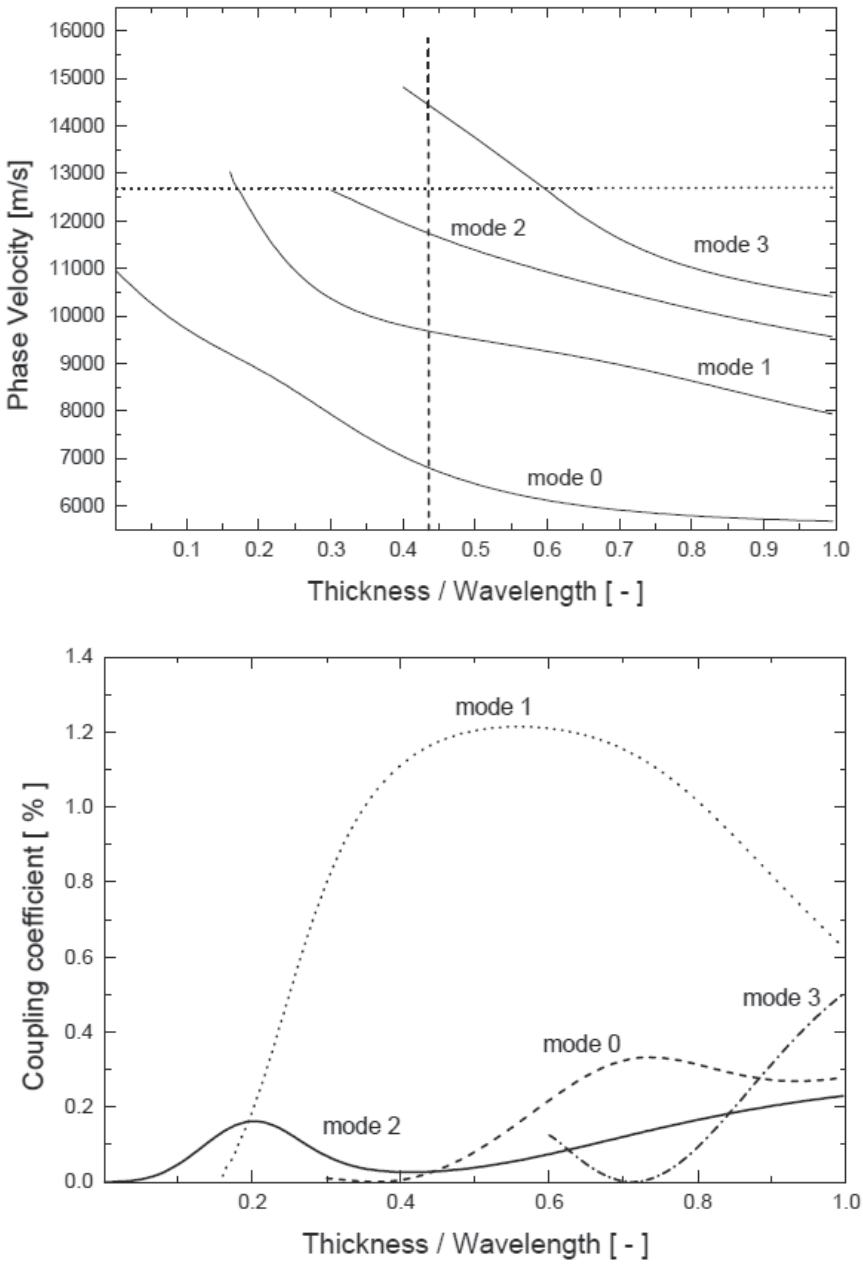


Fig. 10(a). Phase velocity for AlN film as a function of thickness/wavelength ratio for different acoustic wave modes (b). Effective coupling coefficient as a function of thickness ratio of electrode-to-piezoelectric layers for AlN thin-film resonators [Clement et al 2003].

Gold electrodes show the best resonant characteristics. The characteristics of Ag and Cu electrodes are very close to those obtained for gold, but much cheaper. Al and Mo have low resistivity and high Q factors with Mo being one of the most reported electrodes in the AlN film based acoustic devices, because it promotes the growth of highly textured AlN films

[Akiyama et al 2005, Huang et al 2005 a and b, Lee et al 2003, Okamoto et al 2008, Cherng et al 2004]. It was reported that the best-textured AlN films deposited by sputtering on metallic surfaces have been grown on Pt substrates [Lanz & Muralt 2005].
For the AlN FBAR device, the bottom metal layer significantly affects the texture of AlN films and its electro-acoustic properties. AlN films deposited on the materials with fcc lattice structure show a high c-axis orientation, especially for Au and Pt [Tay et al 2005]. Ti has a hexagonal structure similar to that of AlN [Lee et al 2004, Chou et al 2006], while W has a low acoustic attenuation, small mismatch in the coefficient of thermal expansion and high acoustic impedance with AlN, and is thus a good electrode material for AlN devices. Ni has often been chosen because of its surface smoothness, but the AlN film texture on Ni is not as good as that on the other fcc metals. Tantalum [Hirata et al 2007] and iridium [Clement et al 2009] have also been reported as electrodes for AlN film growth. Iridium is of interest as it is a precious metal similar to Pt but considerably cheaper, with a high sound velocity (5300 m/s), and a lower diffusivity in Si than other heavy metals (Au, Pt) [Benda et al 1998]. The thickness ratio of AlN and top or bottom electrodes has been reported to have a significant influence on piezoelectric effect of AlN films [Huang et al 2005 a and b, Akiyama et al 2004]. Lee et al 2002 found that a resonator with a thicker Mo electrode can provide higher Q values than those with thinner Mo electrode.

3.3 Film texture and substrate effect

AlN films with strong texture can have good piezoelectric coefficients, high electromechanical coupling, and acoustic velocities approaching those of the single crystal AlN. The sputtering process parameters significantly affect the orientation of the deposited AlN films. Okano et al. 1992 identified that the *c*-axis orientation increases as the N_2 concentration in the mixture of Ar and N_2 decreases, while Naik et al 1999 have shown that the *c*-axis orientation increases as the sputtering pressure is reduced. AlN films have been reported to show preferred (002) growth orientation on a number of materials which include silicon, quartz, glass, $LiNbO_3$ [Caliendo et al 2003, Lee et al 2004], GaAs [Cheng et al 1998], GaN/Sapphire [Kao et al 2008,Xu et al 2006], SiC [Takagaki et al 2002] and ZnO layer [Lim et al 2001]. For AlN film growth, the texture of film is the result of competitive growth of (100) and (001) planes [Clement et al 2003]. When the (001) crystal growth is favourite, the AlN crystals will grow with a (002) orientation. When (100) crystal growth is more favourite, the other orientations can be dominant, such as (103) (100) (110) and (102) etc. The energy input into the plasma adatoms during film growth is the dominant parameter that controls the film orientation. The possible solutions for better orientation include: higher plasma energy, higher Ar ion energy, application of negative self-bias voltage, shorter target-to-substrate distance and lower pressure. [Clement et al 2003].
Sputtered AlN films normally show a (002) film texture, which results in longitudinal (or Rayleigh) wave modes and is therefore good for sensing in air or gas. However, as explained before, if liquid exists on the sensing surface, excessive damping and attenuation of the propagating wave occurs when the longitudinal mode couples into the liquid**.** This problem can be solved by generating a shear-horizontal (SH) SAW, which propagates on a piezo-material by an in-plane shear horizontal motion [Wingqvist et al 2007], and dramatically reduces SAW coupling into a liquid medium [Mchale 2003]. However, the commonly observed (002) texture in the sputtered AlN films is unsuitable for generating SH–SAWs. In addition, using a pure shear wave is not efficient for driving liquid droplets forward. A good approach to solving the problem is to develop AlN films in which the c-

axis is inclined relative to the surface normal, thus allowing both longitudinal and shear wave modes to be generated [Webber 2006]. These two modes will have different frequencies and thus can be individually controlled for either pumping or sensing purposes. To the best of our knowledge, there are no reports of the application of both the functions (microfluidics and biosensing) on a c-axis inclined AlN based SAW device in liquid conditions. The techniques for the deposition of the inclined AlN film include: (1) using a tilted substrate (up to 45^o) with a controlled position under the sputter-target; (2) using a high energy nitrogen ion beam aimed at the desired angle with respect to the substrate surface normal [Yanagitani & Kiuchi 2007]. Obtaining the inclined AlN films strongly depends on the sputtering pressure, temperature, the oblique incidence of particles [Yang et al 2009]. C-axis inclined AlN films have been deposited on different substrates, including silicon and diamond [Fardeheb-Mammeri et al 2008]. Bjurstrom et al 2004 systematically studied the electromechanical coupling coefficient for both the shear and longitudinal modes at different AlN inclined angles. The k^2 of longitudinal mode has a maximum value for C-axis AlN crystals (θ=0^o), but gradually decreases as angle increases. On the contrary, the k^2 value of shear mode gradually increases as the inclined angle is increased from 0 to a peak value at angle of 45^o. The k^2 value of the two modes reaches to a similar value at angle of 30 to 35^o [Bjurstrom et al 2004].

The acoustic velocity in an AlN/Si SAW device also depends on the orientation of the Si substrate, being about 4700 m/s for Si (111) and 5100 m/s for Si (100) [Clement 2003]. AlN films have been deposited on 128^o $LiNbO_3$ substrate in order to enhance the SAW velocity and improve the temperature stability, i.e., decrease the temperature coefficient of frequency (TCF) [Kao et al 2003, Wu et al 2001 and 2002].

Recently, there has been much research on the deposition of AlN on diamond for SAW devices [Mortet et al 2003, Kirsch et al 2006, Le Brizoual et al 2007, Paci et al 2007, Elmazria et al 2003 and 2009, El Hakiki et al 2007, Wu et al 2009, Shih et al 2009, Iriarte et al 2003, Benedic et al 2008, Lin et al 2009]. The drive for this is that diamond substrates offer a higher phase velocities (6 km/s to 16 km/s) [Wu et al 2008]. Figure 12 shows dispersion curves of the first five Rayleigh SAW modes of IDT/(002) AlN/(111) diamond devices plotted as a function of the film thickness ratio h/λ. The phase velocity of each mode decreases as the film thickness ratio increases. For mode 0, the value of phase velocity is determined by the SAW velocity of (111) diamond, i.e., 10.9 km/s at h/λ =0 and the film thickness ratio h/λ increases, the phase velocity rapidly decreases. At h/λ =3, the velocity of the (002) AlN/diamond is about 5.4 km/s. It can be observed that the harmonic peaks of modes 1, 2, 3, and 4 cut off at the critical point where the phase velocity is equal to the shear bulk wave velocity in (111) diamond (12.3 km/s). For example, the cut-off of mode 1 occurs at h/λ= 0.172, mode 2 at h/λ =0.295, mode 3 at h/λ= 0.594, and mode 4 at h/λ=0.693 [Wu et al 2008]. Similar results have been reported by Benetti et al 2005.

The formation process and growth mechanism of an AlN layer on a (001) diamond substrate has been studied by Imura M et al, 2010. At the initial stage of AlN growth on diamond, the randomly oriented AlN grains are generated and grown three dimensionally with the formation of columnar domains due to the 20% lattice mismatch between AlN and diamond. At the second stage of growth, the c-axis-oriented AlN grain grows by incorporating the randomly oriented AlN grains. This occurs because of the high-growth rate of AlN grains along the [0001] direction [Imura M et al, 2010]. Diamond is a better substrate for epitaxial AlN growth than Si (111) [Imura M et al, 2010], but it is expensive, needs to be deposited at a high temperature, and the resulting surface roughness of the

diamond film is normally quite high. Other alternative choices are diamond-like-carbon (DLC) and nanocrystalline diamond films.

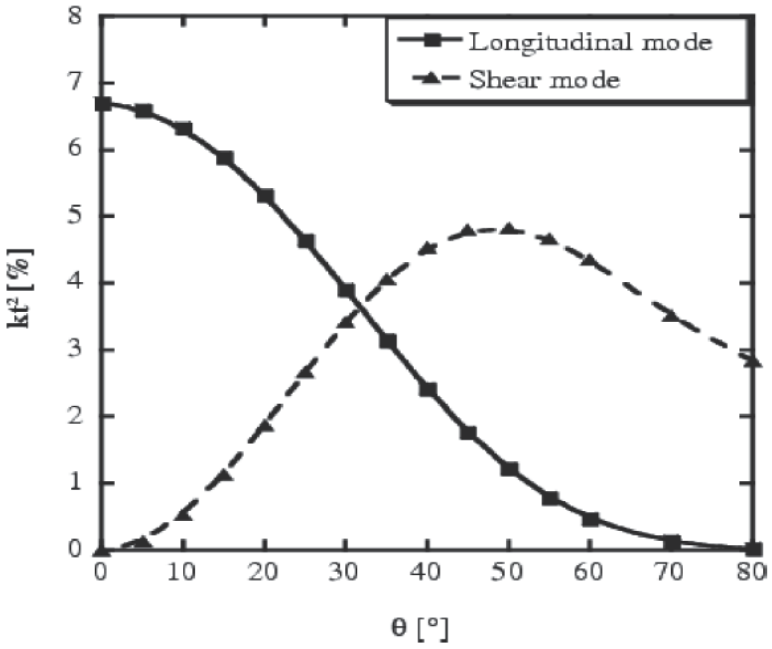


Fig. 11. Electromechanical coupling for both shear and longitudinal modes at different AlN crystal tilt, θ [Bjurstrom et al 2004]

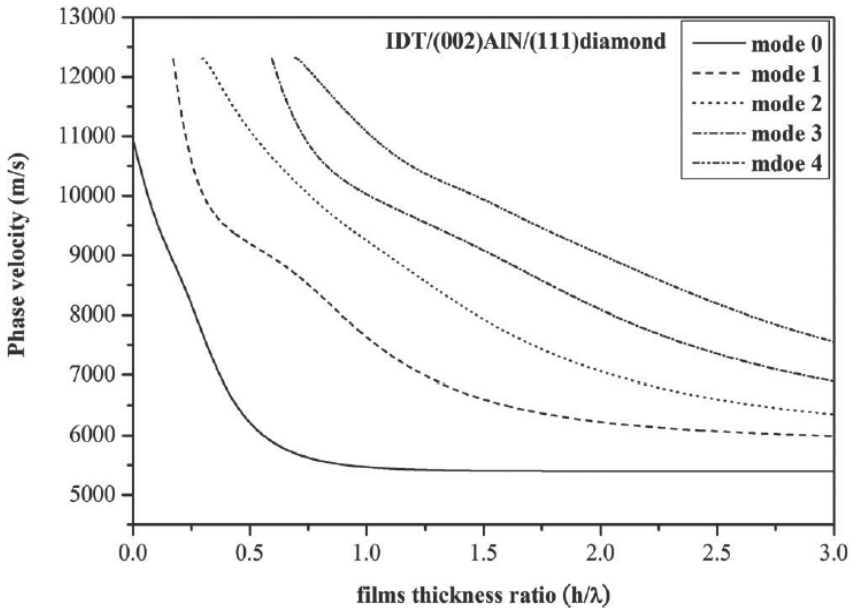


Fig. 12. Phase velocities dispersion curves of the first five Rayleigh SAW modes propagation in the IDT/(002)AlN/(111)diamond [Wu et al 2008].

Recently, AlN has been grown on SiC/Si substrate at a relatively low cost. The CTE of the SiC closely matches that of AlN, and the lattice mismatch is less than 1%. Both materials have been used in applications in high temperature packaging, and AlN thin films have been used as buffer layers for SiC grown on Si substrate [Chung & Kim 2007]. Therefore, SiC is an ideal candidates for a buffer layer of AlN films grown on SiO_2/Si substrates for SAW applications at different temperatures [Hoang & Chung et al 2009].

The AlN's temperature dependent ultrasonic properties (including ultrasonic attenuation coefficient, ultrasonic velocity, and acoustic coupling coefficient constant) have been calculated in the temperature range 200 to 800K by Pandey & Yadav 2009. The total attenuation is mainly dominated by phonon-phonon interaction, and the attenuation decreases from 200– 400 K sharply but it increases gradually from 400–800 K. Thus the temperature 400 K is the characteristic temperature for AlN. The decrease in attenuation from 200–400 K is due to the temperature variation of the thermal relaxation time or thermal conductivity of the material [Pandey & Yadav 2009]. A gradual increase in the attenuation from 400–800 K correlates mainly the temperature variation of the ultrasonic velocities or second order elastic constants. The temperature dependent ultrasonic velocity gives direct information about temperature variation of elastic constants, and the ultrasonic attenuation directly relates to the thermal conductivity or thermal relaxation time at temperature below 400 K and the elastic constants/ velocity of AlN above 400 K [Pandey & Yadav 2009]

During sputtering, particle bombardment can induce large film stress [Iborra et al 2004], and films with large compressive stress can cause buckling–induced delamination in the deposited films and fracture in the released devices. In Ar/N_2 based deposition system for AlN film, the energy of Ar ions colliding with the substrate controls the preferred orientation of the AlN films. The direction and energy of the ions determine the residual stress levels of the AlN films. The film stress or energy of the Ar bombardment can be adjusted by the substrate bias voltage during sputter deposition. Thermal annealing is a good method for post-treatment to reduce the film stress and improve the coating quality [Hung & Chung 2009].

4. MEMS processing and functionalization of AlN films

There are some reports on surface micromachining of AlN [Hara et al 2005], and chromium has been used to form both a good etch mask and electrode [Saravanan et al 2006]. Germanium can be used as the sacrificial layer for the AlN films, instead of common amorphous silicon, SiO_2, or other metal layer. AlN can be etched in aqueous solutions, such as KOH, NaOH, HF/H_2O, HF/HNO_3, tetramethyl ammonium hydroxide (TMAH), and the etch rate is temperature and crystal polarity sensitive [Jasinki et al 2003, Sheng et al 1988, Tan et al 1995]. Dilutate TMAH etches AlN but when with silicic acid and ammonium persulphate, it can be used to etch silicon with a very low etch rate for AlN film, and thus can be used for sacrificial etching [Kar et al 2009]. AlN can be electrochemically etched in electrolytes, such as HPO_3 (60°C to 90°C) or KOH solutions, and the etch rate is strongly dependent on the coating quality (from tens of nm/min up to a few μm/min). The reaction can be expressed as [Zhang & Edgar 2005]:

$$AlN+6KOH \rightarrow Al(OH)_3 \downarrow + NH_3 \uparrow + 3K_2O \downarrow \quad (5)$$

For dry etching process, AlN is normally etched using chlorine based plasma, such as chlorine and BCl_3, rather than a fluorine, as aluminum fluoride is stable and non-volatile [Khan et al 2002 or 2006]. Etching in a Cl-based plasma is normally isotropic, and the volatile reaction product is $AlCl_3$ at high temperatures (above 180°C) or Al_2Cl_6 at a room temperature [Engelmark 2003].

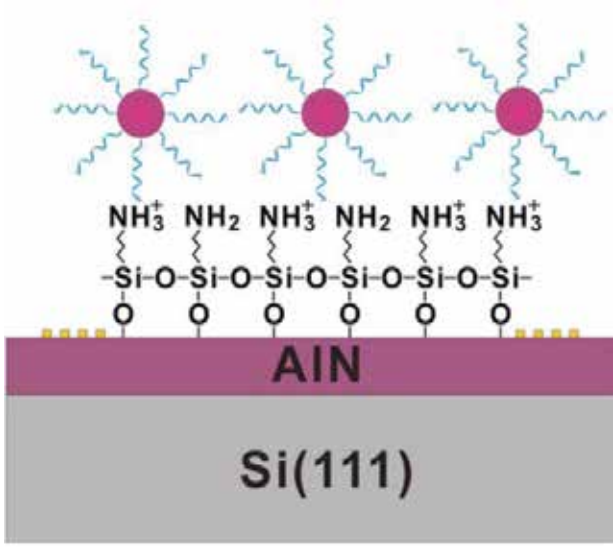


Fig. 13. Schematic diagram of antibody immobilized AlN/sapphire [Chiu et al 2008]

Recently, there have been studies for the surface functionalisation of AlN film for biosensing applications [Chiu et al 2008]. For example, by using silane, a new chemical layer can form on the AlN, and the functional groups on the silane surface can then be used as anchor points for the antibodies. In reference [Chiu et al 2008], antibody immobilized AlN/sapphire was prepared using the process shown schematically in Fig. 13. The AlN films were pre-treated prior to silanization using two methods. In the first method, they were treated by exposure to oxygen plasma. The second method treated the AlN surfaces by ultrasonication in 3/1 (in vol%) piranha solution, followed by rinsing in DI water. Piranha treatment was chosen because it is commonly used as a surface preparation method for silanization of other types of inorganic surfaces. Improved silane surfaces should create a more stable and ordered silane layer for the linkage of antibody, phage or other detecting ligands in the biosensor under development [Chiu et al 2008]. The treated AlN samples were silanized with octadecyltrichlorosilane (OTS). The ability to produce repeatable, homogeneous layers of selected chemical groups by silane derivatization of the AlN surface is considered to be a promising step in the development of a biosensor that uses surface immobilized phage or antibody ligands for analyte detection [Chiu et al 2008].

5. Thin film acoustic devices for biosensor applications

5.1 AlN SAW device

AlN based SAW devices have been reported as promising for high frequency, high sensitivity applications [Chiu etal 2008, Caliendo 2003 a and b, Xu et al 2004, Assouar et al

2002, Mortet et al 2003, Clement at al 2004]. For example, AlN Rayleigh SAW device with velocity of 4590 m/s has been fabricated for surface biofunctionalization using amiosilane molecules as cross-linker to form a monolayer of DNA-Au particle [Chiu etal 2008]. Electrostatic interaction between the positively charged surface amine groups and negatively charged DNA-Au nanoparticle conjugates allows the self-assembly of a probe nanoparticle monolayer onto the functionalized AlN surfaces under physiological conditions. Results showed that Au nanoparticles can play multiple roles in SAW sensing for probe molecule immobilization, signal amplification, and labelling [Chiu et al 2008].
The substrate can have a significant effect on the acoustic velocities of the AlN SAW devices. For example, Clement et al 2007 have reported that SAW velocity depends on the orientation of Si wafers (4700 m/s for Si (111) and 5100 m/s for Si (100)). AlN films deposited on a $LiNbO_3$ substrate have been reported to form a highly sensitive Love mode sensing device [Kao et al 2003 and 2004]. To increase the operating frequency, it is common to use substrate materials with high acoustic velocities, including sapphire, SiC, diamond, etc. Assouar et al 2002 reported a SAW device on a sapphire substrate, which achieved a velocity of 5536 m/s, compared with 5055 m/s measured on a silicon substrate. The acoustic wave velocity associated with sapphire is very close to that of AlN, which limits the acoustic velocity dispersion. AlN/sapphire is hence an attractive structure for SAW devices operating at very high temperature and high frequency applications in harsh conditions [Aubert et al 2010]. Takagaki et al 2002 fabricated AlN SAW devices on SiC substrates, with a higher-order Rayleigh mode and a frequency of 19.5 GHz, corresponding to a velocity above 7000 m/s. Benetti et al 2005 fabricated (002) AlN/diamond/Si SAW devices, with a velocity of 8200 m/s for Rayleigh mode waves and 10784 m/s for Sezawa mode waves.
SH-SAW has been generated using AlN film with laser micromachined grooves [Xu et al 2004]. This required the formation of grooves with periodicity smaller than half the wavelength of the SAW, thus the longitudinal waves is damped and only the SH-SAW can propagate. The so-formed SH-SAW can be used in a liquid biofludic system without significant signal loss, with a high mass sensitivity of 1.35 ng/ml (Xu et al 2004). However, the process is complicated as it needs precision laser microprocessing of the groove.

5.2 Lamb wave devices

In a manner similar to a SAW device, Lamb wave devices on a membrane structure have been used for biosensing in liquid [Muralt et al 2005]. The wave velocity generated in the Lamb wave is much smaller than those in liquids. This minimizes the dissipation of wave energy into the liquid, thus the Lamb wave sensors can be used in a liquid environment [Wenzel & White 1990, Nguyen & White 2000]. The detection mechanism is based on the relative change in magnitude induced by the perturbation on the membrane, and not on the resonant frequency shift. Therefore the sensitivity of these devices increases as the membrane thickness becomes thinner. Lamb wave devices utilizing the lowest order symmetric Lamb wave (S0) propagating in a highly textured 2 μm thick AlN membrane has been successfully demonstrated by Yantchev and Katardjiev in 2007. AlN based Lamb wave resonators can have very high acoustic velocity, up to 10,000 m/s using the lowest symmetric (S0 mode) Lamb wave device, which also shows a low dispersive characteristic [Ling et al 2010]. The main drawback of the Lamb wave biosensor is that there is a practical limit on the minimum film thickness as the thin film becomes more fragile.
Duhamel et al (2006) prepared an AlN/Si based Lamb wave biosensor, and obtained a sensitivity of 200 cm^2/g. However, due to the thin membrane structure, the temperature

sensitivity is also significant as the AlN/Si Lamb wave device has a non-zero temperature coefficient of frequency (TCF) in the range −20 to −25 ppm/◦C [Wingqvist et al 2009]. Therefore, temperature compensation is normally necessary. Different types of temperature compensation methodology have been proposed for AlN Lamb wave devices [Zuo et al 2010, Lin et al 2010, and Wingqvist et al 2009]. For example, AlN was deposited on P^+ doped silicon (which has a positive TCF of 9 ppm/K) to compensate the temperature effect. The most reported method is to use AlN/SiO_2 composite layer structure (as SiO_2 has a TCF of 85 ppm/K) [Bjurstorm et al 2007, Lin et al 2010]. The Lamb wave resonators with almost zero TCF have been fabricated using a composite AlN/SiO_2 membrane structure with different AlN/SiO_2 thicknesses (see Fig. 14) [Wingqvist et al 2009], with a *Q* factors of around 1400 at a frequency of around 755 MHz.

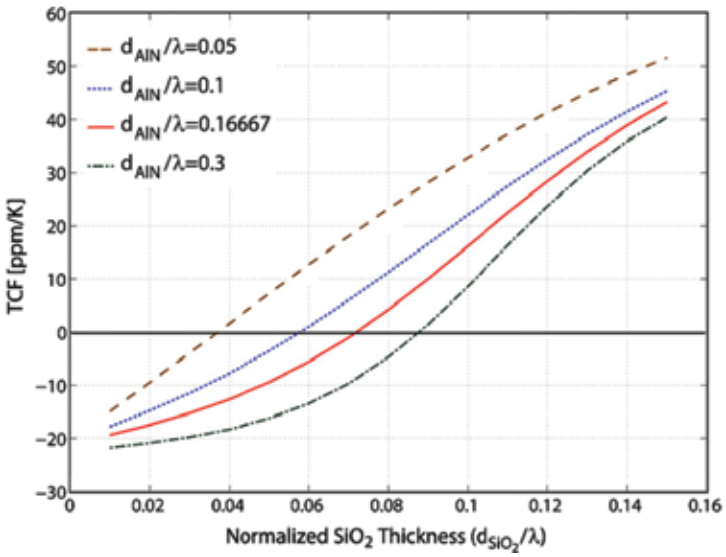


Fig. 14. Temperature coefficient of frequency as a function of the SiO_2 thickness for AlN plates of varying thicknesses normalized to the wavelength (λ) [G. Wingqvist et al, 2009-b]

5.3 AlN FBAR device

Probably, the most of common AlN based acoustic wave biosensor is FBAR structure. Similar to the QCM, an FBAR device (shown in Fig. 16a) consists of a submicron thick piezoelectric film membrane sandwiched between two metallic electrodes [Ruby 2007, Benetti 2005]**.** The frequency shift Δf due to mass loading Δm of an acoustic wave device can be calculated by [Buttry & Ward 1992]

$$\Delta f = \frac{2\Delta m f_o^2}{A\sqrt{\rho\mu}} \tag{2}$$

where A, ρ, μ and f_o are the area, density, shear modulus and intrinsic resonant frequency, respectively. Owing to the much reduced thickness**, t**he FBAR device operates at high frequencies, up to a few GHz, and the attachment of a small target mass can cause a large frequency shift – typically a few MHz. This makes the signal easily detected using simple

electronic circuitry. Figure 15 summarizes the sensitivity range for different types of resonators according to their normal operational frequency ranges [Rey-Mermet at al 2004]. The advantages of the FBAR device includes: (1) the ability to fabricate the device using standard CMOS processing and compatible materials allowing integration with CMOS control circuitry; (2) the significantly reduced size and sample volume. These features together with the intrinsic high sensitivity make the FBAR devices ideal for highly sensitive real time diagnostic biosensor arrays, which provide quantitative results at a competitive cost. However, for the membrane based FBAR design, the membrane fragility and the difficulty in its manufacture are significant issues which have yet to be fully addressed.

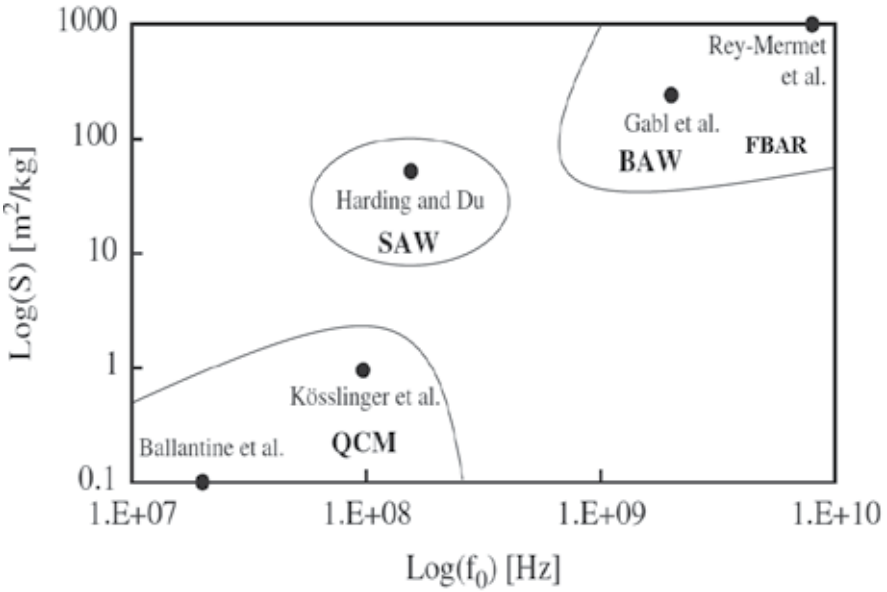


Fig. 15. Sensitivity range for different types of resonators according to their common applied frequency ranges (Rey-Mermet at al 2004)

In addition to the membrane based FBAR structure, there is another common FBAR structure that uses an acoustic mirror deposited between the piezoelectric layer and the substrate (see Fig. 16b). The acoustic mirror is composed of many quarter-wavelength layers of alternating high and low acoustic impedance layers. Due to the high impedance ratio of the acoustic mirror, the acoustic energy is reflected and confined inside the top piezoelectric layer, thus maintaining an excellent resonant bandwidth. This design has a better mechanical robustness and a simpler process control compared with the membrane-based structures. Also cheap substrates, such as glass or plastics can be used, thus the cost can be reduced. Disadvantages for this type of FBAR design is that the process requires thickness and stress control for each layer, increasing the number of the fabrication steps.

There is a third FBAR design which uses a front side etching process [Kang et al 2005]. Initially a sacrificial layer is deposited on the substrate followed by the electrodes and the piezoelectric film depositions. The release of the structure from the substrate is through an air gap made by reactive ion etching of the sacrificial layer. The required selectivity control during the etching process is critical during the fabrication. One disadvantage is the potential liquid trapping inside the gap during biodetection.

An AlN FBAR device has been first reported by Latin et al 1981, and has already been successfully commercialized in the communication industry [Kim et al 2001, Tadigadapa et al 2009], and has also been used as chemical or gas sensors [Benetti, et al 2005]. FBAR biosensors have recently attracted great attention due to their inherent advantages compared with SAW and QCM biosensors: high sensitivity, low insertion loss, high power handling capability and small size [Bjurstrom et al 2004, Kang et al 2005, Loebl et al 2003, Chiu 2007 and 2008]. AlN based FBAR devices have been used to detect carcioembryonic antigen (cEA), a type of glycoprotein associated with breast, colorectal and lung cancer, and the fabricated FBAR device has a frequency of 2.477 GHz, and a sensitivity of 3514 Hz cm^2/ng [Lee et al 2010]. For FBAR, the thickness of the piezoelectric film AlN is comparable with that of the electrode, or bottom layer, being similar to SiO_2 or Si_3N_4. Therefore, the materials for the electrodes and their thickness can influence significantly the performance of the FBAR device.

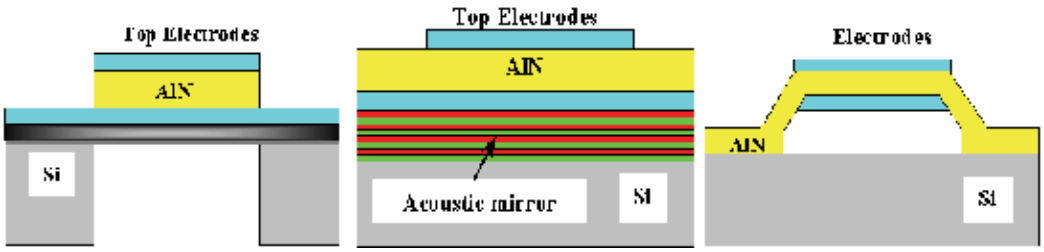


Fig. 16. Types of FBAR resonators: (a) membrane FBAR, (b) air gap FBAR and (c) solid mounted resonator.

For liquid FBAR sensing, there is good reason to deposit an AlN film in which the *c*-axis is inclined relative to the surface normal, thus allowing both longitudinal and shear wave modes to be generated [Weber 2006]. Wingqvist et al. 2007 have fabricated a biochemical sensor based on inclined c-axis AlN for cocaine and heroin detection (see Fig. 17). The FBAR sensor was tested in an immunoassay using avidin/antiavidin detection with a sensitivity of 800 Hz cm^2/ng [Wingqvist et al. 2007]. However, the quality factors was low (100 to 150 for FBAR and 2,000 for QCM) and noise level high, thus the overall detection limit of FBAR is not as good as for QCM devices (detection limit of FBAR was twice as much for a commercial QCM) [Wingqvist et al. 2007]. Wingqvist et al 2009 also used shear mode FBAR devices for multilayer protein sensing, i.e., alternating layers of streptavidin and biotinated BAS, as well as cross-linking of fibrinogen with EDC activation of its carboxyl groups.

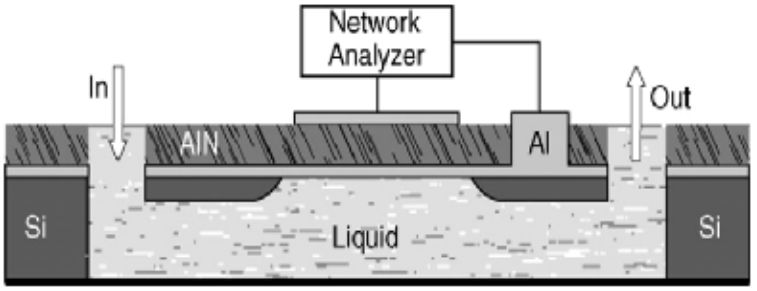


Fig. 17. Schematic picture of lateral FBAR structure comprising the resonator with two electrodes solidly mounted on an acoustic Bragg mirror (from Wingqvist et al 2007a,b)

Another popular method to use FBAR devices in liquid solution is to use lateral field excitation (LFE) of the piezoelectric layer. This requires both signal and ground electrodes being in-plane and parallel on the exposed surface of the AlN film (as can be seen by comparing the conventional longitudinal FBAR electrode design and LFE FBAR design in Fig. 18). A laterally excited AlN thickness shear mode resonator is extremely simple to fabricate and highly sensitive to surface perturbations. The resonator configuration consists of a laterally excited, solidly mounted AlN thin film resonator and the device has been reported to operate stably in biologically equivalent environments such as NaCl in deionized water [Dickherber et al 2008, Corso et al 2007, 2008].

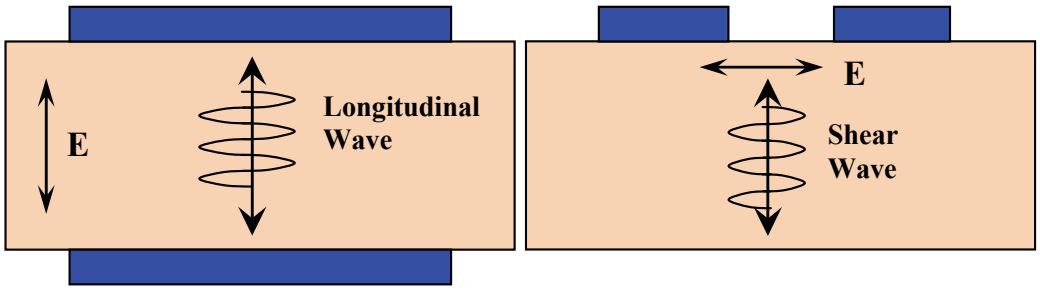


Fig. 18. Comparison of (a) the conventional longitudinal FBAR electrode design; and (b) LFE FBAR design

Xu et al 2010 have proposed a new FBAR of high quality factors *Q*s operating in liquid media. The FBAR is made of a suspended circular shaped AlN ring sandwiched between the top and bottom Au electrodes, which can be excited in a contour mode (Fig. 19). By exciting in its radial-extensional mode, the resonator experiences the shear viscous damping instead of the squeeze damping, which significantly alleviates the acoustic energy dissipated in the contacting liquid. By having a low motional resistance or coupling with liquids, the contour mode FBAR achieved *Q*s up to 189, which is more than 13-19 times than conventional FBAR device in liquids and the resonator was used to test an aptamer—thrombin binding pair, with a mass resolution of 1.78 ng cm^{-2} [Xu et al 2010].

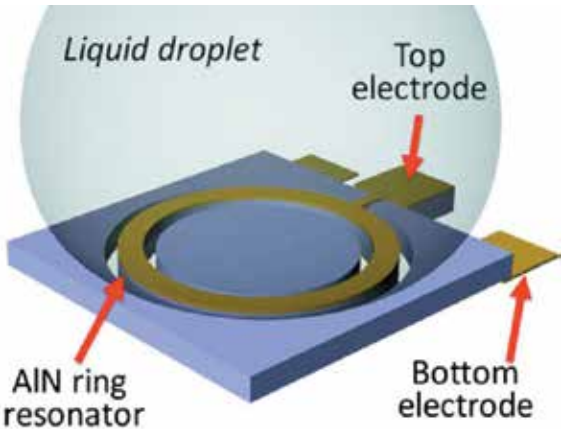


Fig. 19. Schematic figure of the contour-mode AlN FBAR biosensor contacting with a liquid droplet [Xu et al 2010]

Although FBAR based biosensor exhibit a high sensitivity and good resolution, there are some issues to be addressed. For example, they normally have high acoustic wave attenuation and low quality factor due to potential thin film material defects and thin membranes. Other issues include the sensor packaging and the effect of high frequency on biochemistry [Wingquist et al 2007 a and b]. Zhang & Kim 2005 have reported that the second harmonic mode of wave can be excited at a frequency about twice of the fundamental resonance, thus the FBAR using the second harmonic longitudinal mode can have a high Q factor and a low dissipation of acoustic energy into the liquid. Similar to Lamb wave device, the temperature stability of the FBAR is a critical issue, and a composite layer of AlN/SiO_2 is a common method that can be employed to compensate for the temperature effect.

6. AlN film for microfluidic applications

In an AlN based SAW device, the interaction between the longitudinal acoustic wave and liquid droplets can be used to create acoustic streaming which can establish a stable streaming pattern with a double vortex (see Fig. 20). This SAW streaming induces an efficient mixing and agitation within the droplets, which can be utilised to produce good micromixers [Fu et al 2007, Fu et al 2010]. When an RF voltage is applied to the IDTs on a piezoelectric film, the water droplet becomes deformed from its original shape (following the Rayleigh angle) with an increased leading edge and a decreased trailing edge contact angle. After surface hydrophobic treatment, the liquid droplets can be pumped forward, with the droplet movement being a combination of rolling and sliding, which is also dependent upon the power applied and the droplet size.

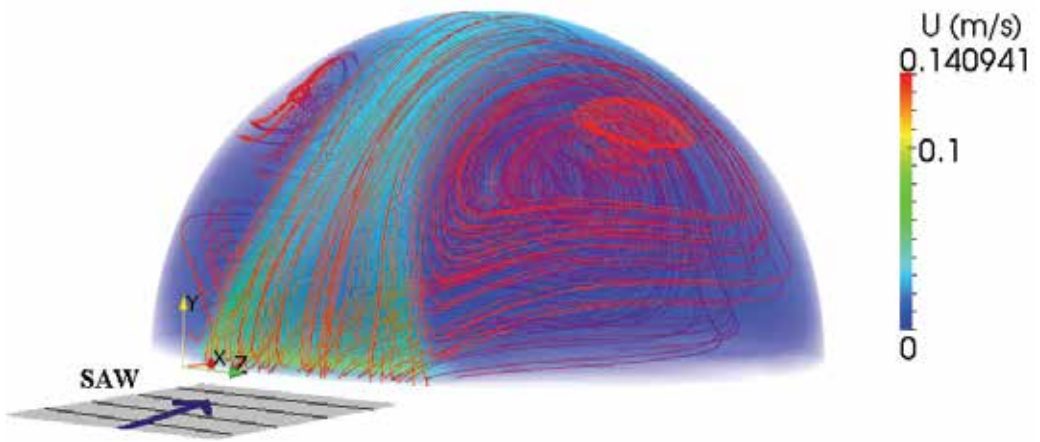


Fig. 20. Numerical 3D illustration showing the droplet SAW interaction leading to 3D complex flow patterns due to SAW energy attenuation and Reynolds stresses formation which in turn producing effective steady force acting in the fluid body "(Courtesy from Mr. Alghane Mansuor)

When the RF power applied to the IDT of an AlN SAW device is sufficiently high, tiny liquid droplets will be ejected from the surface. Ejection of small particles and liquid has many applications ranging from inkjet printing, fuel and oil ejection and bio-technology.

Flexural plate waves or Lamb waves have also been proposed for pumping, agitating and enhancing biochemical reactions [Nguyen & White 1999], with the principle that fluid motion via the travelling flexural wave in an AlN membrane can be used for the transport of liquids. The potential applications include a micro total analysis system (μTAS), cell manipulating systems, and drug delivery systems [Meng et al 2000]. However, there are few studies on microfluidic applications based on the AlN acoustic wave devices, which is a potentially very interesting research topic.

7. Future trends for AlN devices for lab-on-a-chip

The elements required for operating detection as part of a lab-on-a-chip system include: (1) transportation of liquids such as blood or biofluids containing DNA/proteins into an area on which probe molecules have been pre-deposited, (2) mixing/reaction of the extracted DNA or proteins with oligonucleotide or the antibody binders, and (3) detection of an associated change in the physical, chemical, mechanical or electrical signals. Thin film based acoustic wave devices can be used to fabricate lab-on-chip bio-detection systems, which combine the functions of microdroplet transportation, mixing and bio-detection.

Device integration at the device, wafer and system level is critical issue for the lab-on-chip fabrication. Wafer level integration of AlN FBAR device with CMOS fabrication has been reported by Campanella et al 2008. It has electrical connection between FBAR and CMOS. Sharma et al 2010 have fabricated a shear mode AlN solidly mounted resonator microfluidic sensor, which is fully IC compatible, integrating a SMR sensor chip with a PDMS microfluidic channel system. The c-axis AlN film has been used to generate shear mode wave and the AlN SMR device operated at the 1.2 GHz range, with a Q factor of 100 in water.

Acoustic wave technologies can be integrated with other technologies, such as the surface plasma resonance (SPR) method [Homola et al 1999]. SPR sensor technology has been commercialized and SPR biosensors have become an important tool for characterizing and qualifying biomolecular interactions. A combination of SAW microfluidics and SPR sensing would appear to be sensible for both microfluidic and detection functions. A potential problem is that the surface temperature change induced by acoustic excitation may cause changes in refractive index, which is used for SPR sensor detection. A pulse mode SAW signals can be used to minimize this effect. Acoustic wave microfluidic devices can also be combined with liquid or gas chromatography, which can be used to identify the protein or molecules by mass spectroscopy [Sokolowski et al 2006]. Integration of a SAW with optical methods enables the simultaneous qualification of biological soft layers formed on the sensor surface under different density, viscosity, thickness and water content.

For digital microfluidics, there is a need to precisely and continuously generate liquid droplets. AlN acoustic wave technology can be used for the ejection of liquid droplets, but it is rather difficult to precisely control the micro-droplet generation. A potential technology to overcome the drawbacks is to combine electrowetting-on-dielectrics (EWOD) [Li et al 2009] with SAW-microfluidics. In the past ten years, EWOD technology has been successfully developed to dispense and transport nanolitre to microlitre bio-samples in droplet form at the exact volume required [Fair 2007]. However, one of the weaknesses is that EWOD technology does not provide efficient micro-mixing, and requires the integration of other technologies e.g. CMOS to realise bio-reaction and biosensing. A novel idea is to integrate the thin films based SAW devices with the EWOD device to form lab-on-a-chip equipped

with well developed functionalities of droplet generation, transportation by EWOD, mixing and biosensing using SAW technology [Li et al 2010].
Acoustic wave devices can easily be integrated with standard CMOS technology. Dual SAW or FBAR devices can be fabricated next to each other, so that the neighbouring devices can be used as a sensor-reference combination. One of the devices without pre-deposited probe molecules can be used as a reference, while the other one with probe molecules can be used to sense. Using such a combination, the errors due to temperature drift or other interference on the sensing measurement can be minimized. Multi-sensor arrays can easily be prepared on a chip and a judicious selection of different immobilized bio-binders enables the simultaneous detection of multiple DNA or proteins, leading to accurate diagnosis of a disease or detection of multiple diseases in parallel. The creation of these cost-effective sensor arrays can increase the functionality in real time and provide parallel reading functions.
Currently, one limitation of acoustic wave device applications is that they require expensive electronic detection systems, such as network analyzers. A final product aimed at the end user market must be small, portable and packaged into a highly integrated cost effective system. The detection of a resonant frequency can be easily realized using standard oscillator circuits which can measure the sensor losses based on a portable device. The required purposely built electronics for acoustic wave sensing are being developed, but at present they are still bulky and heavy. Fabrication of portable thin film based acoustic wave detection devices is also promising and will enable the system size to be minimised along with reducing the power consumption. A wireless RF signals can be used to remotely power and control/monitor physical, chemical and biological quantities by using acoustic wave devices, without requiring a directly wired power supply. Currently for a lab-on-chip device, sample pre-treatment, purification and concentration, as well as a good interface between the user and the integrated sensing system also need to be developed. A simple, robust, cheap packaging method is also critical for commercialization.

8. Summary

AlN films have good piezoelectric properties and a high electro-mechanical coupling coefficient, and are hence a promising technology for the fabrication of fully automated and digitized microsystems with low cost, fast response, reduced reagent requirement and precision. In this chapter, recent development on preparation and application of AlN films for acoustic wave-based microfluidics and bio-sensors has been discussed. The microstructure, texture and piezoelectric properties of the films are affected by sputtering conditions such as plasma power, gas pressure, substrate material and temperature as well as film thickness. AlN acoustic wave devices can be successfully used as bio-sensors, based on a biomolecular recognition system. Among these biosensors, surface acoustic wave, Lamb wave and film bulk acoustic resonator devices using inclined films are promising for applications in highly sensitive bio-detection systems for both dry and liquid environments. The acoustic wave generated on the AlN acoustic devices can also induce significant acoustic streaming, which can be employed for mixing, pumping, ejection and atomization of the fluid on the small scale depending on the wave mode, amplitude and surface condition. An integrated lab-on-a-chip diagnostic system based on these thin film based acoustic wave technologies has great potential, and other functions such as droplet creation, cell sorting, as well as precise bi-detection can be obtained by integration with other advanced technologies.

9. Acknowledgement

YQ Fu and CS Cherng would like to acknowledge the financial support from International Joint Projects from Royal Society of Edinburgh and National Science Council of Taiwan. The authors would like to acknowledge financial support from the Institute of Integrated Systems, Edinburgh Research Partnership in Engineering and Mathematics (ERPem). They also would like to acknowledge support from Royal Academy of Engineering-Research Exchanges with China and India Awards, Royal Society-Research Grant, Carnegie Trust Funding, and China-Scotland Higher Education Partnership from British council. JKL would like to acknowledge the support of the EPSRC under grant EP/F063865, EP/D051266 and EP/F06294. AJW and YL acknowledge support from The EU (GOLEM STRP 033211) and BBSRC (RASOR BBC5115991). AJW, MD and YQF would like to acknowledge the financial support from Innovative electronic Manufacturing Research Centre (IeMRC) co-ordinated by Loughborough University through the EPSRC funded flagship project SMART MICROSYSTEMS (FS/01/02/10).

10. References

Akiyama, M. T., Kamohara, K. Kano, et al, 2008, Appl. Phys. Lett. 93: 021903.

Akiyama, M., K. Nagao, N. Ueno, et al., 2004. Vacuum, 74: 699-703.

Akiyama, M., N. Ueno, H. Tateyama, et al. 2005. J. Mater. Sci., 40: 1159-1162.

Assouar M. B., O. Elmazria, L. Brizoual, et al, 2002. Diam. Relat. Mater. 11: 413-417.

Aubert T., O. Elmazria, B. Assouar, et al, Appl. Phys. Lett., 96 (2010) 203503.

Auger M. A., L. Vazquez, O. Sanchez, et al, 2005. J. Appl. Phys., 97: 123528.

Baek, J., J. Ma, M.F. Becker, J.W. Keto and D. Kovar, Thin Solid Films 515 (2007), p. 7096.

Ballantine, D. S., R. M. White, S. J. Martin, A. J. Ricco, E. T. Zellers, G. C. Frye, H. Wohltjen, 1996, Acoustic Wave Sensors, Theory, Design and Physical-Chemical Applications, Academic Press.

Barie, N. and M. Rapp; 2001. Biosensors & Bioelectron. 16: 978.

Benda V., M. Cernik and D. Stepkova, Microelectron. J. 29 (1998), p. 695.

Benedic, F., M. B. Assouar, P. Kirsch, P, et al. 2008. Diam. Relat. Mater., 17: 804-808.

Benetti, A., D. Cannata, F. Di Pietrantonio, et al. 2006. Thin Sold. Films, 497: 304-308.

a-Benetti, A., D. Cannata, F. Di Pietrantonio, et al. 2005. IEEE Trans. Ultra. Ferro. Freq. Control. 52, 1806-1811.

b-Benetti M., APL. 87 (2005) 173504

Bjurstrom, J., D. Rosen, I. Katardjiev, V. M. Yanchev and I. Petrov; 2004. IEEE Trans. Ultrason. Ferroelectric and Freq. Control; 51: 1347-1353.

Brizoual, Le L., O. Elmazria O, F. Sarry, M. El Hakiki, A. Talbi, P. Alnot, 2006. Ultrasonics. 45: 100-103.

Brizoual, Le, L. and Elmazria, O, 2007. Diamond. Realt. Mater., 16: 987-990.

Buttry, D. A. and M. D. Ward, 1992. Chem. Rev. 92: 1355.

Caliendo, C., P. Imperatori, E. Cianci, 2003. Thin Solid. Films, 441: 32-37.

Caliendo C., P. Imperatori, Appl. Phys. Lett. 83 (2003) 1641

Campanella et al 2008. IEEE Device Lett., 29 (2008) 28-30.

Kar J P, Bose G, Tuli S, Dangwal A, Mukherjee S, J. Mater. Engng Perf. 18 (2009) 1046-1051.

Chen, Q. M. and Q. M. Wang. 2005. Appl. Phys. Lett. 86: 022904.

Cheng, C. C. Chen, Y.C. Horng R. C. et al. 1998. J. Vac. Sci. Technol., 16: 3335-3340.
Cherng, J. S., C. M. Lin, T.Y. Chen, 2008. Surf. Coat. Technol., 202: 5684-5687.
Cherng, J. S. and D.S. Chang, 2008. Thin Solid Films, 516: 5293-5295.
Cherng, J. S., T.Y.Chen, C. M. Lin, 2009. Ferroelectric, 380: 89-96.
Cheung, T. T. and C. W. Ong, 2004. Diamond Relat Mater., 13: 1603-1608.
Chiu, C S, H. M. Lee, C. T. Kuo, et al. 2008. Appl. Phys. Lett. 93: 163106.
Chiu, C S. 2008. Appl. Phys. Lett., 93: 163106.
Chiu, K H, J. H. Chen, H. R. Chen et al, 2007. Thin Solid Films, 515: 4819-4825.
Chono, K, N. Shimizu, Y. Matsu, J. Kondoh, S. Shiokawa. 2004. Jap. J. Appl. Phys. 43: 2987.
Chou, CH; Lin, YC; Huang, JH, et al. 2006. Integrat. Ferro., 80: 407-413.
Chung, G S, and K.S. Kim 2007, Electron. Lett. 43 (2007), p. 832
Clement, M., L. Vergara, J. Sangrador, et al, 2004. Ultrasonics, 42 : 403-407.
Clement, M., E. Iborra, J. Sangrador, et al. 2003. J. Appl. Phys., 94: 1495-1500.
Clement, M., J. Olivares, E. Iborra, et al., 2009. Thin Solid Films. 517: 4673-4678.
Corso, C. D., A. Dickherber, , W. D. Hunt. 2007. J. Appl. Phys. 101: 054514.
Cote, G. L., R. M. Lec, M. V. Pishko, 2003. IEEE Sens. J., 3: 251-266.
Dickherber, A., C. D. Corso, W. D. Hunt, 2008. Sens Actuat., A 144: 7-12.
Duhamel, R., L. Robert, H. Jia, et al, 2006. Ultrasonics, 44: e893-e897.
Elmazria, O., V. Mortet, M. El Hakiki, et al. 2003. IEEE Trans. Ultrasonic. Ferro. Freq. Cont. 50 : 710-715.
Elmazria, O., Sergei Zhgoon, Laurent Le Brizoual, Frédéric Sarry, Dmitry Tsimbal, and Mohammed Abdou Djouadi, Appl. Phys. Lett., 95, 233503 _2009.
Engelmark, F., G. F. Iriarte and I. V. Katardjiev. 2002. J. Vac Sci Technol B. 20 : 843-848.
Engelmark, F., G. Fucntes, I. V. Katardjiev, et al . 2000. J. Vac Sci Technol A. 18: 1609-1612.
Fair, R. B. 2007. Microfluid Nanofluid, 3: 245-281
Fardeheb-Mammeri, M., B. Assouar, O. Elazria, et al, 2008. Diam. Relat. Mater., 17: 1770-1774.
Franke, T. A. and A. Wixforth, 2008. Chem Phys Chem, 9: 2140-2156.
Fu, Y. Q., J.K. Luo, X. Du, A.J. Flewitt, Y. Li, A. Walton, W.I. Milne, 2010. Sens. Actuat. B. 143: 606-619.
Fu, Y. Q., X.Y.Du, J.K.Luo, A.J.Flewitt, M.I.Milne, 2008. IEEE Sens., 1-3: 478-83.
Galipeau, D.W., P. R. Sory, K. A. Vetelino, R. D. Mileham, 1997, Smart. Mater. Struct. 6: 658.
Gao XD, E.Y. Jiang, H.H. Liu, G.K. Li, W.B. Mi, Z.Q. Li, P. Wu and H.L. Bai, Phys. Status Solidi (a) 204 (4) (2007), p. 1130.
Gizeli, E. 1997. Smart. Mater. Struct. 6: 700.
Grate, W. J., S. J. Martin, R. M. White, 1993. Anal Chem., 65: 940.
Grate JW, 2000, Chem. Rev, , 100 (7), pp 2627–2648
Hakiki M. E., O. Elmazria, P. Alnot, 2007. IEEE transactions on Ultrasonics, Ferro. Freq. Control, 54: 676-681.
Hara, M., J. Kuypers, T. Abe, et al. 2005. Sens. Actuat., A 117: 211-216.
Hirata S., K. Okamoto, S. Inoue, et al. 2007. J. Solid State Chem. 180: 2335-2339.
Si-Hong Hoang and Gwiy-Sang Chung, Microelectronic Engineering, Volume 86, Issue 11, November 2009, Pages 2149-2152.

Homola, J., S. S. Yee, G. Gauglitz, 1999. Sens. Actuat., B 54: 3-15.

Hong, H. S. and Chung, G. S. 2009. J. Korean Phys. Soc. 54: 1519-1525.

Hoummady, M., A. Campitelli, W. Wlodarski, 1997, Smart. Mater. Struct. 6: 647.

Huang, C. L., K. W. Tay, L. Wu. 2005. Solid State Electro. 49: 219-225.

Huang, C. L.; K. W. Tay, L. Wu. 2005. Jap. J. Appl. Phys., 44: 1397-1402.

Iborra, E., M. Clement, J. Sangrador, et al, 2004. IEEE Trans. Ultras. Ferroelectr. Freq. Control, 51: 352-358.

Imura M et al, 2010], Kiyomi Nakajima, Meiyong Liao, Yasuo Koide, Hiroshi Amano, Journal of Crystal Growth 312 (2010) 368–372.

Iriarte, G F,, F. Engelmark and I.V. Katardjiev, J. Mater. Res. 17 (2002), p. 1469.

Iriarte, G.F. 2003. J. Appl. Phys., 93: 9604-9609.

Ishihara, M, K. Yamamoto, F. Kokai, et al. 2000. Vacuum, 59: 649-656.

Jacoby, B. and M. Vellekoop, 1997. Smart. Mater. Structu., 6: 668-679.

Jagadish, C. and S. J. Pearton, 2006. Zinc oxide bulk, thin films and naoctstructures: processing, properties and applications, Elseveier.

Jasinki J, Z. Liliental-Weber, Q. S. Paduano, D. W. Weyburne, 2003. Appl. Phys. Lett., 83: 2811.

Ji, X. H., S. P. Lau, G. Q. Yu, et al, 2004. J. Phys. D. 37: 1472-1477.

Josse, F., F. Bender, R. W. Cernosek. 2001. Anal. Chem. 73: 5937.

Kang, Y. R., S. C. Kang, K. K. Park, Y. K. Kim, S.W. Kim and B. K. Ju. 2005. Sens. Actua. A117: 62.

Kao, K. S., C.C. Cheng, Y.C. Chen, Y. H. Lee, 2003. Appl. Phys., A76: 1125-1127.

Kar J P, Bose G, Tuli S, Dangwal A, Mukherjee S, J. Mater. Engng Perf. 18 (2009) 1046-1051.

Kern, R. S., L. B. Rowland, S. Tanaka, et al. 1998. J. Mat. Res. 13: 1816-1822.

Khan, F. A. et al. 2006. Mater. Sci. Engng, B 95: 51-4.

Kim, E. K., T. Y. Lee, H. S. Hwang, et al.; 2006. Superlatt. & Microstr. 39: 138.

Kim, S. H., J. H. Kim, D. D. Park, G. Yoon, 2001. J. Vac. Sci. Technol., B 19: 1164-1168.

Kirsch, P., M. B. Assouar, O. Elmazria, et al. 2006. Appl. Phys. Lett., 88: 223504.

Kovacs, G., G.W. Lubic, M. J. Vellekoop, A. Venema, 1992. Sens. Actuat., A 43: 38-43.

Kovacs. G. and M. Venema. 1992. Appl Phys Lett. 61: 639.

Kumagai Y., T. Yamane and A. Koukitu, J. Crystal Growth 281 (2005), p. 62.

Kumar, K. S. A. and S. M. Chen, 2008. Analytical Letters 41: 141–58.

Kuznestsova, L. A. and W.T. Coakley, 2007. Biosensors and Bioelectronics 22: 1567-1577.

Lange, K., B. E. Rapp, M. Rapp, 2008. Anal. Bioanal. Chem. 391: 1509-1519.

Lanz R. and P. Muralt, IEEE Trans. Ultrason. Ferr. Freq. Control 52 (6) (2005), p. 936.

Lee, C. K., S. Cochran, A. Abrar, K. J. Kirk, F. Placido, 2004. Ultrasonics, 42: 485-490.

Lee, H.C., J. Y. Park, K. H. Lee, et al. 2004. J. Vac. Sci. Technol. B, 22: 1127-1133.

Lee, J. B., M. H. Lee, C. K. Park, et al. 2004. Thin Solid Films, 447: 296-301.

Lee S. H., K.H. yoon, J. K Lee, 2002. J Appl. Phys., 92: 4062-4069.

Lee, S. H, J. K. Lee, K. H. Yoon. 2003. J. Vac. Sci. Technol., A, 21: 1-5.

Lee T Y, Song J T, Thin Solid Films, 2010, In press.

Li, Y., B.W. Flynn, W. Parkes, et al., Conference of ISSDERC 2009, in press.

Lim, W. T., B. K. Son, D. H. Kang, C. H. Lee, 2001. Thin Solid Film, 382: 56-60.

Lin, Z. X., S. Wu, R. Y. Ro, et al. 2009. IEEE. Trans. Ultraonics. Ferroelec. Freq. Control, 56: 1246-1251.

Ling C M, T T Yen, Y, J. Lai, et al IEEE Trans Ultras. Ferro, Freq. Control, 57 (2010) 524-532.

Lindner, G., 2008. J. Phys. D. 41: 123002.

Liu, Z. F., F.K. Shan, Y.X. Li, B.C. Shin and Y.S. Yu, 2003. J. Crystal Growth 259: 130.

Liu, J. M., N. Chong, H. L. W. Chan, et al. 2003. Appl. Phys., A, 76: 93-96.

Lu Y F, Ren Z M, Chong, TC Cheng, BA, Chow SK, wang J P., J Appl. Phys. 87 (2000) 1540.

Lucklum, R. and P. Hauptmann. 2003. Meas. Sci. Technol. 14: 1854.

Luginbuhl, P., S. D. Collins, G. A. Racine, M. A. Gretillat, N. F. De Rooij, K. G. Brooks, N. Setter. 1997. J. MEMS, 6: 337-346.

Luo, J. K., Y.Q. Fu, Y. F. Li, X.Y. Du, A.J. Flewitt, A. Walton, W. I. Milne, 2009. J. Micromech. Microeng., 19: 054001.

Marx, K. A., 2003. Biomacromolecules. 4: 1099.

Mchale, G. 2003. Meas. Sci. Technol. 14: 1847.

Meng, A. H., N.T. Nguyen and R.M. White, 2000. Biomed. Microdev. 2: 169–174

Mortet, V., M. Nesladek, K. Haenen, et al. 2004. Dia. Relat. Mater., 13: 1120-1124.

Muralt P, 2008. J. Am. Ceramic. Soc., 91: 1385-1396.

Muralt, P., N. Ledermann, J. Baborowski, et al. 2005. IEEE Trans. Ultrasonics, Ferroelectr. Frequen. Control. 52: 2276.

Murochim, N., M. Sugimoto, Y. Matui, J. Kondoh, 2007. Jap. J. Appl. Phys., 46: 4754.

Naik, R. S., R. Reif, J.J. Lutsky and C.G. Sodini, 1999. J. Electrochem. Soc. 146: 691.

Newton, M. I., M. K. Banerjee, T. K. H. Starke, S. M. Bowan, G. McHale, 1999. Sensor & Actuat. 76: 89.

Nguyen, N. T. and R. T.White. 1999. Sens. & Actuat. 77: 229-36.

Okamoto, K., S. Inoue, T. Nakano, et al. 2008. Thin Solid Films, 516: 4809-4812.

Okamoto M., M. Yamaoka, Y.K. Yap, M. Yoshimura, Y. Mori and T. Sasaki, Diamond Relat Mater 9 (2000), p. 516.

O'Hanlon JF, A User's Guide to Vacuum Technology (2nd. ed.), John Wiley and Sons, Hoboken, NJ (1989).

Paci, B., A. Generosi, V. R. Albertini, et al. 2007. Sens. Actuat., A 137: 279-286.

Pandey D.K., R.R. Yadav, Temperature dependent ultrasonic properties of aluminium nitride, Applied Acoustics 70 (2009) 412–415.

Pandey DK, Yadav RR, Appl. Acoustics, 70 (2009) 412-415.

Pearton, S. J., D. P. Norton, K. Ip, Y. W. Heo, T. Steiner, 2005. Prog. Mater. Sci. 50: 293.

Renaudin, A., P. Tabourier, V. Zhang, J.C. Camart, C. Druon. 2006. Sensor & Actuat. B113: 387.

Rey-Mermet, S., J.Bjurstrom, D.Rosen and I.Petrov. 2004. IEEE Trans. Ultrason. Ferroelectric and Freq. Control; 51: 1347.

Ruby, R. 2007. IEEE Ultrasonics Symp. Proc. 1-6 : 1029-1040.

Sanchez, G., A. Wu, P. Tristant, et al. 2008. Thin Solid. Films, 516: 4868-4875.

Sano, A., Y. Matsui, S. Shiokawa, 1998. Jap. J. Appl. Phys. 37: 2979.

Saravanan, S., E. Berenschot, G. Krijnen, M. Elwenspoek, 2006. Sens. Actuat., A130-131: 340-345.

Sharma G., L. Liljeholm, J. Enlund, J. Bjurstorm, I. Katardjiev, K. Hjort, Sens. Actuat., A 159 (2010) 111-116.

Sheng, T. Y., Z.Q. Yu, GJ Collins, 1988. Appl. Phys. Lett., 52: 576.

Shih, W. C., R. C. Huang, Y. K., Peng, et al. 2009. Ferroelectr., 380: 20-29.

Shiokawa, S., J. Kondoh, 2004. Jap. J. Appl. Phys., 43: 2799-2802.

Shiokawa, S., Y. Matsui and T. Morizum. 1989. Jpn. J. Appl. Phys. 28: 126.

Sritharan, K., C. J. Strobl, M. F. Schneider, A. Wixforth, 2006. Appl. Phys. Lett., 88: 054102.

Strobl, C. J., Z. Guttenberg, A. Wixforth; 2004. IEEE Trans. Ultrasonics, Ferrelectric and freq. Control. 51: 1432.

Sudhir G.S., H. Fujii, W.S. Wong, C. Kisielowski, N. Newman and C. Dieker et al., Appl Surf Sci 127–129 (1998), p. 471.

Takagaki, Y., P.V. Santos, E. Wiebicke, et al, 2002. Appl. Phys. Lett. 81: 2538-2540.

Tan, S. S., M. Ye, A. G. Milnes, 1995. Solid State Electro., 38: 17.

Tanosch, K., et al, Sens. Actuato, 2006. A132: 658-663.

Teles F R R, L. P. Fonseca, 2008. Talanta, 77: 606-623.

Toegl, A., J. Scribe, A. Wixforth, C. Strobl, C. Gauer, Z.V. Guttenburg. 2004. Anal. Bioanal. Chem. 379: 69.

Toegl, A., R. Kirchner, C. Gauer, A. Wixforth, 2003. J. Biomed. Technol., 14: 197.

Tseng, W. K., J. L. Lin, W. C. Sung, S. H. Chen, G. B. Lee; 2006. J. Micromech. Microeng. 16: 539.

Vashaei Z., T. Aikawa, M. Ohtsuka, H. Kobatake, H. Fukuyama, S. Ikeda and K. Takada, Journal of Crystal Growth, 311, 2009, 459-462.

Vellekoop, M. J., 1998. Ultasonics. 36: 7.

Vergara, L., M. Clement, E. Iborra, et al. 2004. Diam. Relat. Mater., 13: 839-842.

Weber, J., W. M. Albers, J. Tuppurainen, M. Link, R. Gabl, W. Wersing, M. Schreiter, 2006. Sensors & Actuat. A128: 84-88.

H. Windischmann, Crit. Rev. Solid State Mater. Sci. 17 (1992), p. 547.

Wingquist, G., J. Bjurstrom, L. Liljeholm, et al, 2007. Sens. Actuat., B123: 466-473.

Wingquist, G., J. Bjurstrom, A.C. Hellgren, I. Katardjiev, 2007. Sens. Actuat., B127: 248-252.

Wingqvist G. V. Yantchev, Katardjiev, Sens. Actuat. A., 148 (2008) 88-95.

a-Wingqvist G., Anderson, H., Lennartsson, Weissbach T, Yantchev V., Lyoyd A, Spet Z., Bios. Bioelectron., 24 (2009) 3387-3390.

b-Wingqvist G, L Arapan, V Yantchev and I Katardjiev, J. Micromech. Microeng. 19 (2009) 035018. Wixforth, A., C. Strobl, C. Gauer, A. Toegl, J. Sciba, Z. V. Guttenberg, 2004. Anal. Biomed. Chem., 379: 982.

Wixforth. A. 2004. Superlattices & Microstruct. 33: 389.

Wohltjen, H., et al. 1997. Acoustic Wave Sensor—Theory, Design, and Physico-Chemical Applications, Academic Press, San Diego:39.

Wu, L., S. Wu, H. T. Song, 2001. J. Vac. Sci. Technol., A19: 167.

Wu, S., R. Ro, Z. X. Lin, M S. Lee. 2008. J. Appl. Phys. 104: 064919.

Wu, S., Y.C. Chen, Y.S. Chang. 2002. Jap. J. Appl. Phys., 41: 4605-4608.

Wu, H. P., L. Z. Wu, S. Y. Du, 2008. J. Appl. Phys., 103: 083546.

Wu, S., R. Y. Ro, Z. X. Lin, et al. 2009. Appl. Phys. Lett., 94: 092903.

Xu J. Thakur J S., Zhong F., Ying H., Auner G W, J Appl. Phys. 96 (2004), 212-217.

Xu, J., J. S. Thakur, G. Hu, et al. 2006. Appl. Phys. A, 83: 411-415.

Yanagitani, T. and M. Kiuchi; 2007. J. Appl. Phys. 102: 044115.

Yantchev and Katardjiev 2007 IEEE Trans. Ultrason. Ferroelectr. Freq. Control 54 87–95

Yang, P. F., S. R. Jian, S. Wu, et al. 2009. Appl. Surf. Sci., 255: 5984-5988.

Zhang, D., J. H. Edgar, 2005. Mater Sci. Engng R, 48: 1-46.

Zhang H, Kim E S, 2005, J MEMS, 14, 699-706.

13

Application and Exploration of Fast Gas Chromatography - Surface Acoustic Wave Sensor to the Analysis of *Thymus* Species

Se Yeon Oh[1,2], Sung-Sun Park[3], and Jongki Hong[1]
[1]*College of Pharmacy , Kyung Hee University, Seoul 130-701,*
[2]*Aroma Analytical Laboratory, KOSMO NF Co., Ltd. , Seoul 502-5,*
[3]*Graduate School of Cultural Industry, Sungshin Women's University, Seoul 136-742, South Korea*

1. Introduction

Thymus is an aromatic and medicinal plant of increasing importance in horticulture and economics. *Thymus* is a genus of about 350 species of aromatic perennial herbs in the family Lamiaceae (mint family), and native to Europe, North Africa and Asia. Its essential oil has found diverse applications in pharmacy and medicine. Its volatile phenolic oil, for example thymol and carvacrol, has been reported to have antibacterial, antimycotic, antioxidative, and mammalian age delaying properties. Also *thymus* serves as a flavoring agent for a variety of food products and used as an antiseptic agent for its antimicrobial properties [1-3]. The content of essential oil varies drastically with climate, time of harvest and storage conditions [4-6].

For many years, GC and GC-MS have been used widely for the characterization of the volatile aroma components in *thymus* species. However, traditional GC method requires several routine isolation procedures including solvent extraction [7, 8], steam distillation [9, 10], and simultaneous distillation extraction [11]. These methods involve excessive manipulation of the sample, a very costly, time-consuming procedure, are limited in aroma correlation, and do not allow on-line measurements which may lead to inadequate results. Recently, headspace solid-phase microextraction (HS-SPME) as a successful solvent-free sampling technique has been introduced for purpose of aroma analysis [12, 13]. Especially, aroma analysis demands rapid and simple procedure, because new aroma components may arise from chemical and biochemical reactions promoted by heat and oxidation conditions. Also, aromas are usually composed of complex mixtures of many volatiles, human sensory evaluation by trained panelists is important in aroma analysis. However, it has many limitations which involve a very expensive, time-consuming procedure, and subjectiveness of expert. Therefore, the development in analytical method which provides rapid, simple, low-cost procedure and the clear relationship between their sensory impacts is one of the most desirable subjects in aroma chemistry.

A few years later, a new technique, based on the fast gas chromatography combined with uncoated high quartz surface acoustic wave sensor (GC/SAW, zNose) [14-17] appeared to be one of the suitable methods. Its principle has many similarities comparative to the human

perception system. The advantages of GC/SAW include simplicity, real-time detection of volatiles, non-destructive, portability and lower costs in comparison to a portable GC-MS. Fast GC/SAW permits quantification and pattern recognition by fragrance pattern, called a VaporPrint derived from the frequency of a SAW sensor. Moreover, good sensitivity at the high picogram to nanogram level makes it possible to detect sensitive aroma materials quantitatively [15, 18]. The method validation of GC/SAW and adaptability to a variety of applications were reported in our previous paper [18].
Statistical analysis methods including principal component analysis (PCA) have been successfully applied for the quality control and classification of various herbal medicines or aroma plants. PCA analysis and hierarchical clustering analysis (HCA) as pattern recognition analysis involves the discrimination of chromatographic data of herbal extracts or aroma plants with similar species [19, 20]. Pattern recognition analysis based on the chromatographic data can predict and evaluate the quality control of aroma plants.
The aim of this study is to show the application and exploration of the developed GC/SAW methodology to the analysis of the volatile aroma composition profiles among *thymus* species in order to introduce this advantageous alternative analytical technique in pharmacy, medicine, and horticulture.

2. Experimental

2.1 Materials

Thymus (*T. quinquecostotus, T. quinquecostotus var. japonica, T. mongolicus, T. serpyllum*) plants grown nearby Pocheon city, Kyunggi-Do in South Korea were collected by sunny day sampling in September 2005. The geographical origins of *T. quinquecostotus* and its variety are from South Korea. *T. mongolicus* is in Northeastern Asia, and *T. serpyllum* is in Europe. The medicinal plant material consists of stem and leaves which are raw, elapsed for 5 days at 5 °C and air-dried for 13 days or 16 months. All standard chemicals of analytical grade were purchased from Sigma-Aldrich (St. Louis, Mo, USA) and Tokyo Kasei (Nihonbashi, Tokyo Japan). Organic solvents of a chromatographic grade were obtained from J .T. Baker. The commercially available carboxen-divinylbenzene-polydimethylsiloxane (CAR-DVB-PDMS) SPME fiber (film thickness, 50/30 μm) was purchased from Supelco (Bellefonte, PA, USA) and used.

2.2 GC/SAW (zNose) description

GC/SAW (4100 vapor analyzer, Electronic Sensor Technology, New Bury Park, USA) composed with the fast gas chromatograph and surface acoustic wave sensor is used to detect vapors of the volatile organic compounds. The GC/SAW is especially sensitive to low concentrations.
The uncoated piezo-electric quartz crystal SAW (Surface Acoustic Wave) sensor [14] represents a new class of GC detector. The specificity of the uncoated SAW sensor is based upon the temperature of the crystal surface and the vapor pressure characteristics of the condensates. At a given crystal temperature, analytes with dew points closer to the crystal temperature will interact and be detected better than those with dew points well above the SAW temperature. The high Q crystal is in contact with a thermoelectric element, which controls the temperature for cooling during vapor adsorption and for heating during cleaning of the crystal and operates by maintaining highly focused and resonant surface acoustic waves of 500 MHz on its surface.

2.3 GC/SAW analytical conditions and procedure

About 1.0 g of each air-dried *thymus* sample was weighed into a 40-ml glass vial sealed with a screw cap containing a Teflon/silicone septa. The capped vial was allowed to equilibrate with the headspace in the vial under the 60% humidity and 24 °C for 1 h just before analysis. GC/SAW utilizes two steps to analyze vapors: the sampling process and the injection process. The headspace vapor is swept at 30 ml/min via a pump into the inlet, then the vapor passes through the valve where the compounds are adsorbed onto the Tenax trap inside the system. Switching the valve to the injection process causes helium gas to flow backwards through the Tenax trap and onto the column. During the injection process, the Tenax trap is heated rapidly to 200 °C to desorb the material. Details of this procedure were reported in our previous paper [18]. GC column was heated from 32 °C to 120 °C at a rate of 3 °C/s and the sampling time was 1 s. Helium (99.999%) was used as a carrier gas at 3.2 ml/min (0.053 ml/s). 6% cyanopropyl phenyl polydimethylsiloxane (DB-624, J&W Scientific, Folsom, CA, USA, 1 m x 0.25 mm i.d., 0.25 μm film thickness) fused silica capillary column was used. The set-up temperatures were at 30 °C for sensor, 130 °C for inlet port, and 110 °C for valve. Triplicate measurements per vial were carried out. All analytical procedures were completed within 30 s. The shorter total time-to-result per sample allows several replicated analyses of a sample.

2.4 Headspace solid-phase microextraction (HS-SPME)

About 2.0 g of air-dried *thymus* sample was placed in 25-ml vial sealed with an aluminum cap containing a Teflon /silicone septa. The capped vial was kept to equilibrate under the humidity of 60% and 24 °C for 1 h before HS-SPME sampling. The carboxen-divinylbenzene-polydimethylsiloxane (CAR-DVB-PDMS) SPME fiber (film thickness, 50/30 μm) was used because it was most efficient among the various types of fiber for most volatile organic compounds [12, 21]. The SPME fiber was exposed to the headspace above the *thymus* sample vial at 24 °C for 1 h. After adsorption, the SPME fiber was retracted from the sample vial and immediately inserted into the injection port of the GC-MS where thermal desorption was performed at 240 °C for 1 min.

2.5 GC-MS analysis

The sample analysis was carried out with a Thermoquest-Finnigan ion trap GC–MS (Austin, Texas, USA) equipped with 6% cyanopropyl phenyl polydimethylsiloxane (DB-624, J&W, 30 m x 0.25 mm i.d., 1.4 μm film thickness) and a Hewlett-Packard 6890 Series GC system with an Agilent 5973N Mass Selective Detector (Agilent Technologies, Wilmington, DE, USA) equipped with 5% phenyl polydimethylsiloxane (Ultra 2 column, Agilent, 25 m x 0.25 mm i.d., 0.33 μm film thickness). The oven temperature was initially maintained at 50 °C for 3 min and then programmed to 220 °C for 5 min at a rate 5 °C/min. Injector and transfer line, and quadrupole temperatures were set at 240 °C, 250 °C, and 150 °C, respectively. Helium (99.999%) was used as a carrier gas at 1.0 ml/min. The sample was injected under split mode (split ratio 1:30). The mass spectrometer was run in the electron impact (EI) mode with electron energy at 70eV, scanning the 50.0-400.0 amu. The ion source temperatures of ion trap GC-MS and quadrupole GC-MS were maintained at 200 °C , 230 °C, respectively. Triplicate measurements per vial were carried out.

2.6 Data analysis for pattern recognition

Data transformation for pattern recognition was performed using MS Excel. Fifteen components were chosen based on the corresponding GC/SAW and HS-SPME-GC-MS profile. Especially, fifteen components such as α-pinene, camphene, β-myrcene, *p*-cymene, γ-terpinene, terpinolene, *cis*-sabinene hydrate, camphor, borneol, α-terpineol, thymol methyl ether, thymoquinone, thymol, β-caryophyllene and β-bisabolene were found as characteristic components in GC/SAW and HS-SPME-GC-MS profile. And then the response of each peak was applied from the triplicate measurements. Finally, pattern recognition techniques have been used for the discrimination of the materials. Principal component analysis (PCA) is a pattern recognition technique and statistical analysis. PCA was also carried out using MVSP. 3.1 version (Kovach Computing Service, Anglesey, Wales) in order to classify *thymus* species.

3. Results and discussion

3.1 Identification of volatile herbal aroma compounds for air-dried for 13 days of thymus species by GC/SAW

By using fast GC/SAW, volatile herbal aroma profiles for *thymus* species were obtained. The materials sequentially exit from the column and they land and stick on the SAW sensor. When an analyte adsorbs on the surface of the sensor, the frequency of SAW sensor is altered, which affects the detection signal in direct proportion to the amount of condensate. Fig. 1(A)-(E) shows chromatograms of volatile aroma compounds for air-dried for 13 days of *thymus* species. The area of each peak is correlated to its concentration and is expressed in frequency counts (Cts). The identification of each aroma compounds shown by GC/SAW was carried out by comparison with authentic standards and GC–MS analysis and their relative proportions (% total amounts) are summarized in Table 1. The herbal aroma components of *thymus* species consist mostly of monoterpene hydocarbons (α-pinene, camphene, β-myrcene, and terpinolene), oxygenated monoterpenes (*cis*-sabinene hydrate, camphor, borneol, α-terpineol, and thymoquinone), a monoterpene phenol (thymol), monoterpene phenol precursors (*p*-cymene and γ-terpinene), a monoterpene phenol derivative (thymol methyl ether) and sesquiterpenes (β-caryophyllene and β-bisabolene). The grouping of compounds has an important meaning as responsible for the characteristic aroma of *thymus*.

Fifteen compounds were identified, especially in *T. quinquecostotus* species, which are from Jeju and Mt Gaya in South Korea, the characteristic, distinctive components such as *p*-cymene (26.4%, 24.2%), γ-terpinene (10.3%, 10.5%), and active thymol (29.0%, 33.1%) were constituted 65.7%, 67.8% of the total amounts, respectively. In addition, oxygenated monoterpenes such as *cis*-sabinene hydrate (0.6%, 0.7%), camphor (6.6%, -), borneol (4.7%, 5.7%) and thymoquinone (11.6%, 12.5%) were constituted 23.5% and 18.9% of the total amounts as the secondly most abundance, respectively. Lesser amounts of monoterpene hydrocarbons: α-pinene (0.9%, 0.8%), camphene (0.9%, -), β-myrcene (5.3%, 3.3%), terpinolene (1.2%, 2.4%) were constituted 8.3% and 6.5% of the total amounts, respectively, and β-caryophyllene (2.5%, 6.8%) as sesquiterpene was also found. *T. quinquecostotus var. japonica* species is a variety of *T. quinquecostotus* and its geographical origin is Ulreung island in South Korea. *p*-Cymene (15.8%), γ-terpinene (9.0%), and thymol (33.5%) were also found as characteristic components and constituted 58.3% of the total amounts. Also, oxygenated monoterpenes (32.3%), monoterpene hydrocarbons (5.5%), and sesquiterpenes (3.9%) were found.

Peak No.	t_R (s)	Compound	*T. quinquecostotus* (%)				*T. quinquecostotus var. japonica* (%)		*T. mongolicus* (%)		*T. serpyllum* (%)	
			Jeju, Korea		Mt. Gaya, Korea		Ulreung island, Korea		Northeastern asia		Europe	
			13 D	16 M	13 D	16 M	13 D	16 M	13 D	16 M	13 D	16 M
a	11.14	α -Pinene	0.9(4.86)	1.1(5.72)	0.8(5.89)	1.5(5.43)	0.9(7.63)	1.4(1.49)	–	–	–	–
b	11.92	Camphene	0.9(7.02)	0.5(1.41)	–	–	0.5(7.66)	0.5(5.02)	1.0(4.53)	1.6(10.23)	–	–
c	13.28	β -Myrcene	5.3(8.24)	4.5(4.10)	3.3(4.62)	2.9(3.49)	3.1(6.05)	3.3(1.85)	9.5(2.02)	14.3(14.74)	7.0(4.38)	3.9(0.69)
d	14.50	*p*-Cymene	**26.4**(1.48)	**29.0**(1.19)	**24.2**(7.69)	**30.2**(0.30)	**15.8**(7.57)	**25.1**(1.58)	2.8(5.17)	3.6(18.63)	**15.8**(4.52)	**30.4**(11.42)
e	15.52	γ -Terpinene	**10.3**(7.17)	**9.2**(1.08)	**10.5**(7.31)	**10.6**(0.11)	**9.0**(8.65)	**12.4**(7.99)	–	–	**4.2**(6.12)	–
f	16.30	Terpinolene	1.2(6.34)	1.0(4.36)	2.4(9.54)	0.7(7.69)	1.0(3.90)	0.4(4.30)	3.4(3.00)	2.3(7.49)	2.0(6.38)	–
g	17.44	*cis*-Sabinene hydrate*	0.6(5.83)	1.3(8.65)	0.7(5.63)	2.4(10.67)	0.4(4.23)	0.3(3.18)	2.1(2.30)	1.5(5.59)	2.4(2.28)	0.5(26.33)
h	19.02	Camphor	6.6(3.08)	2.4(9.01)	–	–	–	–	–	–	1.4(9.27)	0.8(0.10)
i	19.82	Borneol	**4.7**(3.00)	**3.6**(5.85)	**5.7**(1.74)	**3.1**(5.85)	**18.4**(1.24)	**9.6**(1.71)	**21.0**(4.94)	**20.0**(6.70)	5.8(3.77)	6.3(3.40)
j	20.56	α -Terpineol	–	–	–	–	–	–	**41.8**(6.26)	**35.9**(7.78)	–	8.1(1.64)
k	21.52	Thymol methyl ether*	–	–	–	–	–	–	–	–	**12.7**(5.19)	**17.5**(18.27)
l	22.48	Thymoquinone	**11.6**(4.41)	**16.9**(3.83)	**12.5**(5.33)	**18.6**(3.56)	**13.5**(5.81)	**18.4**(3.86)	–	–	5.3(7.65)	17.1(12.57)
m	24.34	Thymol	**29.0**(2.28)	**27.5**(1.27)	**33.1**(5.47)	**25.8**(4.80)	**33.5**(4.08)	**24.1**(0.73)	–	–	**34.1**(3.08)	**1.1**(37.71)
n	27.48	β-Caryophyllene	2.5(1.79)	3.0(6.11)	6.8(10.06)	4.2(2.68)	1.5(2.39)	1.9(2.20)	11.3(4.48)	12.9(12.96)	3.5(0.94)	1.5(17.67)
o	30.08	β-Bisabolene*	–	–	–	–	2.4(2.40)	2.6(1.84)	7.1(0.96)	7.9(10.19)	**5.8**(8.55)	**12.8**(10.94)

*Tentatively identified by comparision of GC/MS.: not detected.

% figures are their relative proportions as percent of total amount (Cts), %(RSD), n=3.

The characteristic compounds are indicated in bold.

Column: DB-624 (6% cyanopropyl phenyl polydimethylsiloxane, 1 m x 0.25 mm x 0.25 μm) fused silica capillary column was used.

Table 1. Composition and identification of aroma components for air-dried for 13 days and 16 months of *thymus* species by GC/SAW

Generally, thymol, phenolic monoterpene, defines the essential oil quality because of its active pharmacological properties. It was reported that γ-terpinene and *p*-cymene are the precursors of phenolic monoterpenes such as thymol and carvacrol in *T. vulgaris* by Granger and Passet [22]. It was also proposed that γ-terpinene assumes an important role in the aromatization whose product is *p*-cymene. They reported that *p*-cymene is the precursor of thymol by hydroxylation in *T. vulgaris* L. plant [23]. Moreover, the concentration of *p*-cymene and γ-terpinene are found to vary in coincidence with the variation in their corresponding phenolic monoterpene products [24] and γ-terpinene decreases its concentration and *p*-cymene increases in the essential oil [25].
In *T. mongolicus* species (the origin of Northeastern Asia), active phenolic monoterpene (thymol) and its corresponding monoterpene hydrocarbon precursor (γ-terpinene) were not found, whereas oxygenated monoterpenes, such as borneol (21.0%), α-terpineol (41.8%), were identified as characteristic predominant compounds and constituted 62.8% of the total amounts. *T. serpyllum* species (the origin of Europe) contain almost the same chemical components with those of *T. quinquecostotus*. Especially, a monoterpene phenol derivative (thymol methyl ether: 12.7%) was only found as characteristic component. It was tentatively identified by comparison of its GC-MS data, because its reference standard is not commercially available.

3.2 Discrimination of thymus species using VaporPrint image based on GC/SAW

By comparing the relative contents, chemical composition for characteristic components of air-dried for 13 days of *T. quinquecostotus* species which are originated from Jeju and Mt. Gaya in South Korea were almost same: the active monoterpene phenol (thymol: 29.0%, 33.1%), its corresponding precursors (*p*-cymene: 26.4%, 24.2%, γ-terpinene: 10.3%, 10.5%) and oxygenated monoterpenes (borneol: 4.7%, 5.7%, thymoquinone: 11.6%, 12.5%). It is interesting to note that the components found in the same species of different geographical origin in same country are almost same. Also, the proportion of such components is same.
The GC/SAW provides a visually recognizable fragrance pattern (VaporPrint image) derived from the frequency of SAW sensor. This image is created by transforming the time variable to a radial angle with the beginning and end of the analysis. This image transfers the olfactory response to a visual response [18]. These fragrance images are a useful for comparing delicate differences of various *thymus species* for species identification. The fragrance patterns for air-dried for 13 days of *thymus* species are shown in Fig. 1(A')-(E'). As shown in Fig. 1(A') and (B'), these same species present almost same fragrance patterns.
In its variety species (the origin of Ulreung island in South Korea), the most components are similar in chemical compositions between *original* and its *variety*, specific components which are *p*-cymene (*original* species: Jeju 26.4%; Mt. Gaya 24.2%, its *variety* species: 15.8%), borneol (*original* species: Jeju 4.7%; Mt. Gaya 5.7%, its *variety* species: 18.4%) were shown a different compositions. Also, β-bisabolene (2.4%) was found additionally in *variety* species. Therefore, such differences seem to result in a small differences in fragrance pattern (Fig. 1(A') and (C')).
T. serpyllum species (the origin of Europe) contain almost same chemical components with those of *T. quinquecostotus*, but show a substantial different chemical compositions: active thymol (*T. serpyllum* : 34.1%, *T. quinquecostotus* : Jeju 29.0%; Mt. Gaya 33.1%), active thymol precursors (*T. serpyllum* : 20.0%, *T. quinquecostotus* : Jeju 36.7%; Mt. Gaya 34.7%), oxygenated monoterpenes, (*T. serpyllum* : 14.9%, *T. quinquecostotus* : Jeju 23.5%; Mt. Gaya 18.9%), and sesquiterpenes (*T. serpyllum* : 9.3%, *T. quinquecostotus* : Jeju 2.5%; Mt. Gaya 6.8%) were found

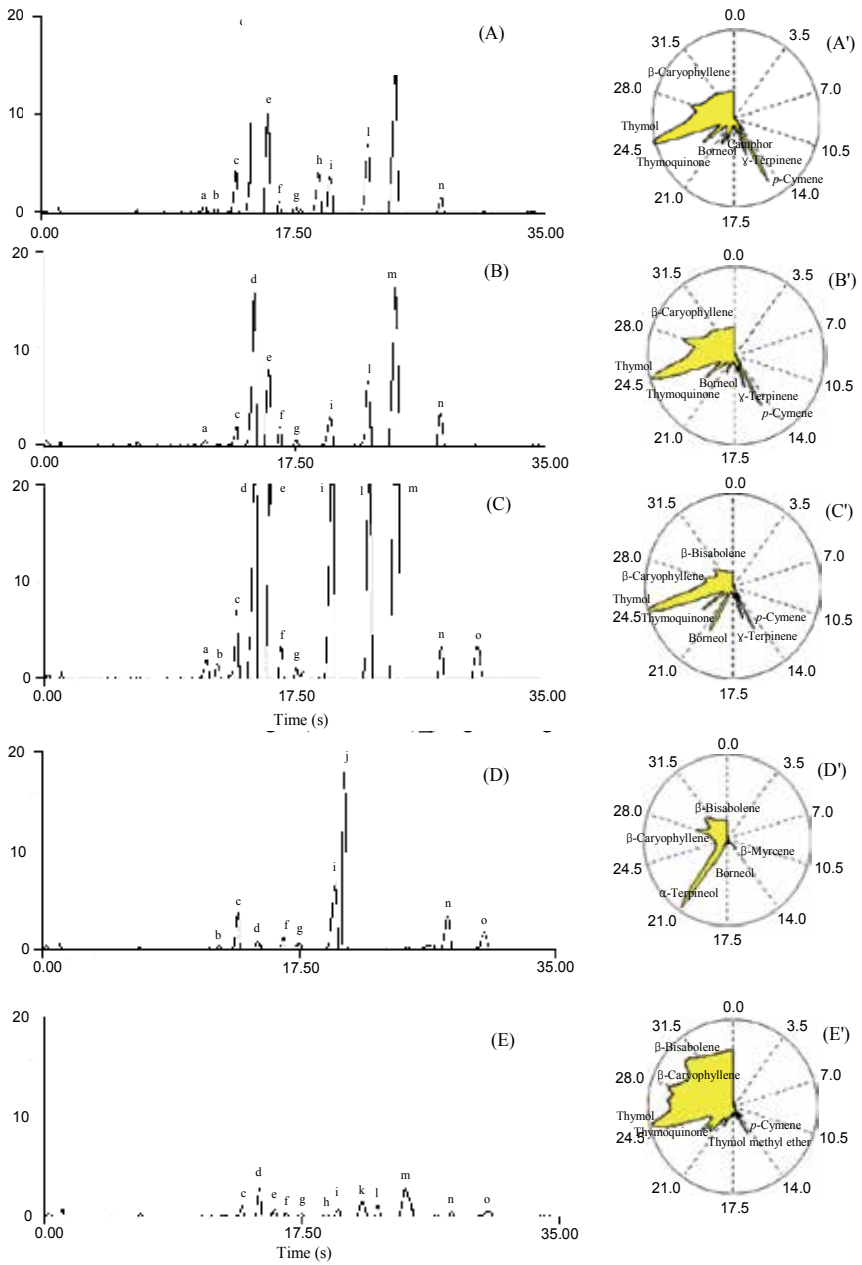


Fig. 1. Comparison of chromatograms for various *thymus* species by GC/SAW. (A) *T. quinquecostotus,* Jeju, Korea; (B) *T. quinquecostotus*, Mt. Gaya, Korea; (C) *T. quinquecostotus var. japonica*, Ulreung island, Korea; (D) *T. mongolicus*, Northeastern Asia; (E) *T. serphyllum*, Europe; (A´)-(E´) their corresponded fragrance patterns using VaporPrint.
Peak identities: a, α-Pinene; b, Camphene; c, β-Myrcene; d, *p*-Cymene; e, γ-Terpinene; f, Terpinolene; g, *cis*-Sabinene hydrate; h, Camphor; i, Borneol; j, α-Terpineol; k, Thymol methyl ether; l, Thymoquinone; m, Thymol; n, β-Caryophyllene; o, β-Bisabolene.

, respectively. Therefore, distinctive differences seem to result in a substantial differences in fragrance pattern (Fig. 1(E')).
As a result, it turned out that each species has own characteristic fragrance pattern owing to its own chemical compositions and its own characteristic fragrance patterns are conducive to species identification.

3.3 Comparison of GC/SAW and HS-SPME-GC-MS method for air-dried for 13 days of thymus species

The compositions of aroma compounds of *thymus* species extracted by HS-SPME using CAR/DVB/PDMS fiber and then analyzed by GC-MS are presented in Table 2. Their GC-MS total ion chromatograms are shown in Fig. 2. The alphabetic numbers of peaks shown in Fig. 2 correspond to the numbers indicated in GC/SAW chromatograms (Fig. 1). These results are comparable to those given by GC/SAW, including the characteristic components and chemical composition. Fifty-four compounds were detected by HS-SPME-GC-MS.

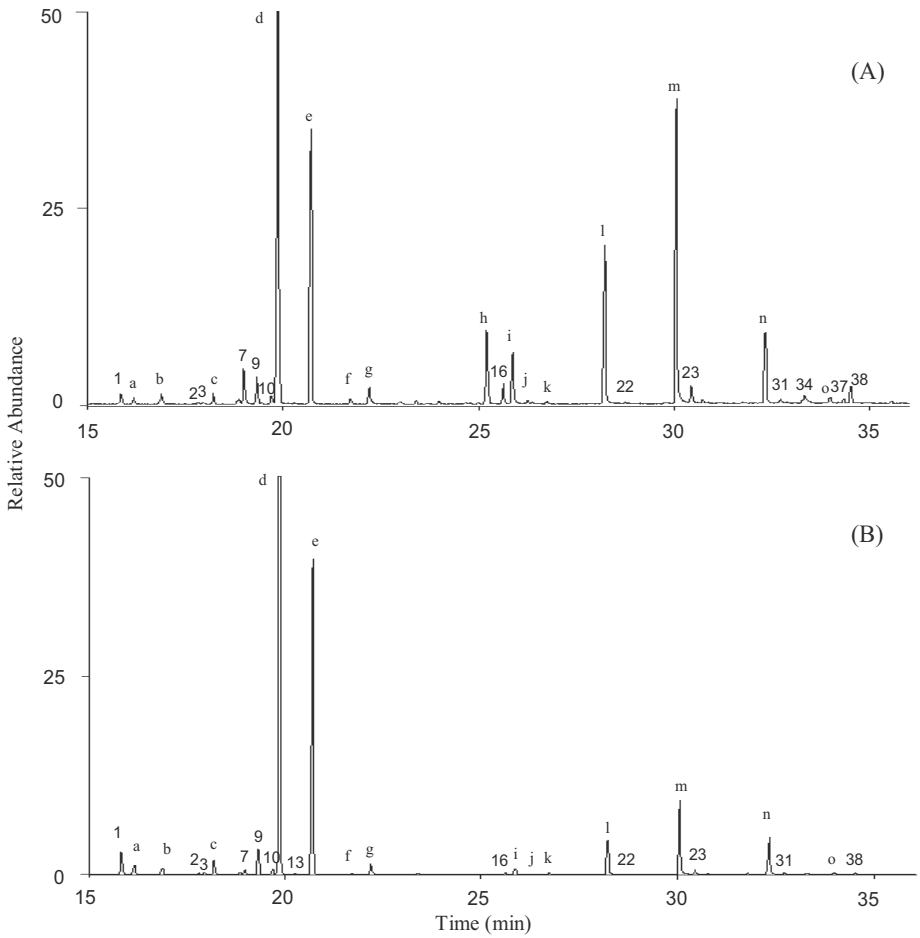


Fig. 2. Comparison of chromatograms for various *thymus* species by HS-SPME-GC-MS. (A) *T. quinquecostotus,* Jeju, Korea; (B) *T. quinquecostotus*, Mt. Gaya, Korea; (C) *T. quinquecostotus var. japonica*, Ulreung island, Korea; (D) *T. mongolicus*, Northeastern Asia; (E) *T. serphyllum*, Europe.

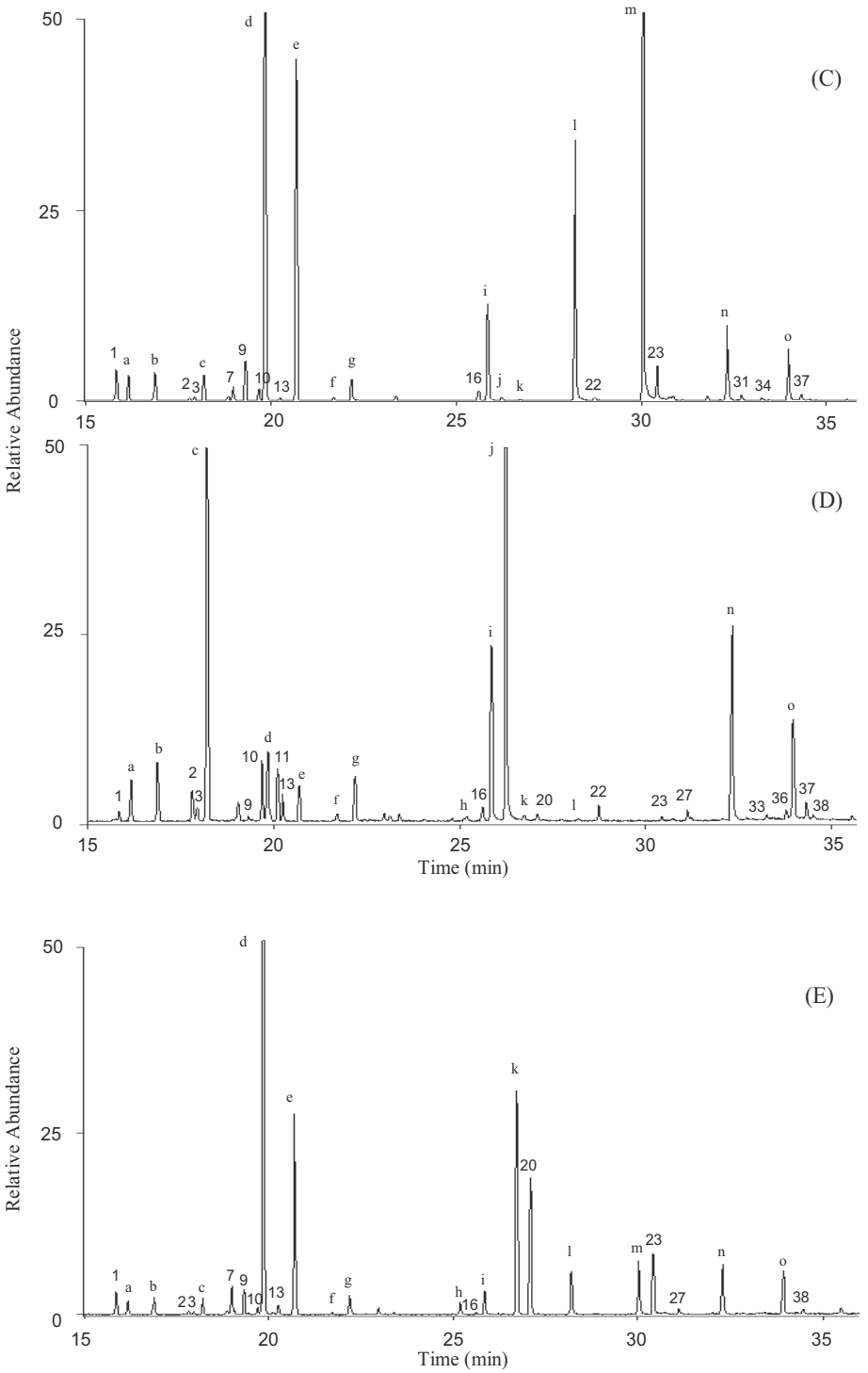


Fig. 2. Continued

As shown in Table 3, characteristic components and analytical tendency for air-dried for 13 days of *thymus* species detected by GC/SAW and HS-SPME-GC-MS are similar, but the abundance ratios between these two methods are different. In *T. quinquecostotus* species, monoterpene phenols (thymol, carvacrol) and monoterpene phenols precursors (*p*-cymene, γ-terpinene) were also dominant (76.0%) like GC/SAW (66.8%). But, the amounts of total monoterpene phenols are 11.7% by HS-SPME-GC-MS, 31.1% by GC/SAW and monoterpene phenols precursors are 64.3% by HS-SPME-GC-MS, 35.7% by GC/SAW. In comparison of *T. quinquecostotus var. japonica* species and its *original* species, these results show similar analytical tendency with GC/SAW. For instance, *p*-cymene decreases from **46.3**% to 22.2% by HS-SPME-GC-MS and from **25.3**% to 15.8% by GC/SAW (an average value is indicated in bold) (Table 1 and 2). While oxygenated monoterpene increases from 11.3% to 17.6% by HS-SPME-GC-MS and from 21.2% to 32.3% by GC/SAW. In *T. mongolicus* species, the characteristic and dominant components are borneol and α-terpineol in similar to the results by GC/SAW. But, the amounts of borneol and α-terpineol are 8.0% and 31.0% by HS-SPME-GC-MS; 21.0% and 41.8% by GC/SAW, respectively (Table 1 and 2). Interestingly, the significant amount of thymol methyl ether and carvacrol methyl ether which are not almost found in other species were found in *T. serpyllum* species. The amounts are 20.5% by HS-SPME-GC-MS and 12.7% by GC/SAW.

3.4 Compositions of volatile herbal aroma compounds of thymus species of elapsed for 16 months by GC/SAW

The compositions of volatile herbal aroma compounds of *thymus* species of elapsed for 16 months extracted by GC/SAW are presented in Table 1. As a result, according to elapse for 16 months, there are slightly differences in the relative quantities of their characteristic constituents than those of 13 days: in *T. quinquecostotus* species and its variety, active thymol precursor *p*-cymene increases from **25.3**% to **29.6**% and from 15.8% to 25.1%, respectively, whereas active thymol decreases from **31.1**% to **26.7**%, 33.5% to 24.1%, respectively (an average value is indicated in bold). However, for *T. serpyllum* species, there appears distinctive differences compared to those dried for 13 days: *p*-cymene also increases from 15.8% to 30.4% and thymol substantially decreases from 34.1% to 1.1%, while thymol precursor γ-terpinene is not found and sesquiterpene β-bisabolene increases 2 times. In *T. mongolicus* species, α-terpineol as the characteristic and dominant component also slightly decreases from 41.8% to 35.9%.

3.5 Principal component analysis for GC/SAW and HS-SPME-GC-MS responses of thymus species

Fig. 3(A) shows principal component analysis (PCA) for GC/SAW responses of air-dried for 13 days and 16 months of *thymus* species. PCA were carried out using MVSP. 3.1 version in order to classify the *thymus* species. As a result, a good classification among four different species: *T. quinquecostotus*, *T. quinquecostotus var. japonica*, *T. mongolicus*, and *T. serpyllum* was obtained. Additionally, the classification of same species which are from different geographical origin in same country, classification of original species and its variety for *T. quinquecostotus*, classification as an air-drying term for 13 days and 16 months for *T. quinquecostotus* showed good results. In these results, it turned out that discrimination of various *thymus* species by using VaporPrint image based on GC/SAW are very clear, which was reproducible data points shown in Fig. 3(A).

Peak No.	Compound	*T. quinquecostotus*				*T. quinquecostotus var. japonica*		*T. mongolicus*		*T. serpyllum*	
		Jeju, Korea		Mt. Gaya, Korea		Ulreung island, Korea		Northeastern Asia		Europe	
		13 D	16 M	13 D	16 M	13 D	16 M	13 D	16 M	13 D	16 M
1	α-Thujene	0.5(43.49)	1.1(56.74)	2.0(15.79)	1.5(16.10)	1.5(9.42)	1.2(18.05)	0.4(13.58)	0.4(25.70)	1.3(15.09)	0.7(5.01)
a	α -Pinene	0.3(37.48)	0.8(77.41)	0.9(14.34)	0.7(22.82)	1.2(9.25)	0.9(24.64)	1.7(15.02)	1.3(27.01)	0.8(10.52)	0.4(7.44)
b	Camphene	0.7(67.69)	1.1(85.93)	0.6(31.30)	0.3(28.20)	1.4(18.31)	1.1(28.02)	2.6(13.67)	2.1(25.67)	1.3(19.92)	0.3(4.70)
2	Sabinene	< 0.1	0.1(59.31)	0.1(30.99)	0.1(16.68)	0.1(18.85)	< 0.1	1.2(16.96)	0.9(21.75)	0.2(28.19)	< 0.1
3	β-Pinene	< 0.1	0.2(73.76)	0.2(16.07)	0.1(23.88)	0.2(20.86)	0.2(24.77)	0.5(22.47)	0.5(27.00)	0.2(23.38)	0.1(4.89)
4	1-Octen-3-ol	–	2.2(40.04)	–	1.4(22.21)	–	1.2(11.25)	–	0.2(10.22)	–	2.0(13.08)
5	3-Octanone	–	0.3(53.59)	–	0.2(26.58)	–	0.1(14.84)	–	1.5(10.16)	–	0.4(13.13)
c	β-Myrcene	0.5(25.79)	1.7(17.33)	1.2(12.58)	1.5(4.93)	1.1(4.97)	2.2(3.00)	**24.8**(8.59)	**23.5**(11.52)	0.9(4.80)	0.4(10.13)
6	3-Octanol	–	0.1(41.30)	–	0.1(23.45)	–	0.1(12.17)	–	0.3(15.29)	–	0.2(9.74)
7	α-Phellandrene	1.9(24.48)	0.2(32.05)	0.4(18.93)	0.2(9.36)	0.6(14.01)	0.3(8.64)	–	–	1.7(19.86)	< 0.1
8	δ-3-Carene	–	0.1(50.19)	–	0.1(13.68)	–	0.1(14.99)	–	–	–	< 0.1
9	α-Terpinene	1.4(36.43)	2.7(28.75)	2.2(22.64)	2.5(7.38)	1.8(5.92)	2.6(6.66)	0.1(16.30)	0.1(30.69)	1.4(18.89)	0.5(7.56)
10	Limonene	0.4(45.29)	0.9(35.81)	0.4(13.16)	0.9(13.86)	0.5(24.61)	0.9(10.27)	2.3(23.00)	2.4(9.28)	0.4(19.80)	0.6(3.34)
d	*p*-Cymene	**38.1**(33.00)	**32.7**(36.98)	**54.5**(23.41)	**39.6**(12.64)	**22.2**(18.11)	**21.3**(12.27)	3.2(14.68)	2.5(5.47)	**39.9**(6.70)	**43.2**(1.31)
11	1,8-Cineol	–	–	–	–	–	–	2.1(14.04)	2.1(19.48)	–	0.1(4.09)
12	*cis*-β-Ocimene	–	–	–	–	–	–	–	0.1(18.48)	–	0.1(3.85)
13	*trans*-β-Ocimene	–	0.1(33.44)	< 0.1	< 0.1	0.1(9.59)	0.1(4.41)	1.0(15.18)	0.7(11.57)	0.5(16.85)	0.4(3.32)
e	γ-Terpinene	**13.9**(34.11)	**16.2**(22.84)	**22.1**(9.03)	**15.7**(3.10)	**13.7**(13.09)	**13.8**(4.66)	1.3(14.61)	1.2(11.52)	11.4(17.20)	2.9(5.33)
14	*trans*-Sabinene hydrate	–	1.3(46.89)	–	1.1(21.89)	–	1.0(13.10)	–	2.0(6.65)	–	0.5(20.68)
f	Terpinolene	0.3(32.70)	0.4(57.43)	< 0.1	0.3(22.47)	0.1(9.98)	0.3(10.42)	0.2(10.12)	0.2(11.34)	0.1(23.26)	0.1(7.02)

Table 2. Composition and identification of aroma components for air-dried for 13 days and 16 months of *thymus* species by HS-SPME-GC-MS

Table 2. Continued

Peak No.	Compound	*T. quinquecostotus*				*T. quinquecostotus var. japonica*		*T. mongolicus*		*T. serpyllum*	
		Jeju, Korea		Mt. Gaya, Korea		Ulreungdo, Korea		Northeastern Asia		Europe	
		13 D	16 M	13 D	16 M	13 D	16 M	13 D	16 M	13 D	16 M
g	cis-Sabinene hydrate	1.0(51.74)	0.3(49.71)	1.0(16.72)	0.3(26.10)	1.1(15.47)	0.3(11.25)	1.8(7.31)	0.7(9.59)	1.3(19.89)	0.5(20.58)
15	Linalool	–	–	–	–	–	–	–	0.2(13.57)	–	–
h	Camphor	4.5(41.59)	2.7(63.50)	–	–	–	< 0.1	0.3(5.47)	0.2(5.53)	0.8(9.78)	1.1(7.45)
16	α-Terpinen-4-ol	1.1(45.81)	0.6(47.48)	0.1(10.29)	0.6(18.66)	0.4(17.14)	0.7(11.71)	0.6(12.86)	0.7(8.45)	0.1(23.35)	0.2(14.57)
i	Borneol	3.2 (43.16)	2.0(50.19)	0.5(21.20)	1.1(29.97)	4.6(4.07)	5.6(15.77)	8.0(3.26)	9.4(6.72)	1.5(8.72)	0.4(16.56)
17	*p*-Cymen-8-ol	–	0.1(65.86)	–	0.1(32.27)	–	< 0.1	–	–	–	< 0.1
j	α-Terpineol	t	0.1(58.05)	t	0.1(17.27)	0.1(28.90)	0.2(76.32)	31.0(9.41)	29.0(5.29)	–	0.1(117.35)
18	*cis*-Dihydrocarvone	–	0.1(78.32)	–	< 0.1	–	0.1(41.41)	–	0.3(10.61)	–	0.1(50.51)
k	Thymol-methyl ether*	0.1(39.53)	0.7(15.22)	0.1(9.00)	0.7(15.22)	t	–	0.2(24.30)	–	12.5(9.52)	14.0(9.92)
19	Isoterpinolene	–	0.1(42.74)	–	0.1(29.69)	–	0.2(25.00)	–	–	–	< 0.1
20	Carvacrol-methyl ether*	–	–	–	–	–	–	0.2(15.35)	0.1(12.11)	8.0(9.11)	7.6(13.06)
21	L-carvone	–	–	–	–	–	0.1(13.69)	–	–	–	–
l	Thymoquinone	8.2(43.35)	10.4(59.57)	3.0(13.81)	10.7(27.19)	11.3(5.91)	13.2(9.43)	0.1(18.86)	–	2.5(11.92)	12.0(19.01)
22	Bornyl acetate	< 0.1	–	t	–	0.1(14.48)	–	0.6(7.83)	0.7(7.52)	–	–
m	Thymol	15.8(22.42)	13.0(38.93)	6.4(7.71)	11.4(7.75)	29.9(11.43)	23.4(4.79)	–	–	3.2(7.09)	0.6(85.23)
23	Carvacrol	0.9(32.29)	0.9(47.45)	0.3(17.45)	0.8(15.69)	1.5(6.63)	1.7(6.52)	0.2(14.18)	–	3.6(11.87)	1.4(16.37)
24	Thymyl acetate	–	–	–	–	–	0.3(10.21)	–	–	–	–
25	Hydroxy-thymoquinone	–	0.2(81.71)	–	0.3(54.28)	–	–	–	–	–	–
26	α-Copaene	–	–	–	–	–	–	–	0.1(20.01)	–	0.1(18.96)

Table 2. Continued

Peak No.	Compound	*T. quinquecostotus*				*T. quinquecostotus var. japonica*		*T. mongolicus*		*T. serpyllum*	
		Jeju, Korea		Mt. Gaya, Korea		Ulreungdo, Korea		Northeastern Asia		Europe	
		13 D	16 M	13 D	16 M	13 D	16 M	13 D	16 M	13 D	16 M
27	β-Bourbonene	–	–	–	–	–	–	0.6(12.17)	0.4(13.81)	0.3(20.20)	0.6(17.43)
n	β-Caryophyllene	4.3(37.81)	3.5(59.49)	3.6(14.10)	5.8(20.06)	3.5(18.40)	3.3(10.46)	8.7(4.79)	9.6(8.29)	3.3(4.69)	2.0(14.43)
28	β-Cubebene	–	–	–	–	–	0.1(17.32)	–	0.1(21.71)	–	0.2(17.29)
29	Aromadendrene	–	0.5(65.47)	–	0.4(26.02)	–	0.6(10.28)	–	0.1(30.37)	–	0.1(20.24)
30	α-Humulene	–	0.2(86.44)	–	0.2(28.35)	–	0.1(11.10)	–	0.2(13.12)	–	< 0.1
31	Alloaromadendrene	0.3(54.19)	0.4(79.58)	0.1(28.43)	0.1(31.62)	0.2(18.55)	< 0.1	–	0.1(11.81)	–	< 0.1
32	α-Amorphene	–	0.2(59.77)	–	0.1(33.44)	–	< 0.1	–	0.1(32.88)	–	0.2(14.72)
33	Germacrene-D	–	0.1(25.69)	–	0.1(25.69)	–	–	0.2(25.84)	1.9(11.81)	–	0.8(22.08)
34	Ledene	0.8(54.35)	0.4(64.40)	–	0.2(30.13)	0.1(20.83)	0.2(10.15)	–	–	–	–
35	Bicyclogermacrene	–	–	–	0.1(14.33)	–	–	–	1.2(9.04)	–	0.2(21.30)
36	α-Farnesene	–	–	–	–	–	–	0.3(17.32)	0.2(10.24)	–	–
o	β-Bisabolene	0.5(38.47)	–	0.2(19.45)	0.2(20.11)	2.4(8.67)	2.5(8.48)	4.9(15.01)	2.4(5.91)	2.6(21.64)	4.5(16.36)
37	γ-Cadinene	0.2(62.95)	0.5(53.92)	–	0.1(23.48)	0.3(17.40)	–	0.7(13.25)	0.1(20.84)	–	0.1(15.00)
38	δ-Cadinene	1.1(47.41)	0.8(51.99)	0.1(33.35)	0.1(20.22)	–	< 0.1	0.2(27.85)	0.1(22.68)	0.2(16.89)	0.3(11.88)
39	Caryophyllene oxide	–	< 0.1	–	0.1(40.71)	–	< 0.1	–	0.1(25.90)	–	–

* Tentatively identified by comparision of MS data with those of library. -: not detected, t: < 0.05%.

% figures are their relative proportions as percent of total area , %(RSD), n=3.

The characteristic compounds are indicated in bold.

DB-624 column was used for analysis of air-dried materials for 13 days and DB-5 column was used for analysis of air-dried materials for 16 months.

Characteristic grouped compounds	*T. quinquecostotus** (%)		*T.quinquecostotus var. japonica* (%)		*T. mongolicus* (%)		*T. serphyllum* (%)	
	HS-SPME	GC/SAW	HS-SPME	GC/SAW	HS-SPME	GC/SAW	HS-SPME	GC/SAW
Monoterpene phenols (thymol, carvacrol)	11.7 (6.7, 16.7)	31.1 (29.0, 33.1)	31.4	33.5	0.2	–	6.8	34.1
Monoterpene phenols precursors (p-cymene, γ-terpinene)	64.3 (52.0, 76.6)	35.7 (34.7, 36.7)	35.9	24.8	4.5	2.8	51.3	20.0
Monoterpene hydrocarbons (α-thujene, α-pinene, camphene, β-pinene, β-myrcene, sabinene, α-terpinene, terpinolene, ocimene, limonene, α-phellandrene)	7.0 (6.0, 8.0)	7.4 (6.5, 8.3)	8.6	5.5	34.8	13.9	8.8	9.0
Oxygenated monoterpenes (1,8-cineol, cis-sabinene hydrate, camphor, α-terpinen-4-ol, borneol, α-terpineol, thymoquinone, bornyl acetate)	11.3 (4.6, 18.0)	21.2 (18.9, 23.5)	17.6	32.3	44.5	64.9	6.2	14.9
Monoterpene phenol derivatives (thymol methyl ether, carvacrol methyl ether)	0.1 (0.1, 0.1)	–	–	–	0.4	–	20.5	12.7
Sesquiterpenes (β-bourbonene, β-caryophyllene, alloaromadendrene, α-farnesene, β-bisabolene, germacrene-D, ledene, γ-cadinene, δ-cadinene)	5.6 (4.0, 7.2)	4.6 (2.5, 6.8)	6.5	3.9	15.6	18.4	6.4	9.3

* In the case of *T. quinquecostotus,* percent value is an average value of Jeju and Mt. Gaya, ( ): It means % value of Jeju and Mt. Gaya, -: not detected.

Table 3. Comparison of composition of aroma compounds for air-dried for 13 days of *thymus* species by HS-SPME-GC-MS and GC/SAW

Fig. 3(B), (B') shows PCA for HS-SPME-GC-MS responses of air-dried for 13 days and 16 months of *thymus* species. As a result, a good classification among species of completely different chemotypes was obtained. But, the classification of *T. quinquecostotus* which are same species of different geographical origin in same country (from Jeju and Mt. Gaya in South Korea), classification of original species and its variety for *T. quinquecostotus*, grouped as an air-drying term 13 days and 16 months for *T. quinquecostotus* appeared to be quite lower than those achieved by GC/SAW.

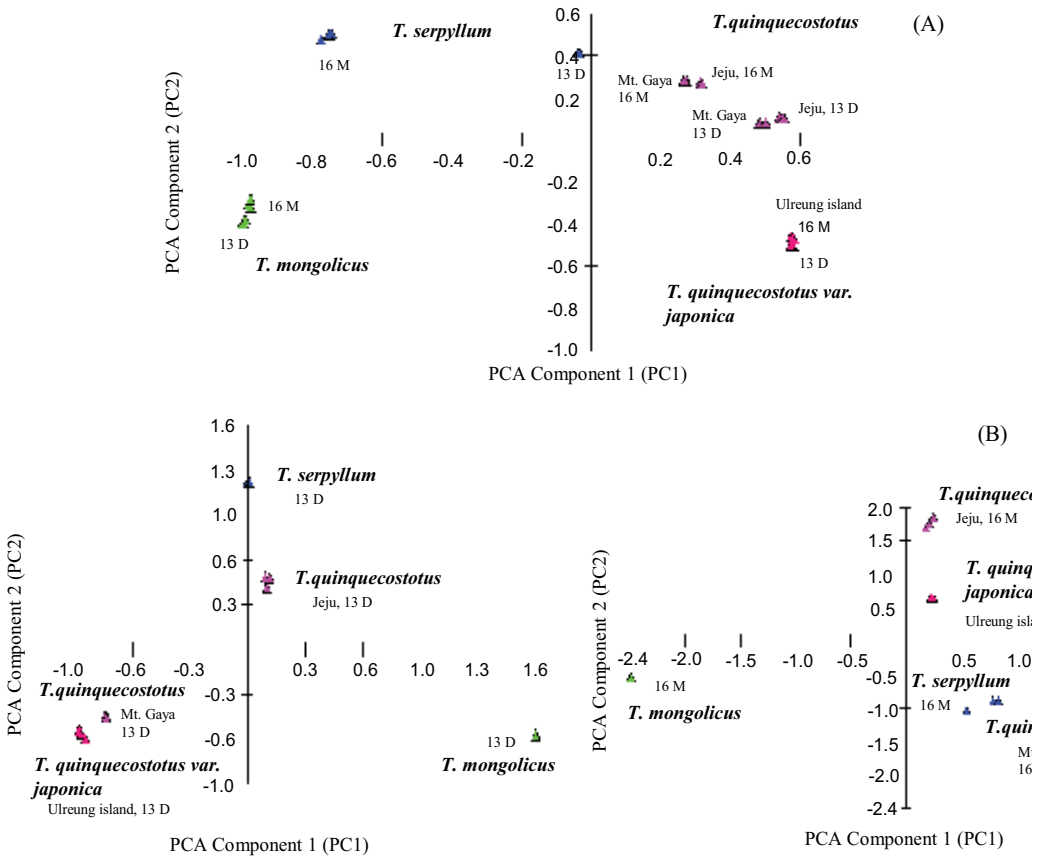


Fig. 3. Comparison of principal component analysis for GC/SAW (A) and HS-SPME-GC-MS (B), (B') responses of *thymus* species. (A) air-dried 13 days and 16 months by GC/SAW; (B) air-dried 13 days by Thermoquest-Finnigan ion trap GC–MS (Austin, Texas, USA) equipped with 6% cyanopropyl phenyl polydimethylsiloxane (DB-624, J&W, 30 m x 0.25 mm i.d., 1.4 μm film thickness); (B') air-dried 16 months by Hewlett-Packard 6890 Series GC system with an Agilent 5973N Mass Selective Detector (Agilent Technologies, Wilmington, DE, USA) equipped with 5% phenyl polydimethylsiloxane (Ultra 2 column, Agilent, 25 m x 0.25 mm i.d., 0.33 μm film thickness).

3.6 Changes of composition of aroma compounds according to elapse for T. quinquecostotus var. japonica by GC/SAW

Fig. 4 shows chromatograms of volatile aroma compounds according to elapse for *T. quinquecostotus var. japonica* by GC/SAW. The changes of composition of their volatile aroma compounds are summarized in Table 4.

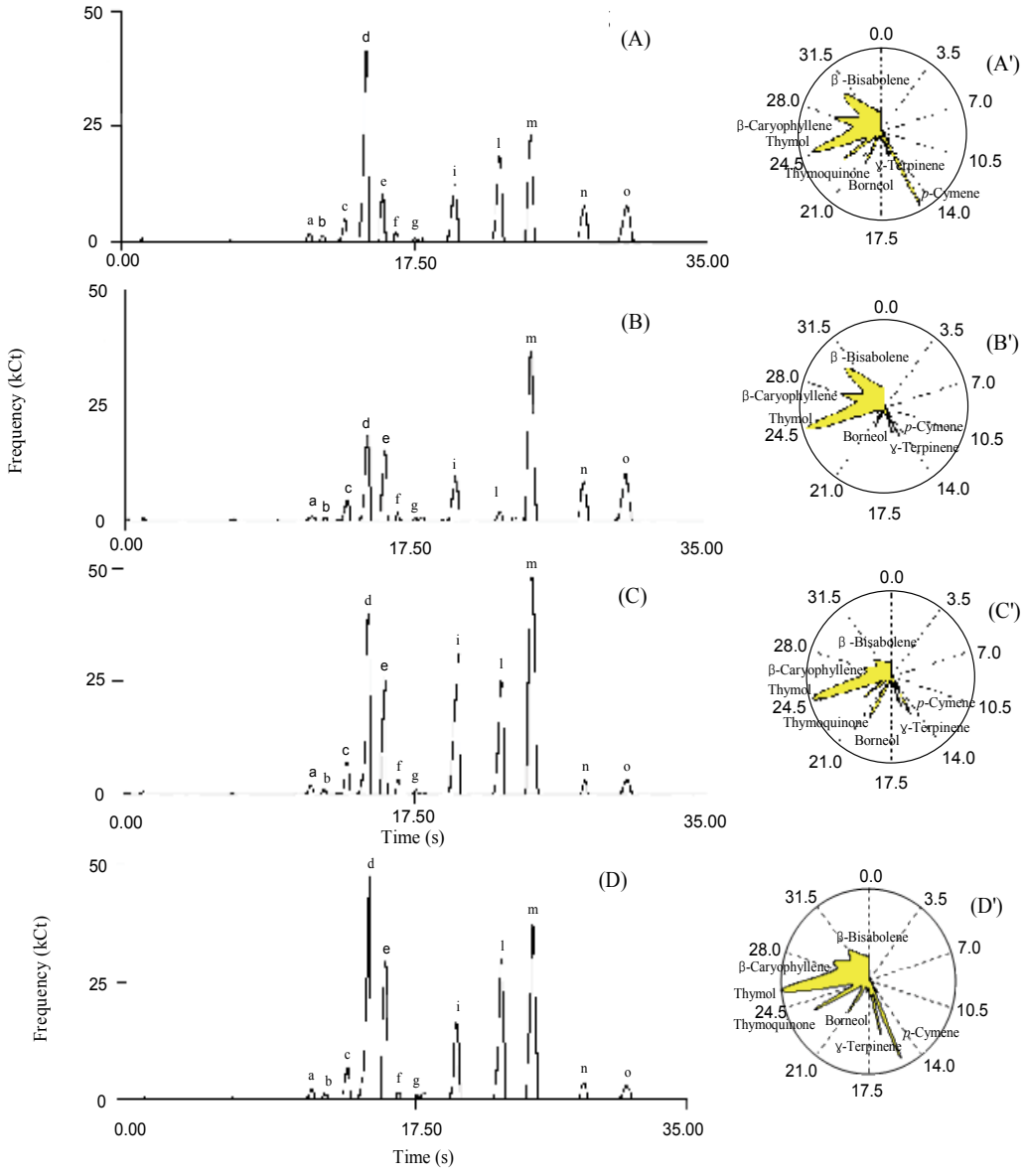


Fig. 4. Changes of composition of aroma compounds according to elapse for *T.quinquecostotus var. japonica* by GC/SAW. (A) fresh raw; (B) elapsed for 5 days at 5° C; (C) air-dried for 13 days at room temperature; (D) air-dried for 16 months at room temperature; (A′)-(D′) their corresponded fragrance patterns using VaporPrint.

Peak No.	t_R (s)	Components	*Thymus-quinquecostotus var. japonica* (Ulreung island)			
			Raw[a]	5 Days[b]	13 Days[c]	16 Months[d]
			%	%	%	%
a	11.140	α -Pinene	1.3(8.20)	0.7(4.21)	0.9(7.63)	1.4(1.49)
b	11.920	Camphene	0.8(2.53)	0.4(1.75)	0.5(7.66)	0.5(5.02)
c	13.280	β -Myrcene	3.5(5.90)	3.5(3.86)	3.1(6.05)	3.3(1.85)
d	14.500	*p*-Cymene	**25.5**(6.24)	**13.6**(4.83)	**15.8**(7.57)	**25.1**(1.58)
e	15.520	γ -Terpinene	**5.2**(4.98)	**10.2**(3.95)	**9.0**(8.65)	**12.4**(7.99)
f	16.300	Terpinolene	1.1(2.70)	0.7(2.64)	1.0(3.90)	0.4(4.30)
g	17.440	*cis*-Sabinene hydrate	0.3(8.68)	0.3(5.22)	0.4(4.23)	0.3(3.18)
h	19.020	Camphor	–	–	–	–
i	19.820	Borneol	**10.5**(2.80)	**9.1**(3.65)	**18.4**(1.24)	**9.6**(1.71)
j	20.560	α -Terpineol	–	–	–	–
k	21.520	Thymol methyl ether	–	–	–	–
l	22.480	Thymoquinone	**14.9**(3.35)	**1.6**(0.70)	**13.5**(5.81)	**18.4**(3.86)
m	24.340	Thymol	**20.4**(1.65)	**37.4**(5.37)	**33.5**(4.08)	**24.1**(0.73)
n	27.480	β-Caryophyllene	7.1(2.46)	8.8(8.92)	1.5(2.39)	1.9(2.20)
o	30.080	β -Bisabolene	9.4(5.03)	13.7(5.19)	2.4(2.40)	2.6(1.84)

Table 4. Change of composition of volatile aroma compounds according to elapse for *T.quinquecostotus var. japonica* by GC/SAW

[a] Fresh sample, it was analyzed as soon as possible after it was picked.

[b] It was elapsed for 5 days at 5°C.

[c] It was dried for 13 days at room temperature.

[d] It was dried for 16 months at room temperature

In the raw sample which are prepared as soon as possible after it was picked, monoterpene phenol (thymol: 20.4%) is lower than its corresponding precursor (*p*-cymene, γ-terpinene: 30.7%). Whereas, in the case of two samples which are elapsed for 5 days at 5 °C and air-dried for 13 days at room temperature, active thymol (33.5%~37.4%) is higher than its corresponding precursor (23.8%~24.8%). Additionally, thymoquinone (13.5%), borneol (18.4%) in 13 days sample increased than those of 5 days sample. After elasping for 16 months, the concentration of active thymol (24.1%) and borneol (9.6%) considerably decreased, whereas thymoquinone (18.4%) increased in high intensity.

As these results, they also show some interesting characteristic feature for the influence of air-drying on the volatile aroma compositions. In the aspect of pharmacological effects, it was found that active monoterpene phenol (thymol) reaches its highest concentrations after it was dried for 5 days or 13 days, which is much higher than in fresh or over-dried for a long times.

4. Conclusion

On the basis of this study, it is concluded that the GC/SAW analytical method shows high speed detection within ten seconds, a relatively high reproducibility compared with HS-SPME method, simplicity, and making it possible to detect sensitive aroma materials

quantitatively. It also enables to discriminate the botanical and geographical origin of *thymus* species by recognizable fragrance pattern analysis.
As a result, a new methodology by GC/SAW can serve as an alternative analytical technique for the analysis of discrimination of *thymus* species that provides second unit analysis, simple, highly sensitive analytical method, and fragrance pattern recognition compared to the conventional HS-SPME-GC-MS technique. In addition, it would be a first report to deal with the volatile herbal actual compositions and discrimination of *thymus* species by GC/SAW and will be applied to a variety of applications.

5. References

[1] G. Sachetti, S. Maietti, M. Muzzoli, M. Scaglianti, S. Manfredini, M. Radice, R. Bruni, Food Chem., 91 (2005) 621.
[2] J. Mastelić, I. Jerković, Food Chem. 80 (2003) 135.
[3] S.J. Lee, K. Umano, T. Shibamoto, K.G. Lee, Food Chem. 91 (2005) 131.
[4] P.H. Gouyon, Ph. Vernet, J.L. Guillerm, G. Valdeyron, Heredity 57 (1986) 59.
[5] M.D. Guillén, M.J. Manzanus, Food Chem. 63 (1998) 373.
[6] R. Venskutonis, L. Poll, M. Larsen, Flavour Fragrance J. 11 (1996) 123.
[7] F. Diemer, J.C. Caissard, S. Moja, J.C. Chalchat, F. Jullien, Plant Physiol. Biochem. 39 (2001) 603.
[8] M. Zimmermann, P. Schieberle, Eur. Food Res. Technol. 211 (2000) 175.
[9] G.B. Nickerson, S.T. Likens, J. Chromatogr. 21 (1966) 1.
[10] M.C. Díaz-Maroto, M.S. Pérez-Coello, M.D. Cabezudo, J. Chromatogr. A 947 (2002) 23.
[11] G. Suvarnalatha, M.S. Narayans, G.A. Ravishandar, L.V. Venkataraman, J. Sci. Food Agric. 66 (1994) 439.
[12] C. Bicchi, S. Drigo, P. Rubiolo, J. Chromatogr. A 892 (2000) 469.
[13] M. Kovaćević, M. Kać, J. Chromatogr. A 918 (2001) 159.
[14] E.J. Staples, Abstr. Pap. Am. Chem. 219 (2000) 236.
[15] E.J. Staples, J. Acoust. Soc. Am. 108 (2000) 2495.
[16] M. Kunert, A. Biedermann, T. Koch, W. Boland, J. Sep. Sci. 25 (2002) 677.
[17] J. Lammertyn, Els A. Veraverbeke, J. Irudayaraj, Sens. Actuators B 98 (2004) 54.
[18] S.Y. Oh, H.D. Shin, S.J. Kim, J. Hong, J. Chromatogr. A 1183 (2008) 170.
[19] M.K. Lee, Y.M. Ahn, K.R. Lee, J.H. Jung, J. O.S. Jung, J. Hong, Anal. Chim. Acta 633 (2009) 271.
[20] M.H. Chun, E.K. Kim, K.R. Lee, J.H. Jung, J. Hong, Microchem. J. 95 (2010) 31.
[21] D. Zabaras, S.G. Wyllie, Flavour Fragr. J. 16 (2001) 411.
[22] R. Granger and J. Passet, Phytochemistry 112 (1973) 1683.
[23] A.J. Poulose, R. Croteau, Arch. Biochem. Biophys. 187 (1978) 307.
[24] M. Hudaib, E. Speroni, A.M. Di Pietra, V. Cavrini, J. Pharm. Biomed. Anal. 29 (2002) 691.
[25] M.J. Jordán, R.M. Martínez, K.L. Goodner, E.A. Baldwin, J.A. Sotomayor, Ind. Crops Prod. 24 (2006) 253.

14

Application of Acoustic Waves to Investigate the Physical Properties of Liquids at High Pressure

Piotr Kiełczyński
Laboratory of Acoustoelectronics,
Institute of Fundamental Technological Research,
Polish Academy of Sciences,
ul. Pawińskiego 5B, 02-106 Warsaw,
Poland

1. Introduction

Monitoring and studying the pressure effect on liquid physical properties are becoming increasingly important in the food (Le Bail et al., 2003), chemical (Schettino et al., 2008), cosmetic (Sonneville-Aubrun et al., 2004) and pharmaceutical industry (Masson et al, 2001). High-pressure research of the physical properties of liquids has been stimulated by the fast development of such technologies as biodiesel production (Demirbas, 2008), high pressure food processing and conservation (Bamberger et al., 1999). High pressure processing enables inactivation of pathogenic microorganisms without decreasing the nutritional values and organoleptic properties. Rheological data provide information on molecular structure of the processed food. The knowledge of viscosity changes with pressure is also very important for food-processing plant design. Viscosity measurement on-line is necessary for control of food quality at different stages of the process.
Knowledge of the effect of pressure on the viscosity of polymer melts containing dissolved gases and on the viscosity of carbonaceous materials used in the impregnation process (Kosinskii, 2009) is also very important. Acoustic measurements offer also a potentially practical manner for the in situ characterization of reservoir fluids (e.g., crude oil) under reservoir conditions of pressure (Ball et al., 2002). The rheological properties of liquid lubricants (Bair et al., 2001) under high pressure determine friction and wear and they are fundamental properties for tribological evaluation of rolling bearings, gears and traction devices. High-pressure technologies (up to 1 GPa) have proved a great potential in modern bioengineering as a method of modification of biotechnological materials. The knowledge of physical properties (e.g., viscosity, compressibility) of treated substance is essential for understanding, design and control of the process technology. Measurement techniques for "in-situ" determining of physical parameters of liquids under high pressure allow insight into the phenomena governing the microstructural modifications occurring in the treated substance. High-pressure transitions in liquids can be investigated by the measurement of the viscosity or the acoustic wave phase velocity in function of hydrostatic pressure.

Pressure is an important parameter in processes of glass making when pressures in regime of some hundreds of megapascals occur e.g., in injection molding or pressure-assisted sintering. Changes in the melt viscosity upon pressurization result in changes in the glass transition temperature and the working regime for glasses (Del Gaudio et al., 2009).
In earth science interesting information can be obtained from the high-pressure rheological investigations of molten minerals (Bingwell et al., 2004). Moreover, oil-based drilling fluids have become widely used in oil industry because of their distinct advantage over water-based drilling fluids. However, unlike water-based drilling fluids, oil-based drilling fluids show significant rheological properties dependence on temperature and pressure. It is highly recommended to measure rheology of the actual mud system at high pressures prior to drilling operations (Zhao et al., 2008).
The properties of water at high pressures are investigated for the purpose of studying the physical chemistry, geophysics, planetology, and the most diverse problems in science and engineering. At pressures lower than 1 GPa, the rheological properties of water, including viscosity, are essential for the investigation of hydrodynamic and heat transfer processes. At present, the rheological properties of water are used in designing power plants of some types, as well as in developing new technologies in the food industry and medicine where, in particular, the shock compression of water has come to be used to lithotripsy (Mineev & Funtikov, 2005).
Direct measurements of the physical properties such as density, compressibility and isobaric heat capacity are very difficult under conditions of very high pressure. The speed of sound is closely linked with these thermodynamic properties and can be measured relatively easily and with high accuracy over wide ranges of temperature and pressure. An understanding of the pressure dependence of sound speed, attenuation, and relaxation frequencies can provide valuable information as to transport quantities such as fluid viscosity and thermal conductivity along with ratios of specific heats. Additionally, sound speed is closely related to derivatives of the equation of state. Therefore, the precision of these derivatives is often substantially better when they are deduce from the speed of sound rather than obtained from the analysis of classical pVT data.
Up to date, high pressure viscosity measurements were performed only in laboratory conditions, using conventional mechanical methods (Kulisiewicz & Delgado, 2010), developed as early as in the second half of the nineteenth century (Shames, 2002), (Ferguson & Kemblowski, 1991). The conventional mechanical methods followed works of such eminent scientists as Stokes, Navier, Poiseuille, Couette, et al. The common factor of all conventional mechanical methods is their inability to measure the viscosity on-line, without interfering with the industrial process controlled. Since on-line monitoring is necessary in process automation a need for new real-time monitoring methods emerged.
In this work new ultrasonic methods for the measurement of the viscosity of liquids under high pressure are presented (Kiełczyński et al., 2008a). These methods employ SH (shear horizontal) surface waves of the Love and Bleustein-Gulyaev (B-G) type. The energy of the SH surface wave is concentrated in the vicinity of the waveguide surface. Thus, the SH surface wave velocity and attenuation strongly depend on the boundary conditions on the waveguide surface which is viscoelastically loaded. Application of these SH surface waves extends considerably range of measuring pressures (up to 1 GPa). Moreover, the viscosity measurement is simplified and can be computerized. This enables on-line measurements of liquid viscosity.

2. Mechanical measuring methods for the measurement of liquid viscosity

Among the mechanical methods, the methods using rolling ball (King et al., 1992), falling ball (Nakamura et al., 2005), falling needle (Sha, 1997), and falling cylinder (Schaschke et. al., 2008) are the most popular. Rotational viscometers of Couette type (Matveev et al., 2005) form another group of high-pressure viscometers. The critical parts of rotating viscometers are seals. The third group of viscometers is based on the Hagen-Poiseuille formula for capillary flow (Ripple, 1992). Similarly, a modified capillary tube viscometer is a high-pressure extrusion slit die viscometer (Lan & Tseng, 2002). Another type of viscometers is a sliding plate viscometer. In these viscometers, the medium to be tested is charged in between two parallel sliding plates. After shear rate and shear stress are measured, the viscosity can be readily evaluated according to the Newton equation of viscosity (Koran & Dealy, 1999). However, it is very difficult to extend conventional methods to determine the viscosity at high pressure. One of the problems is to control the trajectory of the falling (rolling) ball and to track its movements. The resetting of the sinker or rolling ball also present difficulties. An eccentric fall of the sinker can cause significant errors in determining viscosity based on sinker descent time. Falling sinker viscometers and rolling ball viscometers have very long measuring times at high viscosities. Moreover, capillary type viscometers pose problems with pressure gradients.
Conventional mechanical methods and devices for measuring viscosity of liquids possess many disadvantages:

1. presence of moving parts
2. measurements are tedious and time consuming
3. require special sophisticated equipment
4. large dimensions
5. difficult to computerize

The application of rotary viscometers is limited due to the problems with generated heat and leakage during the transmission of the rotation into high-pressure chamber. Due to inherent limitations, the conventional methods cannot operate in real-time, and are only laboratory methods.
There exist also other methods employing different physical phenomena, e.g., magnetic field (Mattischek & Sobczak, 1997), (Royer et al., 2002) and light scattering (Fukui et al., 2010), for measuring the viscosity of liquids at high pressure. However, they need very complicated equipment and specially developed high-pressure chambers.
This is why, their use for measuring liquid viscosity at high pressure is very limited.

3. Ultrasonic methods

3.1 Bulk acoustic waves

Due to the disadvantages of the mechanical methods a need for new measuring methods arose. To this end, ultrasonic methods for the measurements of the viscosity of liquids under high pressure were proposed. Ultrasonic waves are mechanical disturbances, propagating in a material medium, at frequencies above 20 kHz. Present day technology enables for routine generation and detection of ultrasonic waves in the frequency range from ~20 kHz to ~2 GHz. However, the frequency range used in acoustic viscosity sensors is usually limited to 1-20 MHz. The ultrasonic methods due to their accuracy and relative simplicity can be applied in the study of liquid state. Ultrasonic velocity and attenuation measurements have proved to be useful in investigations of structures of liquids and interactions between the molecules.

Standing waves (resonators) and travelling waves (waveguides) were used to investigate the rheological properties of liquids at high pressure. For example, a torsionally oscillating piezoelectric quartz rod was applied as an ultrasonic viscosity sensor (Phillippoff, 1963), (Collings & McLaughlin, 1971), (Ruttle & Stephenson, 1975). In this type of ultrasonic sensors bulk type waves were applied. The acoustic energy of bulk waves is distributed in the entire volume of the resonator. The contact with a measured liquid takes place on the surface of the resonator. This results in the moderate sensitivity of this type of viscosity sensors.

A high-pressure (up to 300 MPa) torsional shear wave rheometer has been developed by Kulisiewicz (Kulisiewicz et al., 2007). This measurement system uses transmission of small amplitude torsional shear waves generated and detected by piezoelectric elements. In order to determine the complex shear modulus the measurement of the time of flight of the bulk torsional acoustic wave travelling between driver and sensor plates (distance 0.3–1 mm) is used to assess the wave velocity. To perform the viscosity measurement a very complicated calibration procedure is needed.

Modified crystal plate (langasite) resonators were also used to measure the viscosity of liquids at high pressure (Andle et al., 2008). This attempt was not successful because of the enormous troubles in the construction of the resonator. The structure of the resonator is fragile and not robust. Moreover, the range of measuring pressures was very modest (up to 60 MPa).

To overcome the disadvantages of the bulk wave methods, the author has proposed to use the SH surface acoustic waves of the Love and Bleustein-Gulyaev (B-G) type (Kiełczyński & Płowiec, 1989). At the beginning, the measurement of the liquid viscosity was carried out at the atmospheric pressure.

Subsequently, SH surface waves, i.e., Love waves and acousto-electric Bleustein-Gulyaev waves were used as a tool to measure the rheological parameters of liquids at high pressure (Kiełczyński et al., 2008a), (Kiełczyński et al., 2008b).

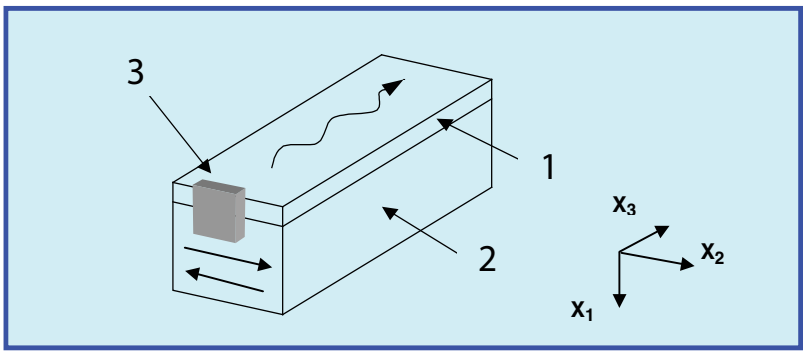
4. Surface acoustic waves

4.1 Love waves

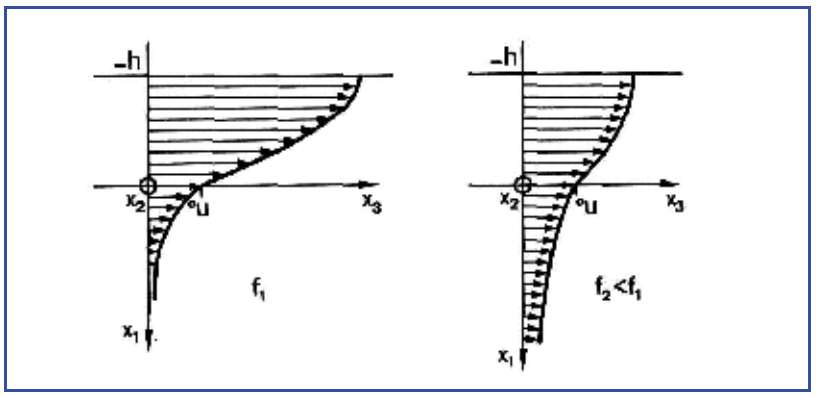
The Love wave propagates in a semi-infinite layered structure shown in Fig.1. Here, an elastic isotropic layer is rigidly attached to an isotropic and elastic half-space. Love waves can exist in special layered structures where phase velocity of the SH volume wave in the surface layer is smaller than that in the substrate, (Achenbach, 1973), (Farnell, 1978), (Royer&Dieulesaint, 2000). Mechanical vibrations of the shear horizontal surface wave are performed along the x_2 axis parallel to the propagation surface ($x_1 = 0$) and perpendicularly to the direction of propagation x_3. The energy of Love waves is concentrated in the vicinity of the surface. The amplitude $f(x_1)$ of the surface Love wave should vanish for $x_1 \to \infty$. The penetration depth of the Love wave is of the order of the wavelength. At low frequencies the energy of the Love wave propagates mainly in the substrate. As the frequency increases the fraction of energy travelling in the surface layer increases. This improves sensitivity to surface perturbations like liquid viscous loading.

The propagation of Love waves in the layered waveguides is governed by the differential problem (Sturm-Liouville problem). Solving this problem, we obtain a set of pairs $\left(\beta_i, f_i(x_1)\right)$, namely, the eigenvalue β_i, and eigenvector $f_i(x_1)$ correspond to the propagation constant and distribution of the mechanical displacement with depth x_1 of the Love wave. The index $i = 1$ refers to the fundamental mode. Higher modes of Love waves

are labeled by $i > 1$. Similar Sturm-Liouville problem describes propagation of light waves in planar optical waveguides and motion of quantum particles in a potential well (Schrödinger equation).



(a)



(b)

Fig. 1. a) Excitation of the Love wave in the layered waveguide by means of the PZT plate transducer (3). Cu surface layer (1) is deposited on a steel substrate (2), b) Love wave amplitude distribution with the depth x_1 for two different frequencies ($f_1 > f_2$).

The Love wave has a multimode character. In the present paper, we have restricted our attention to the propagation of the fundamental mode of Love waves.
Love waves are excited by the plate transducer (3) attached to the waveguide face, see Fig.1. The sending-receiving transducer (3) is excited to shear vibrations parallel to the waveguide surface and generates impulses of the Love wave that propagate along the waveguide surface. Theoretical and experimental analysis of the generation of SH surface waves by means of a plate transducer is presented in (Kinh & Pajewski, 1980).

4.2 Bleustein-Gulyaev (B-G) waves

Bleustein-Gulyaev (B-G) waves are shear horizontal acousto-electric waves, and they have no elastic counterpart (Royer & Dieulesaint, 2000), (Nakamura, 2007). If there is no piezoelectric effect, B-G wave degenerates to the shear bulk wave.

The distribution of the B-G wave mechanical displacement is similar to that of the Love wave. The B-G wave is capable of propagating along the surface of some crystals, e.g., with $6mm$ or $mm2$ symmetry (Zhang et al., 2001), as well as along the surface of properly polarized piezoelectric ceramics, see Fig.2.
Metallization of the PZT ceramic surface lowers the penetration depth of the B-G wave. In this case the penetration depth is of the order of a wavelength. Hence, in the metallized surface condition the B-G wave is more sensitive to liquid loading. B-G waves are excited similarly as Love waves, using the plate transducer (Kiełczyński et al., 2004), see Fig.2.

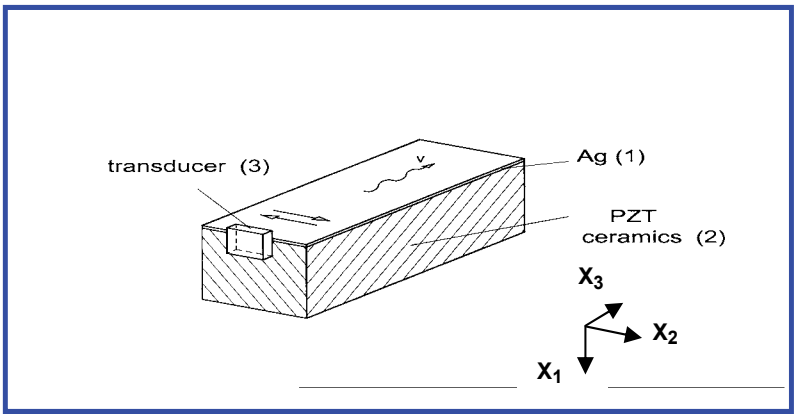


Fig. 2. Excitation of the B-G wave in a piezoceramic PZT waveguide (2) covered on the surface by a very thin metallic (Ag) layer (1) by means of the PZT plate transducer (3). PZT ceramics (both in the transducer and waveguide) is polarized along the axis x_2 .

The Love wave is a dispersive wave (i.e., the phase velocity is dependent on frequency) and can exhibit higher waveguide modes than fundamental one. By contrast, the B-G wave is a nondispersive wave. Moreover, an advantage of B-G wave for liquid sensing application is that B-G wave has no multiple modes. This makes that inverse determination of liquid properties by utilizing B-G wave is easier than that by utilizing SH surface waves of the Love type. Both types of SH surface waves are widely used in resonators, sensors and delay lines.

5. Application of SH surface waves for determining the rheological parameters of liquids at atmospheric pressure

To overcome the drawbacks of the bulk wave method, shear horizontal (SH) surface acoustic waves (SAW) such us:
1. Love waves and
2. Bleustein-Gulyaev (B-G) waves

have been introduced for the viscosity measurements under ambient pressures (Kiełczyński & Płowiec, 1989). These waves have only one SH component of mechanical displacement perpendicular to the direction of wave propagation and parallel to the waveguide surface. The energy of these waves is concentrated in the vicinity of the surface being in contact with a measured liquid. In consequence, the sensitivity of the viscosity sensors using SH surface acoustic waves (SAW) can be several orders larger than the sensitivity of the sensors employing bulk shear acoustic waves.

To measure the viscosity of liquid Rayleigh waves were also applied. Rayleigh waves have at least two components of vibrations i.e., longitudinal and vertical transverse, which cannot be separated. When Rayleigh waves propagate at a solid-liquid interface, the surface normal displacement radiates compressional waves into the liquid. Consequently, Rayleigh waves can be completely attenuated within the propagation range of the sensing device. Therefore, Rayleigh waves are impractical for use in the measurements of liquid viscosity. However, Rayleigh waves can be successfully applied in gas phase sensors.

In measurements of liquid viscosity, the effect of an investigated liquid on the properties of acoustic waves propagating in waveguides is primordial. The liquid presented on the waveguide surface loads it mechanically. The value of this load is proportional to the value of the mechanical impedance Z_L of a liquid medium (Kiełczyński et al., 2004). The mechanical impedance of a layer of liquid loading the surface of the SH surface wave (i.e., Love or B-G wave) waveguide is equal to the characteristic shear impedance of the liquid Z_L for plane waves:

$$Z_L = \left(\rho_L \cdot G_L\right)^{1/2} \tag{1}$$

where: $G_L = G' + jG''$ is the complex shear modulus of the liquid defined as the ratio (T/S) of the shear stress T to the shear strain S, ρ_L is the liquid density and $j = \left(-1\right)^{1/2}$.

In general, liquid loading of the sensor surface changes the phase velocity v and the attenuation α of the SH surface wave. The complex propagation constant γ of the SH surface wave changes (Ballantine et al., 1997):

$$\frac{\Delta\gamma}{\beta} = \frac{\Delta\alpha}{\beta} - j\frac{\Delta v}{v_0} \tag{2}$$

where: $\gamma = \alpha + j\beta$, $\beta = \omega/v$, v_0 is the phase velocity of the non-perturbed SH surface wave on the free surface, and ω is the angular frequency of the SH surface wave.

Significant experimental indications result from Eq.2. Namely, (1) by measuring the time delay between two subsequent echoes, one can determine the relative change in phase velocity of the surface wave $\frac{\Delta v}{v_0}$, and (2) by measuring the amplitudes of the subsequent impulses of the surface waves, we can determine the relative change in the surface wave attenuation $\frac{\Delta\alpha}{\beta}$. In this way, the relative change in the complex propagation constant $\frac{\Delta\gamma}{\beta}$ of the surface wave is determined experimentally. Knowledge of the change in complex propagation constant γ is fundamental to the established nondestructive method used to determine the rheological parameters of a liquid medium.

By applying the perturbation method one can prove that the change in the complex propagation constant γ of the SH surface wave produced by viscoelastic liquid loading is as follows (Auld, 1973):

$$\Delta\gamma = -j\left(\frac{\left|v_2\right|^2_{x_1=0}}{4P}\right)Z_L = -jKZ \tag{3}$$

where: v_2 is the SH surface wave amplitude on the waveguide surface ($x_1 = 0$), P is the mean power on the unit width of the SH surface wave. The coefficient K is the characteristic quantity for each SH surface wave waveguide and depends solely on the material parameters of the waveguide and frequency (Kiełczyński & Płowiec, 1989).
Knowing the change in the complex propagation constant $\Delta\gamma$ from the experiment, we can calculate the complex shear impedance of a liquid $Z_L = R_L + jX_L$. Subsequently, by separating the real and imaginary parts of the Eq.1 we can calculate the real G' and imaginary G'' parts of the complex shear modulus G_L of the liquid and, consequently, the rheological parameters of a viscoelastic liquid.

6. Application of SH surface waves for measuring the viscosity of liquids at high pressure

The Love wave and the Bleustein-Gulyaev (B-G) wave method for measuring the viscosity of liquids at high pressures have been established in the Laboratory of Acoustoelectronics of the Institute of Fundamental Technological Research, Polish Academy of Sciences in Warsaw, Poland (Kiełczyński et al., 2008a), (Kiełczyński et al., 2008b).
The SH SAW method for measuring the viscosity of liquids at high pressures possesses many advantages:

1. absence of moving parts
2. operation in real time
3. short measuring time
4. high sensitivity
5. low power consumption
6. small dimensions, simple and robust construction of the sensor
7. possibility of computerization
8. output signal is electrical
9. no leakage problems
10. no heating caused by shear

6.1 Measuring set up

High-pressure chamber was designed and fabricated in the Institute of Physics at Warsaw University of Technology (Rostocki et al., 2007). High pressure was generated in a thick-walled cylinder of 17 mm internal diameter with a simple piston and Bridgman II sealing system. The piston-cylinder assembly was working with a 20–tonne hydraulic press, driven by hand operated pump. The maximum pressure in this arrangement is limited to about 1.2 GPa due to the hydraulic press working range. For pressure measurement, a typical 500 Ω manganin transducer was used. Its resistance was measured with a precise HP 34970 multimeter. An accuracy of the pressure measurement was better than ± 0.5 MPa. All experiments were carried out at the temperature 293 K. Temperature was measured with the Cu – Constantan thermocouple placed inside the chamber. The described previously viscosity sensor (B-G or Love waveguide, see Figs.1, 2 and 4) was placed inside the high-pressure chamber, see Fig.3.
The piezoelectric transducer attached to the SH surface wave waveguide, manganin coil and thermocouple were connected with the external measuring setup by an electrical multichannel lead-through.

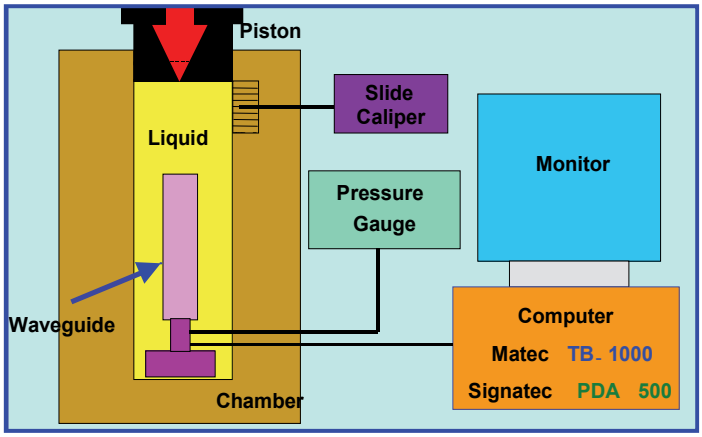


Fig. 3. Ultrasonic set up for measuring the viscosity and pressure of liquids under high pressure.

Fig. 4. Love wave waveguide (Cu surface layer on a steel substrate) connected to the high-pressure lead-through (on the left).

In the setup for measuring viscosity using the SH surface wave, see Fig.3, the sending-receiving piezoelectric transducer is driven by the TB-1000 pulser-receiver computer card (Matec, USA). The TB-1000 pulser generates the rf tone burst with a frequency f = 2 MHz and length equal to 0.5 μs. The repetition period equals 0.4 ms. The SH surface wave impulse generated by the transducer is reflected in multiple ways between two opposite edges of the SH surface wave waveguide (Fig. 4). The signals received by the transducer, see Figs.5a, b, are amplified by the TB-1000 receiver and sent into the PDA-500 digitizer card (Signatec, USA). This card samples and digitizes the input analog signals. The stored signals are then analyzed by computer software. For each measurement, the ultrasonic signal is averaged 1024 times in order to improve the signal – to – noise ratio. A computer program which controls the operation of the pulser–receiver card and digitizer card was written in C language.

6.2 Theoretical background

In this paper, the liquids investigated under high pressure are treated as the Newtonian liquids. The model of a Newtonian liquid was used by (Philippoff, 1963). He stated that the majority of oils in the considered shearing rate (about 1 MHz), and under high pressure are

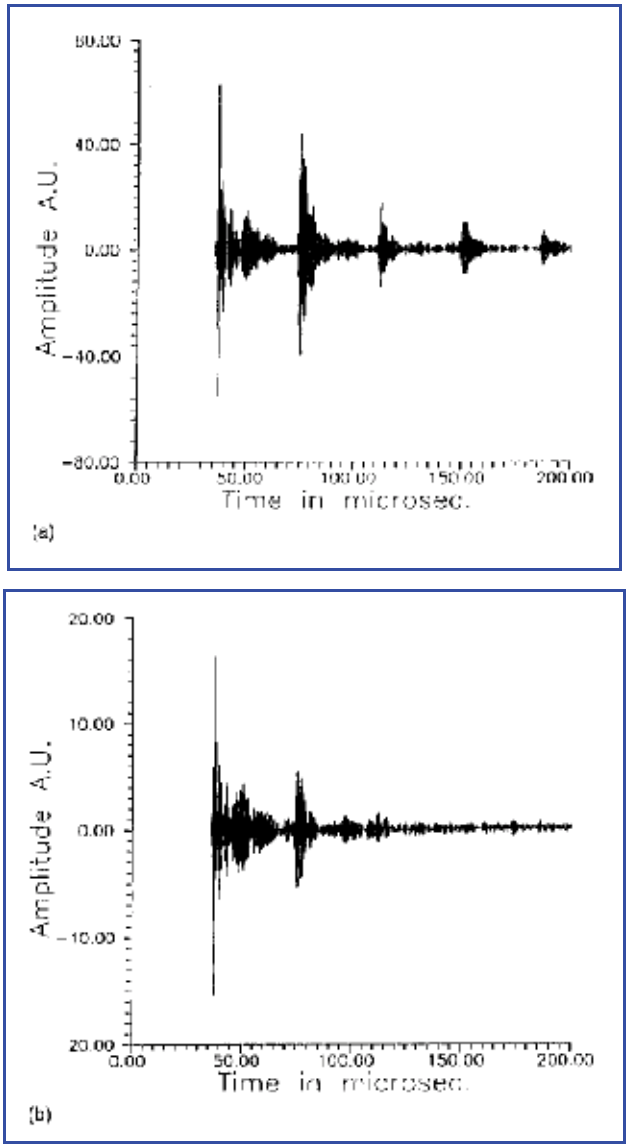


Fig. 5. (a) Oscillogram of the SH surface wave impulses reverberating in the waveguide unloaded with an investigated liquid, and b) Oscillogram of the SH surface wave impulses reverberating in the waveguide loaded with an investigated liquid.

the Newtonian liquids. This can justify the use of a Newtonian liquid model in our paper. For the case of a Newtonian (viscous) liquid, the shear mechanical impedance Z_L (defined as a ratio of the shear stress to the shear vibrational velocity) can be expressed as follows (Landau&Lifshitz, 1958):

$$Z_L = R_L + jX_L = \left(\frac{\rho_L \omega \eta}{2}\right)^{1/2} (1 + j) \tag{4}$$

where: η is the viscosity, ρ_L is the density of a liquid and $j=(-1)^{1/2}$.
So that, we may regard formula (5) as holding for the liquids considered in the paper.

$$\eta=\frac{2R_L^2}{\omega\rho_L}=\frac{2X_L^2}{\omega\rho_L} \tag{5}$$

where: R_L and X_L is a real and imaginary part of the mechanical shear impedance of a liquid. The shear mechanical impedance of a liquid $Z_L=R_L+jX_L$ can be determined from the measurement of the change in attenuation and time of flight of wave-trains that propagate in the waveguide loaded by a liquid (Kiełczyński et al., 2004), see Fig.6.
The real part R_L of the shear mechanical impedance of a liquid can be expressed as, see Fig.6:

$$R_L=\frac{\ln\left(A_1^0/A_1^1\right)}{2KL} \tag{6}$$

where: A_1^0 and A_1^1 represent amplitudes of the first echo of the SH surface wave for an unloaded $\left(A_1^0\right)$ and loaded $\left(A_1^1\right)$ waveguide respectively, L is the length of the waveguide covered with an investigated liquid.

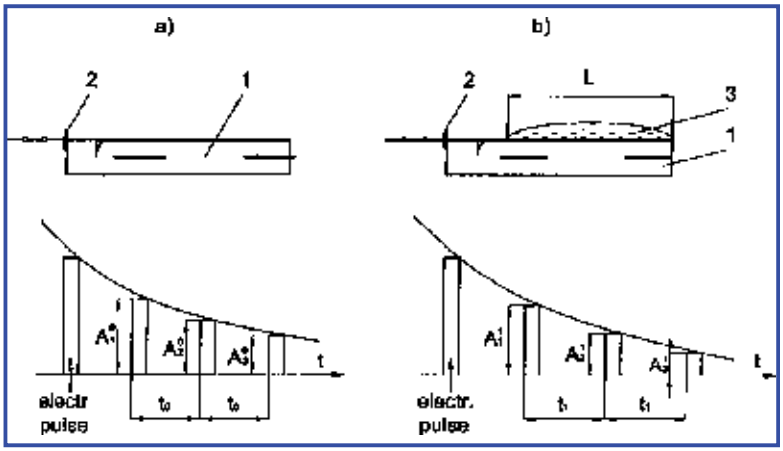


Fig. 6. Scheme of the SH surface wave measuring method, (a) free (nonloaded) waveguide surface and (b) waveguide surface loaded with a viscoelastic liquid. 1) waveguide of the SH surface wave, (2) sending+receiving transducer, and (3) layer of an investigated viscoelastic liquid.

6.3 Experimental results (Love waves)

An example of variations in viscosity of liquids as a function of hydrostatic pressure measured by the Love wave method is presented in Fig.7 (Kiełczyński et al., 2008b), (Rostocki et al., 2010).
Castor oil is a vegetable oil, that is a triglyceride in which approximately ninety percent of fatty chains are ricinoleic acid. Oleic and linoleic acids are the other significant components. Castor oil and its derivatives have applications in the manufacturing of soaps, lubricants, hydraulic and brake fluids, paints, dyes, coatings, inks, cold resistant plastics, waxes and polishes, nylon, pharmaceuticals and perfumes.

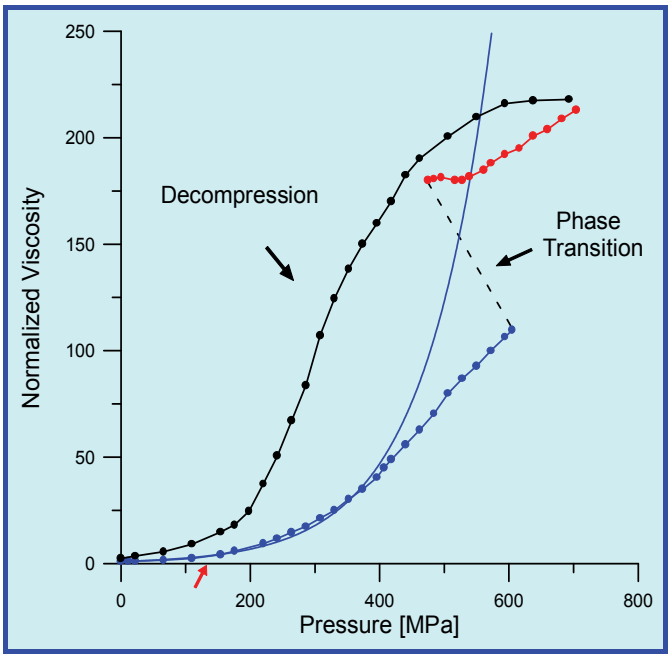


Fig. 7. Variations in viscosity of castor oil, as a function of hydrostatic pressure, measured by the Love wave method, $f = 2\,MHz$. Red arrow indicates the hydrostatic pressure such as on the bottom of the Marianas Trench.

The pressure was generated in 10 MPa steps then kept constant for about 2-5 minutes. During that time the pressure was carefully observed. That allowed to identify pressure drop due to the first order phase transition and to observe whether the system is reaching thermodynamic equilibrium. After approaching 0.6 GPa the pressure was kept constant for about 20 hours to enable the phase transformation to occur. During the phase transition the small drop of pressure and increment of viscosity was observed.

As it can be seeing in Fig.7, the experimental curve up to about 400 MPa is almost tangential to the exponential curve which represents the Barus formula $\eta(p) = \eta_0 \exp(\alpha p)$, (continuous curve in Fig.7), where: η_0 is the viscosity at atmospheric pressure and α is the viscosity – pressure coefficient. Above 400 MPa the experimental points are raising slower than the theoretical prediction. Finally, at 600 MPa when the pressure rise was stopped for about 20 hours the viscosity has risen to the new value characteristic for the high-pressure phase of castor oil. The further increment of viscosity was rather linear function of pressure.

6.4 Experimental results (Bleustein-Gulyaev waves)

Similar as in the case of Love waves, measurements of high-pressure liquid viscosity were also performed using the Bleustein-Gulyaev wave method. Fabrication of the B-G wave waveguide is easy and its construction is simpler that that of the Love wave. On the other hand, Love wave waveguides are more robust and mechanically resistant.

A triglyceride and unsaturated fat: a triolein ($C_{17}\,H_{33}\,COO)C_3\,H_5$ was investigated. Triolein is a model liquid in investigations of high-pressure phenomena in the natural oils that are very important in biodiesel technologies as well as in high-pressure food processing.

Most of the natural oils like castor oil, soybean oil, rapeseed oil, etc. consist of triglycerides of various fatty acids. The statistical characteristics of the molecular composition of these oils have made difficulties for the interpretation of the phenomena observed at high pressure. Therefore experiments have been concentrated upon well-defined triacylglycerol (triglyceride) structures. The triolein is very good model-liquid for the whole group of triglycerides since the phase transition takes place within only several dozen minutes after the application of sufficient pressure. That allows for more detailed and accurate studies than similar phase transition in vegetable oils taking even up to ten days from the application of sufficient pressure to the start of the phase transition.

Variations in viscosity of triolein as a function of hydrostatic pressure measured by the B-G wave method is presented in Fig.8 (Kiełczyński et al., 2008a). Up to about 500 MPa the viscosity was increasing exponentially according to the known empirical Barus formula (continuous curve in Fig.8). After approaching 700 MPa the compression was stopped and the piston in the high-pressure chamber was fixed to enable the phase transformation to occur undisturbed. During the phase transition a pressure drop of about 100 MPa was observed in the chamber. The viscosity showed the further rise despite the pressure drop. It means that volume occupied by the resulting high-pressure phase diminishes. After the termination of the phase transformation process, the further increase of viscosity with increasing pressure was observed. The changes of viscosity during the decompression process inducing the high-pressure phase decomposition have shown large hysteresis (upper curve). Large hysteresis indicates existence of large internal friction forces. The phase transition in the case of triolein (Fig.8) lasts 1 hour. By contrast the phase transition in castor oil (Fig.7) is completed after 20 hours.

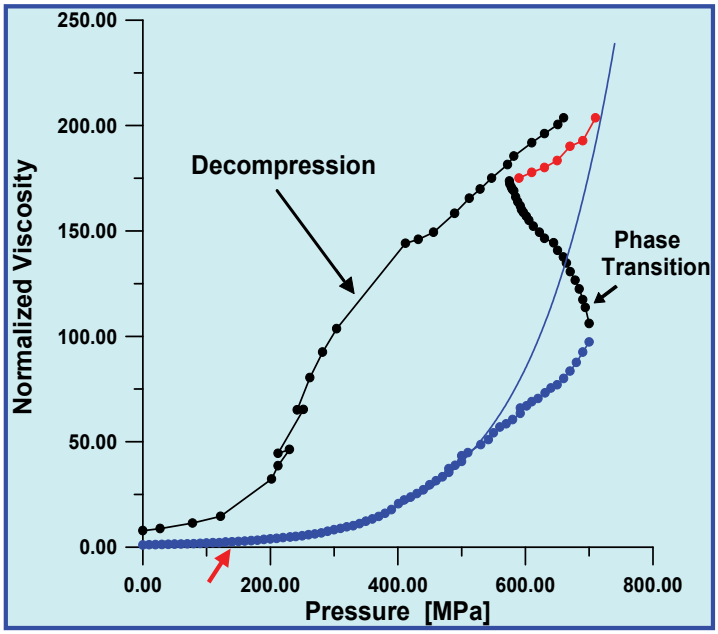


Fig. 8. Variations in viscosity of triolein, as a function of hydrostatic pressure, measured by the Bleustein-Gulyaev (B-G) wave method, $f = 2\,MHz$. Red arrow indicates the hydrostatic pressure such as on the bottom of the Marianas Trench.

7. Measurement of sound speed in liquids at high pressure

For the measurements of the phase velocity of longitudinal ultrasonic waves we have constructed the setup (Fig.9) especially designed to obtain a low level of parasitic ultrasonic signals (Kiełczyński et al., 2009).
High-pressure chamber and computer cards are the same as in the viscosity measurements described previously. A special mounting of transducers in the high-pressure chamber was fabricated. The transducers were 5 MHz $LiNbO_3$ (Y_{36} cut) plates (Boston Piezo-Optics Inc., USA). The phase velocity of the longitudinal ultrasonic wave was measured using a cross-correlation method (Sugasawa, 2002) to evaluate the time of flight (TOF).

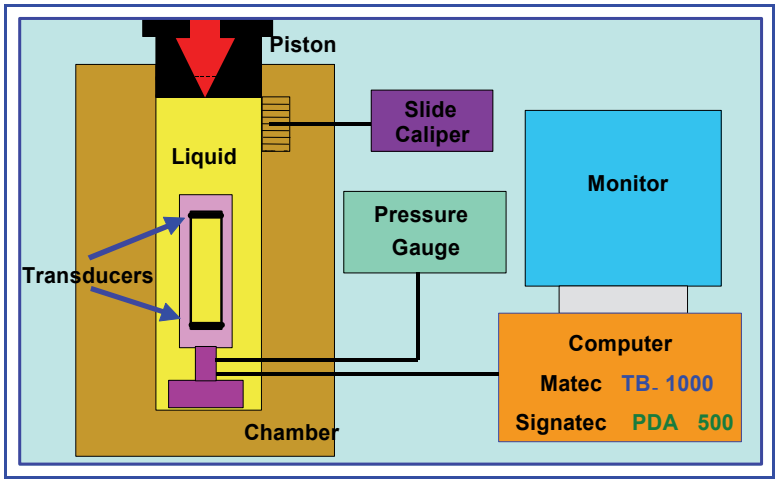


Fig. 9. Ultrasonic setup for measuring the phase velocity of longitudinal acoustic waves in liquids as a function of hydrostatic pressure.

For the sending and receiving of wave pulses the TB-1000 pulser-receiver computer card (Matec, USA) was used. The TB-1000 pulser generated the rf tone burst with a frequency 5 MHz and length equal to $0.3\,\mu s$. The longitudinal wave impulse generated by the sending transducer propagated in investigated liquid and was detected by the receiving transducer. The PDA-500 digitizer card (Signatec, USA) sampled and digitized the signals received by the transducer and amplified by the TB-1000 receiver. The stored signals were then analyzed by computer software. For each measurement, the ultrasonic signal was averaged 1024 times in order to improve the signal-to-noise ratio. A computer program that controls the operation of the pulser-receiver card and digitizer card was written in C++ language. The time of flight of the ultrasonic pulses (see Fig.10) was evaluated by applying the cross-correlation method (Sugasawa, 2002), (Viola & Walker, 2003). The cross-correlation method is a global differential method. Due to this reason, the cross-correlation method does not depend on the trigger level and delays in cables and amplifiers. The change in the height of the column of a liquid caused by the piston movement was measured by a digital caliper. The piezoelectric transducers and manganin coil were connected with the external measuring setup by an electrical multichannel lead-through.

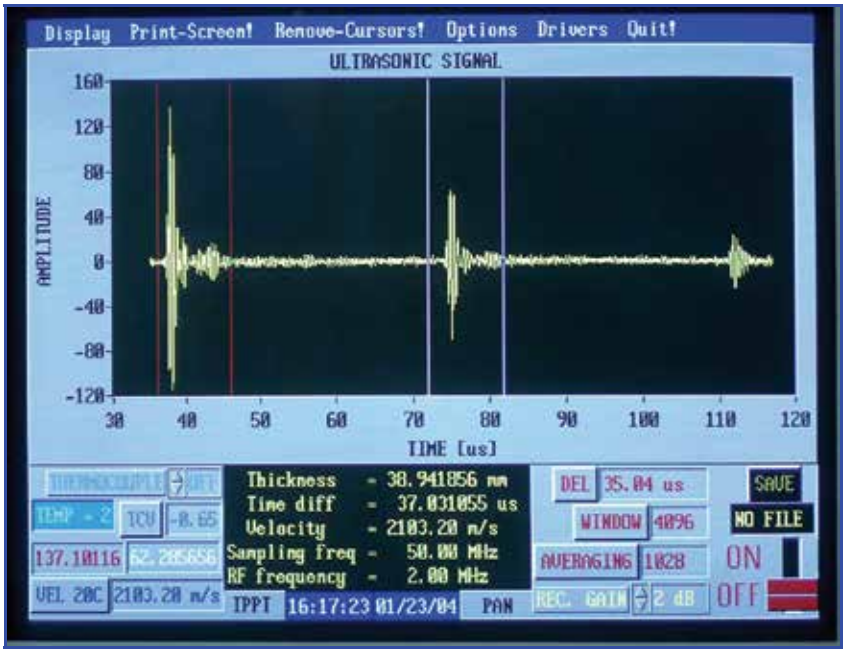


Fig. 10. Time of flight (TOF) between two ultrasonic impulses (delimited by cursors) is evaluated by using the cross-correlation method.

The sound velocity v_L was calculated using the formula:

$$v_L = \frac{l_0}{\Delta t} \tag{7}$$

where: l_0 is the distance between sending and receiving transducer, Δt is time of flight (TOF) of the ultrasonic signal.

The isothermal compressibility β_T is given by the formula:

$$\beta_T = -\frac{1}{V}\frac{\partial V}{\partial p} \tag{8}$$

where: V is the volume of a liquid in the chamber for a given value of the hydrostatic pressure p.

7.1 Results

The measurements of the phase velocity (Fig.11) and isothermal compressibility (Fig.12) of triolein were carried out in function of hydrostatic pressure up to 650 MPa.

The pressure was generated in 20 MPa steps then kept constant about 2 min. that allowed to control whether the system was reaching equilibrium. Up to 450 MPa the phase velocity was increasing monotonically with pressure (arrow 1 in Fig.11). After approaching 450 MPa the compression was stopped, and the piston in the high-pressure chamber was fixed to enable the phase transformation to occur undisturbed. During the phase transition a pressure drop of about 150 MPa was observed in the chamber. It means that the volume occupied by the resulting high-pressure phase of triolein diminished. The phase velocity showed the further rise despite the pressure drop (arrow 2 in Fig.11). Finally the phase velocity has risen to the

new value characteristic for the high-pressure phase of triolein. Once the phase transition was completed the pressure was further increased up to about 650 MPa (arrow 3 in Fig.11). The phase velocity of longitudinal waves in high-pressure phase has increased monotonically. After approaching 650 MPa the decompression process was started (arrow 4 in Fig.11). At the point marked in Fig.11 by **a** the decomposition of the high-pressure phase started. Between points marked by **a** and **b** two phases coexisted in triolein.

Compressibility is an important property. It enters into many pressure-dependent thermodynamic expressions, and is an essential parameter for the design and use of any high-pressure equipment. Compressibility is dependent on the intermolecular forces acting within the substance, that is, it is the result of the balance between attractive and repulsive potentials. Compression results in decreasing the average intermolecular distance and reducing rotational and translational motion. Compressibility of liquids decreases with pressure, since the initial "free volume" has largely disappeared, and the repulsive potential is stronger than the attractive at high pressure. (Barbosa, 2003)

The isothermal compressibility of triolein presented in Fig. 12 was calculated using Eq.8. The volume changes ΔV were determined from the changes of the height of the triolein column measured by the slide caliper. The arrows indicated by numbers 1, 2 and 3 in Fig.12 refer (similarly as in Fig.11) to the low-pressure phase, phase transition and high-pressure phase respectively.

It is worth noticing that the value of isothermal compressibility β_T during the phase transition is negative. Moreover, the isothermal compressibility of high-pressure phase is different than that of low-pressure phase.

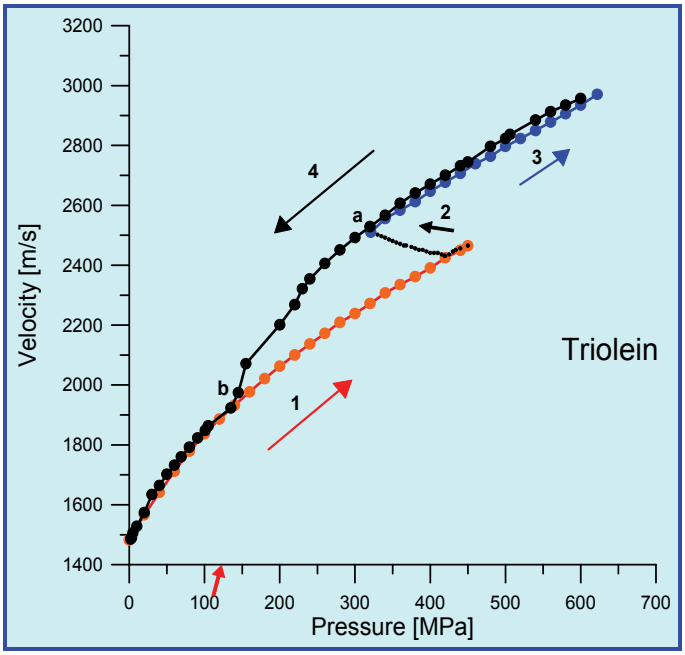


Fig. 11. Phase velocity v_L of longitudinal acoustic waves in triolein in function of hydrostatic pressure. (1) refers to low-pressure phase, (2) indicates the phase transition, (3) refers to high-pressure phase, and (4) indicates the decompression, $f = 5\,MHz$. Red arrow indicates the hydrostatic pressure such as on the bottom of the Marianas Trench.

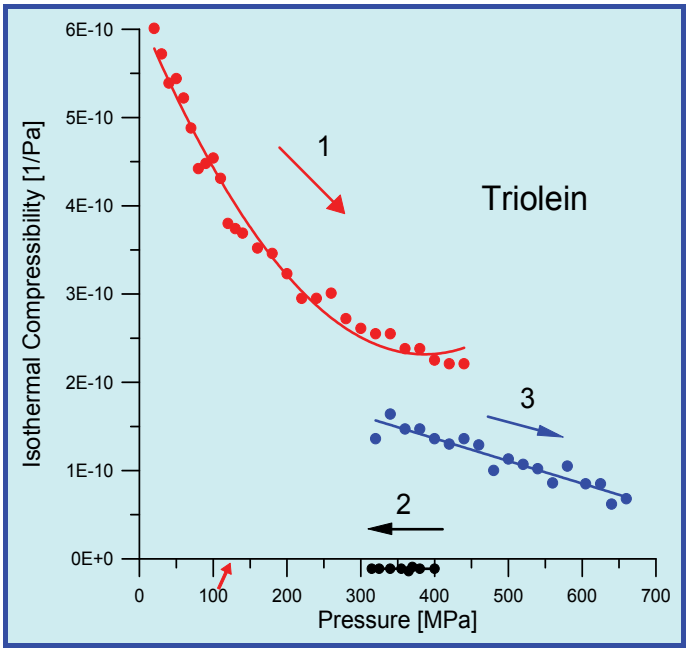


Fig. 12. Isothermal compressibility of triolein β_T as a function of hydrostatic pressure. During the phase transformation (arrow 2) β_T is negative. $f = 5\,MHz$. Red arrow indicates the hydrostatic pressure such as on the bottom of the Marianas Trench.

7.2 Possibility of measurement of various physical (thermodynamic) parameters

The speed of sound is a particularly interesting property to study as it provides an indirect way to all of the observable thermodynamic properties of a single fluid phase. This way is a convenient one experimentally because sound speed measurements may be made quickly and accurately over ranges of temperature and pressure by means of largely automated apparatus.

The knowledge of the thermodynamic properties of pure organic liquids is of practical interest to industries in different fields, such as chemical, pharmaceutical industries, and food technology, because the applied industrial procedures are influenced by the temperature and pressure dependence of the used liquids.

The seven thermodynamic variables: pressure, volume, temperature, entropy, and the three components of the vector fluid velocity, can be related to one another through the equation of state, the equation of energy conservation, and also through the equation of continuity for mass and momentum, and the second law of thermodynamics. (Heydemann&Houck, 1969), (Stallard et al., 1969).

The densities, isobaric heat capacities, isobaric thermal expansions, isentropic compressibilities, isothermal compressibilities, and internal pressures as functions of temperature and pressure can be calculated using the experimental speeds of sound under elevated pressures together with the densities and heat capacities at atmospheric pressure (Oakley et al, 2003a), (Oakley et al, 2003b).

The bulk modulus of biodiesels determines the spray characterisics upon injection. As the fuel injection in the engine is approximately an adiabatic process, the adiabatic bulk modulus seems to be more useful than the isothermal one in estimation of the fuel injection

timing. The only experimental method that leads directly to adiabatic modulus is the acoustic one, based on the measurement of the speed of sound. The method is relatively simple tool for determination of thermodynamic properties, especially at high pressures (Dzida & Prusakiewicz, 2008).
The measurements of ultrasonic velocity can be applied for studying the nature of molecular systems and physicochemical properties of liquid mixtures and solutions. The results are interpreted in terms of molecular interaction between components of the liquid mixtures. A study of the thermophysical properties as a function of pressure and temperature in a homologous series of chemical compounds is of great interest not only for industrial applications (for example, the petroleum industry), but also for fundamental aspects for understanding the influence of the chain length of the components on the liquid structure and the developing models for an accurate representation of the liquid state. To this aim, ultrasonic speed measurements under pressure have been successfully applied (Daridon et al., 2002).
The experimental determination of the non-linearity parameter B/A is possible using the measurement of the ultrasonic wave velocity change due to an isentropic change of the static pressure (Khelladi et al., 2009), (Plantier et al., 2002).

8. Measurement of the physical properties of liquids during phase transition

High hydrostatic pressure can change molecular structure and intermolecular interactions in liquids. (Kulisiewicz et al., 2007), (Nithya et al., 2009). Rheological parameters describe macroscopic properties of a material. However, they are governed by molecular structure and mutual interactions of molecules in the material. Therefore, rheological parameters can be correlated with micro-structural parameters of a liquid (Delgado et al., 2010).
Investigation of phase transitions is important in lubricants, since rheological properties of lubricants can change during phase transitions. Investigations of the phase transitions in vegetable oils (e.g., castor oil or olive oil) and in the triglycerides (e.g., triolein) are of great importance. Vegetable oils are usually excellent boundary lubricants. They show higher viscosity index than mineral oils and they are environmentally friendly. In general, vegetable oils are highly attractive substitutes for petroleum based oils. Unfortunately their high-pressure behavior was not yet systematically investigated. The most important is to determine the range of pressures when phase transition (solidification) begins (Mia et al., 2007).
Investigation of phase transitions is also very important in food industry and in food conservation. Phase transitions can modify irreversibly the molecular structure and quality of food products. Media with high molar volumes like edible oils and fats exhibit phase transition at pressure levels about several hundred megapascals.
Investigation of phase transitions was impossible with conventional mechanical methods. By contrast, the proposed novel SH surface wave methods enable for the measurement of the rheological parameters of liquids during phase transitions.
As it is seen in Figs.7, 8, 11 and 12 during phase transitions a step change in liquid viscosity, phase velocity and compressibility occurs. This phenomenon is a clear indication that phase transitions in a liquid were initiated.
The kinetics of the phase transition, as a function of pressure, was investigated during phase velocity measurements in triolein. Pressure changes, occurring during phase transition, were registered with the piston locked in a fixed position, see Fig. 13. At stable pressure

conditions (450 MPa) the pressure remained constant in the first 30 minutes. Subsequently, a rapid decrease of pressure was observed, due to a phase transition in triolein. After about 60 minutes the pressure level stabilized on 330 MPa. This means that phase transition was complete. As a result, a new high-pressure phase in triolein with different microstructure has emerged. Physical properties of the new high-pressure phase are different than those in the low-pressure phase. The high-pressure induced phase transitions in vegetable oils are the phase transitions of first order.

To the author's knowledge, the measurement of the phase velocity of longitudinal acoustic waves in liquids during the phase transitions was not reported in the scientific literature and is a novelty. Measurement of the viscosity and sound speed in liquids during phase transitions is an original author's contribution.

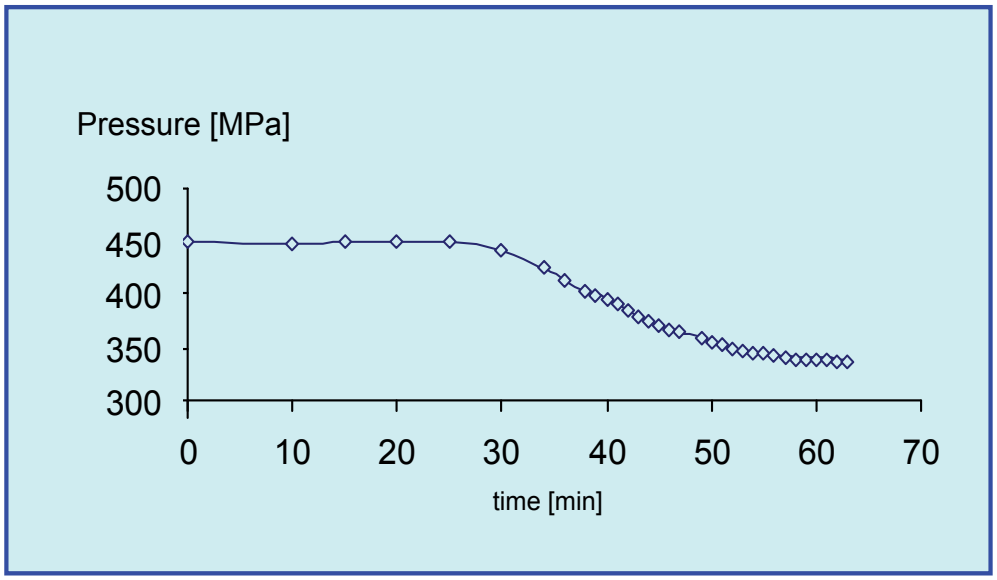


Fig. 13. Variation in pressure on time during the phase transition in triolein.

9. Summary

In this Chapter new methods for measuring the viscosity of liquids at high pressure are presented. Measurement of liquid viscosity at high-pressure is important in tribology in rolling bearings, in design and exploitation of ship diesel engines, in the chemical, pharmaceutical and cosmetic industries as well as in bio-fuels, and food conservation.

Based on the SH surface Love and Bleustein-Gulyaev waves, we designed novel methods and devices to characterize viscosity of liquids at high pressure. The SH SAW viscosity sensor is electrically responsive. Owing to this fact, modern methods of the digital signal acquisition and processing can be efficiently used. The measuring setup operates in real-time and can be employed for measuring liquid viscosity under high-pressure in the course of the technological processes. In general, the SH SAW method has high sensitivity and high reliability. The sensitivity of this method can be several orders larger than the sensitivity of the methods employing bulk acoustic waves. Application of this method will provide real-

time process monitoring and control thereby reducing down time and increasing product quality in food, chemical, cosmetic, pharmaceutical and petroleum industry.
The SH SAW method can be computerized. This enables continuous (on-line) monitoring of the rheological parameters of a liquid "in situ" in the processing line. Small dimensions of the viscosity sensor and the absence of moving parts are substantial advantages of this method. Only an electrical lead-through is needed. Therefore, it is very easy to assemble the sensor into the high-pressure chamber.
Ultrasonic methods using SH surface waves enable the measurement of the liquid viscosity in difficult access places such as pipelines and tanks.
In general, conventional mechanical methods enable the measurement of viscosity only up to about 200 MPa. In some cases, extending the measuring range up to 400 MPa may be possible after overcoming the enormous difficulties. Classical methods are in principle mechanical methods and can not be applied "in situ".
Application of ultrasonic waves extends considerably the range of pressures (up to 1 GPa) employed during the measurement of viscosity.
We measured the viscosity of liquid not only in the exponential range but also during the phase transitions, at high pressure phase and during the decompression. This is a novelty.
The measurements of the rheological properties of liquids during the phase transitions are not possible using conventional mechanical methods. On the other hand, application of the ultrasonic methods enables both the detection of phase transitions and investigation of their kinetics. This makes it also possible to determine the changes in microstructure occurring during the phase transitions.
To the author's knowledge, the measurements of liquid viscosity and sound speed under high-pressure during the phase transition and during the pressure decompression have not been reported in the scientific literature. This is an original author's contribution.
In future research it would be desirable to measure the physical properties of liquids in function of temperature and pressure. In this way, additional thermodynamic parameters could be determined (e.g., molar volume or adiabatic specific heat). The measurements should be extended into the non-Newtonian liquids (e.g., slurries, drilling fluids). In future the developed method could be applied in industry for on-line operation to investigate lubricants and bio-fuels (bio-diesels) and to control the food processing.

10. References

Achenbach J.D. (1973). *Wave propagation in elastic solids*. North-Holland, ISBN: 0 7204 2367 8, Amsterdam.

Andle J.; Haskell R.; Chap M. (2008). Electrically isolated thickness shear mode liquid phase sensor for high pressure environments. (2008) *IEEE International Ultrasonics Symposium Proceedings*, pp. 1128-1133, ISBN: 978-1- 4244-2480-1, Beijing, China, November 2008.

Auld B.A. (1973). *Acoustic Fields and Waves in Solids*, Wiley, ISBN: 0-471-03700-1, New York, Vol. II, Chap. 12.

Bair S.; Jarzynski J.; Winer W.O. (2001). The temperature, pressure and time dependence of lubricant viscosity. *Tribology International*, Vol. 34, No 7, (July 2001) (461-468), ISSN: 0301-679X.

Ball S.J.; Goodwin A.R.H. & Trusler M.J.P. (2002). Phase behavior and physical properties of petroleum reservoir fluids from acoustic measurements. *Journal of Petroleum Science and Engineering*, Vol. 34, No 1, (2002) (1-11), ISSN: 0920-4105.

Ballantine D.S.; White R.M.; Martin S.I.; Ricco A.J.; Zellers E.T.; Frye G.C.; Wohltjen H. (1997). *Acoustic Wave Sensors*, Academic Press, ISBN: 0-120-77460-7, London.

Bamberger J.A.; Bond L.J.; Greenwood M.S. (1999). Ultrasonic measurements for on-line real-time food process monitoring. *Sixth Conference on Food Engineering Proceedings*, 1999 AIChE Annual Meeting, Paper PNNL-SA-32024, Dallas, USA

Barbosa R.D. (2003). High pressure and temperature dependence of thermodynamic properties of model food solutions obtained from in situ ultrasonic measurements. PhD. Thesis, Graduated School of the University of Florida, USA.

Bingwell D.B.; Courtial P.; Giordano D.; Nichols A.R.L. (2004). Viscosity of peridotite liquid. *Earth and Planetary Science Letters*, Vol. 226, No. 1-2, (September 2004) (127-138), ISSN: 0012-821X.

Collings A.F.; Laughlin E. Mc. (1971). Torsional crystal technique for the measurement of viscosities of liquids at high pressure. *Trans. Faraday Soc*. Vol. 67, (1971), (340-352).

Daridon J.L.; Carrier H.; Lagourette B. (2002). Pressure dependence of the thermophysical properties of n-Pentadecane and n-Heptadecane. *International Journal of Thermophysics*, Vol. 23, No. 3 (May 2002) (698-708), ISSN: 0195-928X.

Delgado A.; Kulisiewicz. L.; Rauh C.; Benning B. (2010). Basic aspects of phase changes under high pressure. *Annals of the New York Academy of Sciences*, Vol. 1189, (March 2010) (16-23), ISSN: 1749-6632.

Del Gaudio P.; Behrens H. (2009). An experimental study on the pressure dependence of viscosity in silicate melts. *Journal of Chemical Physics*, Vol. 131, No. 4, (July 2009) (044504-1-14), ISSN: 0021-9606.

Demirbas A. (2008). Relationships derived from physical properties of vegetable oil and biodiesel fuels. *Fuel*, Vol. 87, No 8-9, (July 2008) (1743-1748), ISSN: 0016-2361.

Dzida M.; Prusakiewicz P. (2008). The effect of temperature and pressure on the physicochemical properties of petroleum diesel oil and biodiesel fuel. *Fuel,* Vol. 87, No. 10-11, (August 2008) (1941-1948), ISSN: 0016-2362.

Farnell G.W. (1978). Properties of elastic surface waves, In: *Acoustic surface waves*, Oliner A.A. (Ed.), (26-81), Springer, ISBN: 3-540-085785-0, Berlin.

Ferguson J,; Kemblowski Z. (1991). *Applied Fluid Rheology*, Springer, ISBN: 1851665889, Berlin-New York.

Fukui K.; Asakuma Y.; Maeda K. (2010). Determination of liquid viscosity at high pressure by DLS. *Journal of Physics; Conference Series*, Vol. 215, (2010), 012073-1-4, ISSN: 1742-6588.

Heydemann P.L.M.; Houck J.C. (1969). Self consistent ultrasonic method for the determination of the equation of state of liquids at very high pressure. *Journal of Applied Physics*, Vol. 40, No 4, (February 1969) (1609-1613), ISSN: 0021-8979.

Khelladi H.; Plantier F.; Daridon J.L.; Djelouah H. (2009). Measurement under high pressure of the nonlinearity parameter B/A in glycerol at various temperatures. *Ultrasonics*, Vol. 49, No 8, (December 2009) (668-675), ISSN: 0041-624X.

Hou Y.Y.; Kassim H.O. (2005). Instrument techniques for rheometry. *Review of Scientific Instruments*, vol. 76, (October 2005), 101101-1-19, ISSN: 0034-6748.

Kiełczyński P.; Płowiec R. (1989). Determination of the shear impedance of viscoelastic liquids using Love and Bleustein-Gulyaev waves. *Journal of the Acoustical Society of America*, Vol. 86, No 2, (August 1989) (818-827), ISSN: 0001-4966.

Kiełczyński P.; Pajewski W, Szalewski M.; Balcerzak A. (2004). Measurement of the shear storage modulus and viscosity of liquids using the Bleustein-Gulyaev wave. *Review of Scientific Instruments*, Vol. 75, No. 7, (July 2004) (2362-2367), ISSN: 0034-6748.

Kiełczyński P.; Szalewski M.; Siegoczynski R.M.; Rostocki A.J. (2008). New ultrasonic Bleustein-Gulyaev wave method for measuring the viscosity of liquids at high pressure. *Review of Scientific Instruments*, Vol. 79, (February 2008) (026109-1-3), ISSN: 0034-6748.

Kiełczyński P.; Szalewski M.; Rostocki A.J.; Gładysz J. (2008). Investigation of high-pressure phase transitions in castor oil using SH surface acoustic waves. *IEEE International Ultrasonics Symposium Proceedings*, pp. 2154-2157, ISBN: 978-1-4244-2480-1, Beijing, China, November 2008.

Kiełczyński P.; Szalewski M.; Rostocki A.J.; Zduniak M.; Siegoczyński R.M.; Balcerzak A. (2009). Investigation of high-pressure phase transitions in vegetable oils by measuring phase velocity of longitudinal ultrasonic waves. *IEEE International Ultrasonics Symposium Proceedings*, pp 1563-1566, , ISBN: 978-1-4244-4390-1, Rome, Italy, September 2009.

Kinh N.V.; Pajewski W. (1980). Generation of acousto-electrical waves using a source of transverse vibrations. *Archives of Acoustics*, Vol. 5, No 3, (July 1980) (261-274), ISSN: 0137-5075.

King H.E.; Herboltzheimer E.; Cook R.L. (1992). The diamond-anvil cell as a high-pressure viscometer. *Journal of Applied Physics*, Vol. 71, No 5, (March 1992) (20171-2081), ISSN: 0021-8979.

Kosinskii V.V. (2009). Effect of the major process factors on the high-pressure impregnation of rigid porous substrates with viscous media. *Powder Metallurgy and Metal Ceramics,* Vol. 48, No1-2, (January 2009) (13-20), ISSN: 1068-1302.

Koran F.; Dealy J.M. (1999). A high-pressure sliding plate rheometry for polymer melts. *Journal of Rheology*, Vol. 43, No. 5 (September 1999), (1279-1290), ISSN: 0148-6055.

Kulisiewicz L.; Baars A.; Delgado A. (2007). Effect of high hydrostatic pressure on structure of gelatin gels. *Bulletin of the Polish Academy of Sciences,* Vol. 55, (June 2007), (239-244), ISSN: 0239-7528.

Kulisiewicz L; Delgado A. (2010). High-pressure rheological measurement methods. A review. *Applied Rheology*, Vol. 20, No. 1, (January 2010), (13018-1-15), ISSN: 1430-6395.

Lan H.Y.; Tseng H.C. (2002). Study of the rheological behavior of PP/supercritical CO_2 mixture. *Journal of Polymer Research*, Vol. 9, No 3 (September 2002), (157-162), ISSN: 1022-9760.

Landau L.D.; Lifshitz E.M. (1958). *Continuum Mechanics*. Pergamon Press, London, Chap. 2.

Le Bail A.; Boillereaux L; Davenel A.; Hayert M.; Lucas T.; Monteau J.Y. (2003). Phase transitions in foods: effect of pressure and methods to assess or control phase transition. *Innovative Food Science and Engineering Technologies*, Vol. 4, No 1, (March 2003) (15-24), ISSN: 1466-8564.

Masson P.; Tonello C.; Balny C. (2001) High pressure biotechnology in medicine and pharmaceutical science. *Journal of Biomedicine and Biotechnology*, Vol. 1, No 2, (February 2001) (85-88), ISSN: 1110-7243.

Mattischek J.P.; Sobczak R. (1997). High-pressure cell for measuring the zero shear viscosity of polymer melts. *Review of Scientific Instruments,* Vol. 68, No 5, (May 1997) (2101-2105), ISSN: 0034-6748.

Matveev V.A.; O.F. Orlov O.F.; Berg V.I. (2005). Measurement of the viscosity of a liquid at high pressures. *Measurement Techniques*, Vol. 48, (October 2005) (1009-1013), ISSN: 0543-1972

Mia S.; Hayashi S.; Ohno N. (2007). High pressure tribological behavior of vegetable oils as lubricant. *Proceedings of the International Conference on Mechanical Engineering*, (29-31 December 2007), Dhaka, Bangladesh, Paper Id FL-07.

Mineev V.N.; Funtikov A.I. (2005). Measurements of the viscosity of water under shock compression. *High Temperature*, Vol. 43, No 1, (January 2005), (141-150), ISSN: 0018-151X.

Nakamura K. (2007). Shear-horizontal piezoelectric surface acoustic waves. *Japanese Journal of Applied Physics*, Vol. 46, No 7B (July 2007), (4421-4427), ISSN: 0021-4922.

Nakamura Y.; Kurosaki Y. (2005) Micro-rheometry of pressurized lubricants and micro-nanorheology. *Microsystems and Technology*, Vol. 11, (August 2005), (1127-1131), ISSN: 0946-7076.

Nithya R.; Nithyanathan S.; Mullainathan S.; Rajasekaram M. (2009). Ultrasonic investigation of molecular interactions in binary mixtures at 303 K. *E-Journal of Chemistry,* Vol. 6, No. 1, (January 2009) (138-140), ISSN:0973-4945.

Oakley B.; Barber G.; Worden T.; Hanna D. (2003). Ultrasonic parameters as a function of absolute hydrostatic pressure. I. A Review of the data for organic liquids. *Journal of the Physical and Chemical References Data*, Vol. 32, (April 2003) (1501-1533), ISSN: 0047-2689.

Oakley B.; Hanna D.; Shillor M.; Barber G. (2003). Ultrasonic parameters as a function of absolute hydrostatic pressure. II. Mathematical models of the speed of sound in organic liquids. *Journal of the Physical and Chemical References Data*, Vol. 32, (April 2003) (1535-1544), ISSN: 0047-2689.

Philippoff W. (1963). Viscoelasticity of polymer solution at high pressures and ultrasonic frequencies. *Journal of Applied Physics*, Vol. 34, (May 1963) (1507-1511), ISSN: 0021-8979.

Plantier F.; Daridon J.L.; Lagourrette B. (2002). Measurement of the B/A nonlinearity parameter under high pressure: Application to water. *Journal of the Acoustical Society of America*, Vol. 11, No 2, (February 2002) (707-715), ISSN: 0001-4966.

Ripple D. (1992). A compact high-pressure capillary viscometer. *Review of Scientific Instruments*, Vol. 63, (May 1992), (3153-3155), ISSN: 0034-6748.

Rostocki A.J.; Wiśniewski R.; Wilczyńska T. (2007). High-pressure transition in rape seed oil. *Journal of Molecular Liquids,* Vol. 135, (February 2007) (120-122), ISSN: 0167-7322.

Rostocki A.J.; Siegoczyński R.M.; Kiełczyński P.; Szalewski M. (2010). An application of Love waves for the viscosity measurement of triglycerides at high pressure. *High Pressure Research,* Vol. 30, No 1, (January 2010), (88-92), ISSN: 0895-7959.

Royer D.; Dieulesaint E. (2000). *Elastic waves in solids*. Springer, ISBN 3-540-65932-3, Berlin.

Royer J.R.; Gay Y.J; Adam M.; Simone de J.M.; Hakan S.A. (2002). Polymer melt rheology with high-pressure CO_2 using a novel magnetically levitated sphere rheometer. *Polymer*, Vol. 43, No 8, (April 2002) (2375-2383), ISSN: 0032-3861.

Ruttle S.G.R.; Stephenson M.I. (1975). A high-pressure ultrasonic viscometer. *Ultrasonic International Conference Proceedings*, IPC Sci. Technol., (1975) (224-227), Ed. Browne L.J., Guilford, England.

Schaschke C.J.; Abid S.; Flether J.; Heslop M.J. (2008). Evaluation of a falling sinker-type viscometer at high pressure using edible oil. *Journal of Food Engineering*, Vol. 87, No 1, (January 2008) (51-58), ISSN: 0260-8744.

Schettino V.; Bini R.; Ceppatelli M.; Citron M. (2008). Activation and control of chemical reactions at very high pressure. *Physica Scripta,* Vol. 78, No. 5, (November 2008) (058104-1-15), ISSN: 0031-8949.

Shames I.H. (2002). *Mechanics of Fluids*. Mc Graw-Hill, ISBN: 0071198899, New York.

Sha Zhen-Shun. (1997). The improvement on the falling ball viscometer. *Review of Scientific Instruments*, Vol. 68, (April 1997), (1809-1811). ISSN: 0034-6748.

Sonneville-Aubrun O.; Simonnet J.T.; L'Alloret F. (2004). Nanoemulsions: a new vehicle for skincare products. *Advances in Colloid and Interface Science*, Vol. 108-109, No 1-3, (May 2004) (145-149), ISSN: 0001-8686.

Stallard J.M.; Rosenbaum I.J.; Davis C.M. (1969). Ultrasonic method for determining thermodynamic properties of liquids with results for mercury. *Journal of the Acoustical Society of America*, Vol. 45, No 3, (March 1969) (583-586), ISSN: 0001-4966.

Sugasawa S. (2002). Time difference measurement of ultrasonic pulses using cross-correlation function between analytic signals. *Japanese Journal of Applied Physics*, Vol. 41, part 1, No 5B, (May 2002) (3299-3307), ISSN: 0021-4922.

Viola F.; Walker W. (2003). A comparison of the performance of time-delay estimators in medical ultrasound. *IEEE Trans. on Ultrasonics, Ferroelectric and Frequency Control*, Vol. 50, No. 4, (April 2003) (392-401), ISSN: 0885-3010.

Zhang C.; Caron J.J.; Vetelino J.F. (2001). The Bleustein-Gulyaev wave for liquid sensing applications. *Sensors&Actuators B*, Vol. 76, No 1-3 (June 2001) (64-68), ISSN: 0925-4005.

Zhao S.Y.; Yan J.N.; Shu Y.; Zhang H.X. (2008). Rheological properties of oil-based drilling fluids at high temperature and high pressure. *Journal of South University of Technology*, Vol. 15, No 1, (September 2008) (457-461), ISSN: 1005-9784.

15

Pressure and Temperature Microsensor Based on Surface Acoustic Wave in TPMS

Tianli Li[1], Hong Hu[2], Gang Xu[1], Kemin Zhu[1] and Licun Fang[1]
[1]*Shenzhen University,*
[2]*Harbin Institute of Technology Shenzhen Graduate School*
China

1. Introduction

The tire pressure monitoring system (TPMS) can not only make the driver more safety, but also save fuel and protect the tire. Tire safety is attracting the driver's attention, the United States had developed laws to enforce the TPMS installation in the car. In this paper, the basic structure and the implement method of TPMS are introduced. The SAW theory and some surface acoustic wave (SAW) temperature and pressure sensors which suit for the TPMS application are illustrated, because the passive sensor is becoming the focus in the TPMS research field. Passive SAW sensor is the good choice for TPMS, according to its wireless, passive, small size, zero age rate etc. The wireless passive SAW TPMS is one of the most important research direction.

For the typical applications of automotive TPMS, a novel microsensor based on SAW is reported in this paper. The kernel structure and design theory of this sensor with a single sensing unit are introduced. With the theory of SAW delay line, the effects of temperature and pressure on the microsensor were able to be reflected by the variations of the radio frequency (RF) echo signals. The accurately measured temperature and pressure values were obtained by using of a weight factor in the data process. The excellent agreement between the pressure and temperature results measured by the sensor and the direct measurement data is presented. The practical results in the certain ranges of pressure and temperature demonstrated that the microsensor is able to measure temperature (0kPa-200kPa) and pressure (20^{o}C-100^{o}C) at the same time. For SAW sensors the temperature measurement accuracy can reach 0.05^{o}C, and the pressure measurement accuracy can reach 7.2kPa. In the areas of TPMS, where reliability and durability are really demanding, the reported microsensor has its practicability and potential market with its advantages of simple structure, and wireless and passive working mode.

2. Background of tire pressure monitoring system

Security, economy and comfort of the automobile are the basic demands for the customers. Automakers have made significant improvements on vehicle safety in recent years, such as Antilock-Braking System (ABS), Traction Control System (TCS), 4 Wheel Steering (4WS) and Electronic-Stability Program (ESP) in today's vehicle products. The applications of these technologies can contribute to highway-accident reduction, i.e., ensuring that the vehicle is in good condition for use.

Tire condition is one of the greatest contributors to safety. Paradoxically, studies in the United States reveal that although about 85% of the population recognizes the importance of maintaining properly inflated tires, most drivers wait until the vehicle's service interval to have a car service person check the tires. Tire pressure may have an influence on any crash that involves braking, since low tire pressure can result in increased stopping distances. Road safety can cost as many as 40,000 lives every year across Europe.

With vehicle-service intervals that now approach 50,000 km for some models, lack of attention to routine maintenance clearly creates unnecessary hazards. Under-pressure tires sometimes are not easy to be observed, because the circumference changes relatively little from 1 to 3 bar of internal pressure, and 30% under-pressure is nearly undetectable to a casual observer. Fig. 1 shows that it is difficult for an ordinary driver to tell the difference between two tires with different pressures (David, 2004). A study from the U.S. DoT (Department of Transportation) that surveyed 11,530 vehicles nationwide showed that 27% of passenger cars and 33% of light trucks operated with one or more substantially under-pressure tires. The consequences of tire under-pressure are increased fuel costs, reduced tire and tread life, tire blowouts and tread separation, dual wheel assembly problems and additional high temperature problems. Incorrect tire pressure can affect the handling, braking and stability of a vehicle and, according to industry statistics, may be responsible for hundreds of thousands of accidents each year. One of the automotive safety requirements that is becoming more prominent is the use of TPMS, which alerts the driver when tire pressure or temperature drops to an unsafe level, just like the warning lights that signal low fuel level.

Fig. 1. A fully inflated tire (left) and one with only 70% pressure (right) (David, 2004)

Unsurprisingly, vehicle manufacturers and regulatory authorities are keenly interested in techniques for TPMS. This fits to visions and strategies of vehicle manufacturers, automotive suppliers and tire manufacturers. This development was mainly driven by vehicle manufacturers. Although automakers recognized this requirement as long ago as 1970, but a cost-effective solution simply wasn't available at that time. Today's technologies can satisfy the requirement in an affordable way, so the U.S. DoT mandates that most vehicles built from now on should carry a TPMS, and all vehicles sold in the United States must be equipped with TPMS starting from 2007. Key legislation appeared in the TREAD (Transportation Recall Enhancement, Accountability, and Documentation) Act of November 2001, with main support work coming from the U.S. NHTSA (National Highway Traffic Safety Administration).

TPMS not only provides enhanced safety for drivers and passengers, offers more convenience, but also eliminates guesswork to the tire pressure and temperature. TPMS will warn drivers when a tire is abnormal, preventing tire damage and subsequent accidents. TPMS will not only help prevent accidents, but save as much as $1.7 billion each year in fuel and vehicle maintenance costs, as under-pressure tires shorten tire life and increase fuel consumption, according to NHTSA. TPMS meets the demand for improved vehicle safety, performance, reliability and fuel efficiency. The future business for TPMS is set to be huge. Consultant firm strategy analytics expects TPMS to become the fastest-growing automotive electronics system, reaching 30 million units by 2010.

3. TPMS implementation methods

TPMS turns tire pressure checking into an automatic process. TPMS can measure the tire pressure and temperature parameters automatically utilizing sensors inside the tires at any time. According to the measurement results, TPMS can give drivers an alarm when the pressure or temperature in tires is abnormal so that the drivers can take the appropriate corrective action (Wang et al., 2003). In this way, the tire pressure and temperature are kept at normal level, and the resistance of tire rolling is reduced, tire abrasion and fuel consumption are decreased, tire's service-life is increased, so the economy and safety of the car are improved greatly.

Currently, two methods of TPMS are being developed, which are indirect TPMS and direct TPMS (David, 2004). Direct TPMS measures the tire pressure and temperature through the sensors which are installed in the tires. Indirect TPMS obtains tire pressure through external software algorithms by analyze rotational speed of each wheel.

3.1 Indirect TPMS

Indirect methods use wheel speed sensors and ECU of the ABS system which are already existed in the car to infer low tire pressure by looking for a wheel that is spinning faster than the others. The technique works by comparing the rotational speed of each wheel in normal driving mode, since a tire's rolling radius depends on the air pressure inside. This method minimizes implementation cost by taking advantage of the fact that ABS appears in virtually every vehicle product. However, the radius also depends on many other variables, which do not make ABS-based TPMS very reliable. Problems that might occur using indirect TPMS are listed as follows:

1. The system needs calibration before it can sense different tire conditions. In addition, tire changing requires resetting the system to relearn the dynamic relationship between each wheel.
2. In a test carried out by the U.S. DoT, an indirect TPMS didn't detect some kinds of low tire pressures when two low-pressure tires were on the same side or axle, or all four tires were under-pressure to a similar degree.
3. Slip at the wheels disturbs the pressure-sensing algorithm.
4. Speed, acceleration, uneven tire wear and production tolerances affect rolling radius.
5. The system is unable to detect tire deflation of typically less than 30%.

3.2 Direct TPMS

The direct TPMS utilizes sensors installed inside tires to measure and feedback the pressures and temperatures directly. Wireless technologies for data transmission have to be used,

because the wheel is a rotating system which can't be connected by a wire. The direct TPMS uses RF technology for transmitting sensor data to the vehicle. The most commonly used frequency for transmitting the measured tire information to the receiver is about 433 MHz. This frequency can be freely used in Europe. In the United States a similar license-free frequency is 315 MHz. The receivers consist of antenna, processor, memory and a user interface. Products for the aftermarket usually have a receiver that contains all of these components in one box.

The direct TPMS can be classified as three classes according to the sensor installation place. First, clamp-on-rim sensors can be installed on the well bed of the rim with a stainless steel clamp, which is showed in the left of Fig. 2. This fixing method can usually be used in aftermarket products, when the same product must suit a large variety of cars. Secondly, valve-attached sensors can be fixed on the bottom end of the tire valve, as shown in the middle of Fig. 2. In this case, the sensor is actually located on the very same spot as when using a clamp on the rim well bed. Different rims require different valves, which means that this fixing method is better for the original equipment market, where it only has to fit a specific car model with limited variety of wheel types. Thirdly, valve-cap-integrated sensors are to try and squeeze the sensor electronics inside a valve cap, which is showed in the right of Fig.2. This fixing method is easy and suitable for aftermarket and especially heavy vehicles, but this method may make the tire sensor more dangerous for exposure outside.

Fig. 2. Picture from three types

There are two types of direct TPMS according to the power method of the sensor. One type uses the active sensors, that is to say the TPMS contains a component for electric power supply by batteries. The main components of an active battery operated in TPMS are battery, processor, memory, sensors, radio component, and antenna. Some of these components are usually integrated on a single chip to save weight and space and to reduce power consumption. The battery is the most problematic component. It limits the operation time of the sensor, and the existing batteries have also temperature limitations. In very cold or high temperature they may not work properly or may even be destroyed completely. The other type uses the passive sensors, that is to say the sensor needn't use battery as a power, the sensor can get the working energy from the others methods such as the RF singles or an generator near the sensors.

3.3 Comparison between indirect TPMS & direct TPMS

Compared with the indirect approach, the direct method demonstrated its advantages in many aspects. First, direct measurement provides a much more accurate indication of tire pressure. Secondly, it can report individual tire status. Thirdly, the signal is always available even when the vehicle is parked. The direct TPMS can ensure tires are properly inflated and running optimally at all the time. The direct method also has downsides that include higher installation cost, a limited battery life of five to ten years, and the potential for sensor damage during the installation of spare tires or through driver error by kerbing the tire. From Table 1, one can see clearly about the difference between indirect TPMS and direct TPMS.

Performance	Direct	Indirect
Accuracy	Higher	Lower
Detect Multiple Low Tires	Yes	No
Affected by Acceleration	No	Yes
Affected by Loose Surfaces	No	Yes
Speed Range	0 to vehicle max	20 to 110 km/hr
Incremental Cost	Higher	Lower
Vulnerable to Tire Mount Process	Yes	No
Power Source at Each Wheel	Yes	No

Table 1. PMS Implementation Methods(Freescale Semiconductor, Inc.)

4. Products and researches for TPMS

A lot of works have been done on TPMS in the USA, Japan, Germany, and UK. Austria. Continental corp., Siemens corp., and Nokia corp. have provided some productions of TPMS (Wang et al., 2003). As required by the market demands, almost all of the companies are focus on direct TPMS products, only a few companies concentrate on indirect TPMS. The tire manufacturer Continental AG is one of the companies who supplies indirect TPMS products. Its product is called Deflation Detection System (DDS) which is installed in BMW M3.

There are many companies that have done a lot of work on direct TPMS. For example, Freescale Semiconductor (USA), Philips Semiconductors and Sensonor (Norway) manufacture commercial components for TPMS products. SmarTire (Canada) is the oldest supplier in the TPMS market, its products can be ascended to 1990. RoadSnoop Pressure Watch is made from the Finnish tire manufacturer Nokian Tyres plc. Both of them have aftermarket products as shown in Fig. 3, and their sensors are fixed on the rims with stainless steel bands. SmarTire's receiver is fixed on the dashboard or on the windscreen and connected to the cigarette lighter of the car for power supply. RoadSnoop's receiver is a small wireless battery-operated device, which can be put on any place inside the car, where it can be easily seen and heard. Schrader Electronics from UK is the manufacturer of the standard valve-attached TPMS sensors which are used in cars like Aviator, Pathfinder, Peugeot 607, Citroen C8, and etc.. Information of lots of products similar to the Schrader's can be found on the Internet. These products also contain valve fixed sensors, such as Beru

and Siemens VDO Automotive from Germany, Pacific Industrial and Omron from Japan. U.S. companies such as Fleet Specialties with "Tire Sentry" and Advantage Enterprises with "Pressure Pro" have the valve-cap-integrated tire sensors used in the direct TPMS.

Fig. 3. SmarTire's TPMS product (Left) & RoadSnoop's Pressure Watch (Right)

The direct TPMS using passive sensors is not available in commercial products. The German IQ-Mobil GmbH is developing a batteryless TPMS, which is called RDKS. This product is only available in prototype. Fig. 4 shows the size of the chip and how the transponder is mounted on the tire valve.

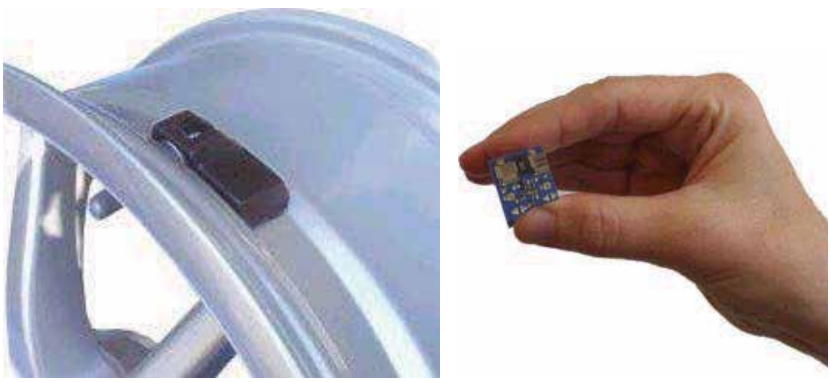


Fig. 4. IQ-Mobil's sensor attached on a tyre valve and the size of the electronics

More technical details can be found about TPMS sensors. For instance, Freescale Semiconductor introduces the MPXY8020 tire pressure monitoring sensor, SensoNor uses SP12T sensor, Philips Semiconductors has P2SC (Philips' signal conditioning chip) family, etc. The MPXY8020 sensor is comprised of a capacitive pressure sensing element, a temperature sensing element, and an interface circuit with the wake-up feature. All the parts of the sensor are integrated on a single chip. Some companies' TPMS is similar to MPXY8020, the others produce pressure sensor elements using the piezoresistive technology. Some detailed features about the sensors are shown in Table 2.

Sensor Model (Company)	Max . Operating Pressure	Pressure Resolution	Pressure Accuracy	Temperatur e Range	Temperatu re Accuracy
MPXY8020A (Freescale)	637.5 kPa	2.5 kPa	±7.5 kPa	-40 °C to +125 °C	±4 °C
Sensor (SmarTire)	538 kPa	--	±10 kPa		±3 °C
SP12T (SensoNor)	1400 kPa	2.97 kPa	±28 kPa		±4 °C
Pressure Watch (RoadSnoop)	350 kPa	--	±12 kPa	-40 °C to +120 °C	±6 °C
RDKS (IQ-Mobil)	1200 kPa	0.2 kPa	--	-40 °C to +170 °C	--

Table 2. Comparison of some commercial sensors based on TPMS

Up to now three different techniques of information transmission have been applied to direct TPMS: radio transmission employing an active sensor unit inside the tire, inductive transmission and radio transmission via reflection (passive transponder). Currently, the TPMS consisting of an active, battery powered sensor unit inside each wheel dominates the market. A typical TPMS product contains active semiconductor circuits, a sensor circuit, a wake-up unit and a transceiver unit. All the units are powered by a lithium battery which limits the lifetime of the sensor units, i.e., the battery must be replaced when the power was run out. Although the TPMS companies declare that the battery can be used 5 to 10 year for normal vehicle use, tire sensors will use the power of the battery more quickly if the car keeps working. It will be more complex to design if it contain active semiconductor circuits and the wake-up unit in each of the tire, and will make the sensor part heavier. In order to guarantee a life time of at least 5 to 10 years the battery needs to have several hundred mAh capacity, which causes the battery to be relatively big and heavy. Furthermore, in some cases, remote sensor systems are affected by strong thermal, mechanical, or electromagnetic loads so that batteries, semiconductors, and active elements are likely to be damaged. So a better way should be found to avoid these problems. The obvious solution would be to replace the battery with some other component to get energy, because the electrical power supply of sensor is necessary. For example, the battery can be replaced by an inductive transmission or a local power generation. The main benefits are the increase of reliability and environmental friendliness and the reduction of maintenance efforts. Surface Acoustic Wave sensors are a good choice. The SAW sensor needn't the power supply unit and the wake-up unit, and only an antenna is needed for the transceiver unit. This means that the sensing devices gets the necessary energy from the radio signal which is obtained through the antenna. The circuit design is much simpler than the traditional sensor of TPMS.

The recent research mainly focus on developing the TPMS with wireless passive SAW sensors. The SAW sensor is small, light, reliable, stable, passive, and sensitive. It is not affected by strong thermal, mechanical and electromagnetic loads. From Table 3, one can see the differences between the active sensor and the passive sensor.

In addition, the SAW sensor works very well in bad environment, closed chambers, moving and rotating parts of engine. So it can be embedded in the surface of the tire or fixed around the rim. It is one of the best choices for the TPMS sensors. In 1996, Alfred Pohl and F.Seifert started to research the wireless passive SAW sensor used in the tires in University of Technology Vienna (Pohl & Seifert, 1997). They designed the SAW pressure sensor based on

	Active Sensor	Passive Sensor
Measuring Range	Same	
Operating Temperature Range	Same	
Survival Temperature Range	Same	
Basic Principle	Capacitive	SAW
Size	Large	Small
Weight	Heavy	Light
Typical TX Range	Long	Short
Wake Up	Yes	No
Battery Needed	Yes	No
Design Complexity	High	Low
Aging Rate	Normal	Low
RF Transmitter	Yes	No

Table 3. The comparison between an active sensor and a passive sensor

the SAW delay line and verified the design. But the signal processing was very difficult because of the noise disturbance, and they didn't go further in this aspect. In the same year, W. Buffll and M. Rusko et al. designed the wireless passive SAW sensor based on two SAW resonators with different frequencies in Germany (Buff et al., 1998). This sensor had better precision and not affected by movement and rotation. Since the sensor was sensitive to both pressure and temperature, it was affected by cross-disturbance. In 1998, Reinhard Steindl and Alfred Pohl designed the SAW hybrid sensor based on the combination of the SAW delay line and the conventional pressure sensor (Steidl et al., 1998). This sensor could measure both pressure and temperature, and its precision was high, but there were questions in practical realization and signal processing.

In conclusion, the research of wireless passive SAW sensor is still in the early period and currently there is no manufacturing solution available because of questions in theory and technique. In order to apply the SAW sensor in TPMS, not only the design methods of wireless passive SAW sensors but also the arithmetic in dealing with the feedback sensing data should be considered. In recent years, some novel materials and technologies have be developed in this field, thus it is possible for wireless passive SAW sensors to make better performance through novel sensor design and new fabrication technology.

5. Application of wireless passive SAW sensors in TPMS

A schematic drawing of the TPMS as an example of a wireless SAW hybrid sensor system is shown in Fig. 5 As the functional principle of the wireless SAW sensor system has already been described (Schimetta et al., 2000) , only a short survey should be given here. The measurement cycle is initiated by a high frequency electromagnetic burst signal emitted from the wheel arch antenna of the central transceiver unit. This signal is received by the antenna of the SAW transponder unit mounted on the rim. The IDT connected to the antenna transforms the received signal into a SAW. In the IDT the reflected acoustic waves which include the sensor information are reconverted into an electromagnetic pulse train

and retransmitted to the central transceiver unit, where the received signal is amplified, down converted and analyzed. The antennae of the transceiver were set at every wheel arch and connected with the transceiver with twisted-pair. The transceiver sends wireless signals with every antenna to the SAW sensors in the tires and receives the reflected signals from the SAW sensors in the tires. In addition, the transceiver sends the received signals to the computer and display unit by CAN bus. The signals are processed in the computer unit and the tire state is displayed in the display unit.

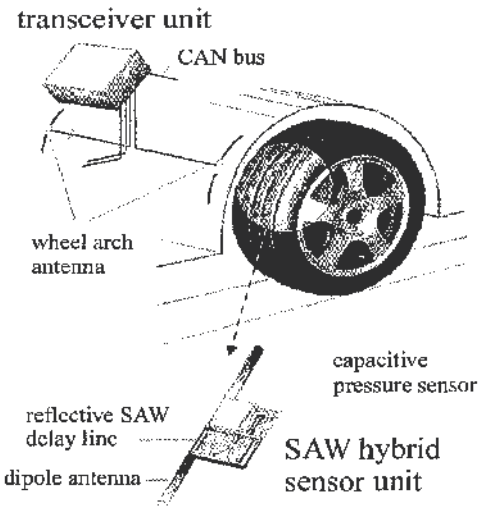


Fig. 5. Schematic drawing of a SAW sensor system applied to TPMS (Schimetta et al., 1997)

The transceiver begins to send and receive pulse signals periodically as long as the car starts. In every period, firstly the transceiver sends the RF interrogation signal to the first tire sensor and receives its reflected signals, then the transceiver sends the signals to the second tire sensor and receives the return signals from the second tire sensor. In this way the transceiver does on the third and the fourth tire sensors. The tire code, pressure and temperature information are all included in the reflected signals. The computer unit processes these reflected signals. First of all, it recognizes the tire code and calculates the tire pressure and temperature, then stores the data as the tire state information, finally every tire pressures and temperatures are averaged in some periods as each tire pressure and temperature. The differences between the tire pressure, temperature and the correct values are calculated. The alarm is given to the driver in the display unit if the difference is out of the secure valve, otherwise only the pressure and temperature are displayed in the display unit.

5.1 Principles of wireless SAW sensors

The applicability of passive SAW devices for remote sensing was found for decade years. SAW sensors can be built with a SAW delay line element connected to an antenna. The SAW delay line consists of a substrate, an interdigital transducer (IDT), and a reflector. The working sequence of the wireless passive SAW sensor are illustrated in Fig. 6:

1. The transceiver sends RF interrogation signal which is received by the antenna of the SAW sensor.

2. The IDT which is connected to the antenna, transforms the received signal which is an electrical RF voltage applied between the two opposing electrode combs into a SAW.
3. The SAW propagates on the piezoelectric crystal and is partially reflected by reflectors placed in the acoustic path.
4. The reflected waves are reconverted into an electromagnetic pulse train by the IDT and are retransmitted to the radar unit.
5. The high frequency electromagnetic signal is amplified and down converted to the base band frequency in the RF module of the radar unit.
6. Then the sensor signals are analyzed with a digital signal processor.
7. Finally the measurement results can be transferred to a personal computer for post processing and data storage.

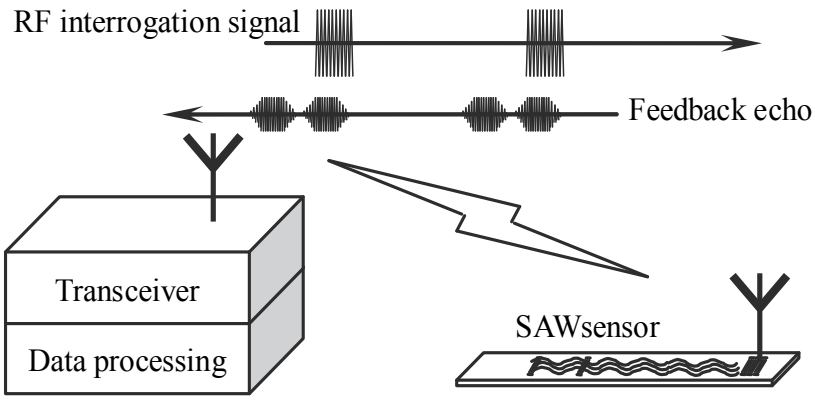


Fig. 6. Principle of a wireless SAW sensor

Fig. 6 illustrates suggested principles for SAW remote sensor device, which basically can be utilized in two different ways. The sensor signal can be produced by SAW device itself which means that the delay time is varied due to, e.g., varying temperature or applied pressure causing stress and a deformation of the device. Alternative configurations for this approach include the application of chirp-transducers and SAW resonators (Reindl et al., 1998). Another sensor device, which changes its impedance under the influence of the quantity to be sensed, is attached to a second IDT acting as reflector structure. This load impedance determines the amplitude and phase of the reflected SAW burst (Steidl et al., 1998).

The velocity of a SAW is approximately the factor 100 000 smaller than the velocity of light or radio signals. Therefore the propagation velocity of SAW allows a long delay time to be realized within a small chip. A time delay of 1 us requires a chip length between 1.5mm and 2mm, depending on the substrate material which cause the different SAW transmitting velocity, whereas in 1us a radio signal propagates 300m in free space. Therefore, pulse response of SAW sensors with time delays of several microseconds can be separated easily from environmental echoes, which typically fade away in less than 1-2us. If the reflectors are arranged in a predefined bit pattern like a bar code an RF identification system can be realized with a readout distance of several meters. SAW transponders are small, robust, inexpensive, and can withstand extreme conditions. Fig. 7 shows a typical response signal of a SAW ID-tag together with the interrogation impulse and environmental echoes (Reindl et al., 1998).

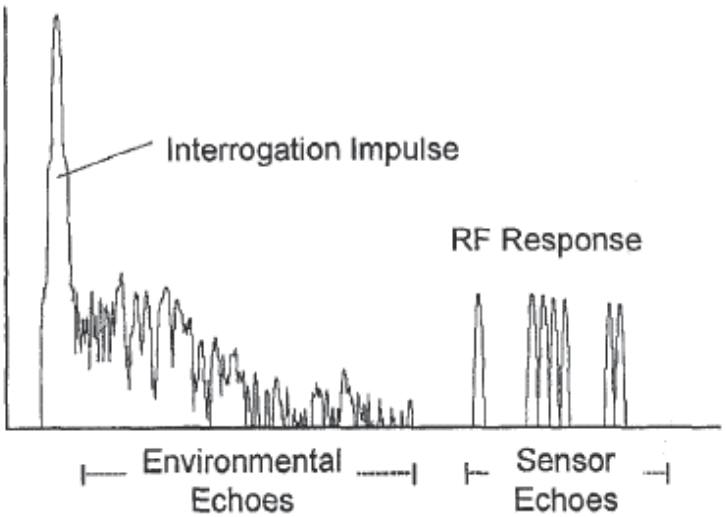


Fig. 7. Interrogation pulse, environmental echoes, and RF response of a SAW reflective delay line (Reindl et al., 1997).

5.2 Wireless passive SAW sensors

A schematic drawing of a SAW pressure sensor is shown in Fig. 8 The SAW propagates on a quartz diaphragm, bending under hydrostatic pressure. To bend the diaphragm in a defined manner, there has to be a constant reference-pressure at the other side of the diaphragm. This is realized by a hermetically closed cavity with the reference pressure inside. Therefore with a sand-blast unit a blind-hole was structured into a quartz cover plate, which is of the same substrate material as the diaphragm (Scholl et al., 1998).

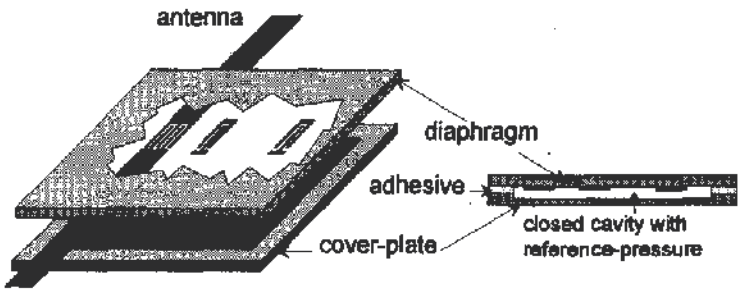


Fig. 8. Schematic drawing of a SAW pressure sensor (Scholl et al., 1998)

A monolithically packaged SAW radio transponder and pressure sensor are developed for the application to a TPMS (Oh et al., 2008), showed in Fig. 9 The device contains the wireless transponder, which converts analog signal into digital one without any auxiliary electronic circuits and transmits the converted data wirelessly. The realization of the mechanical A/D conversion is possible since the SAW radio transponder is connected to the touch-mode capacitive pressure sensor. The SAW radio transponder and touch-mode sensor are fabricated using a surface micromachining and a bulk micromachining technologies, respectively. The performance of the integrated, passive and wireless pressure sensor meets

the design specifications such as linearity, sensitivity and noise figure. This approach can increase the accuracy of signal detection, if more A/Ds are used, but the number of the A/D are restricted by the MEMS fabrication method, so the sensor can not reach the high accuracy. Paper (Schimetta et al., 2000) proposed the concept of using hybrid sensors to achieve the pressure sensor, includes SAW sensor and the corresponding non-contact capacitive pressure sensor, the corresponding matching circuit are needed between them. The sensing structure relatively complex, and can only measure pressure changes.

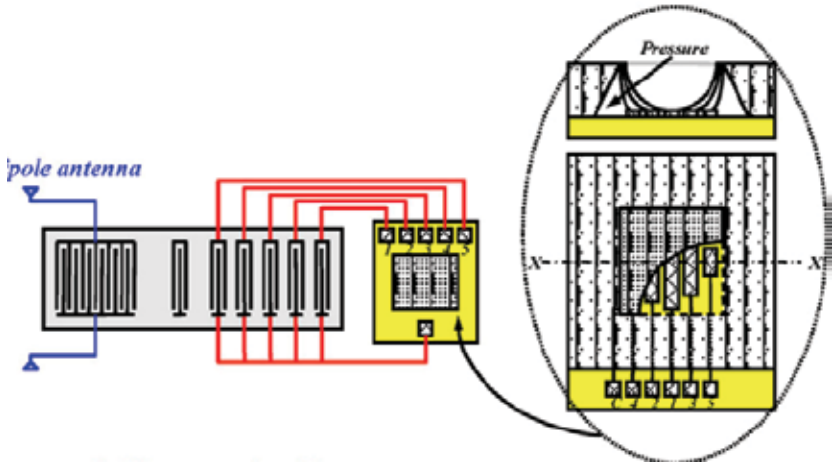


Fig. 9. A schematic illustration of embedded MEMS A/D converter with SAW wireless transponder (Oh et al., 2008).

An U.K. company Transense is developing SAW sensor technology for tire monitoring purposes. It's sensor uses the SAW device as a diaphragm between the side of the sensor subjected to tire pressure and a sealed reference chamber. The energy needed is provided from the signal of the receiver component. The Triple SAW Pressure Device provides temperature compensated pressure measurement from a single quartz die operating in a simple bending mode. Fig. 10 shows how the SAW sensor is used in TPMS.

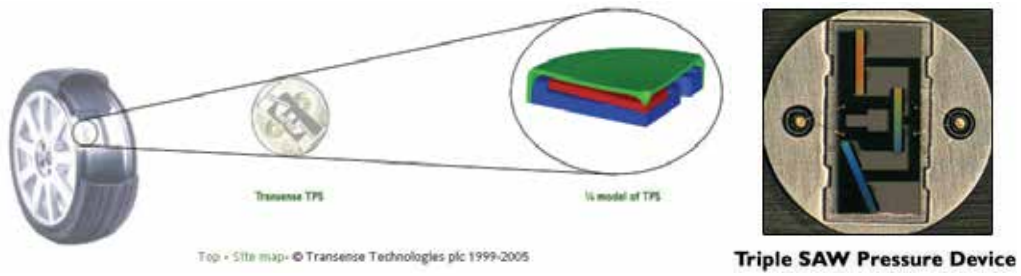


Fig. 10. SAW sensor used in TPMS.

The important of TPMS is introduced, and the TPMS implement method is discussed in this section. For the disadvantage of active sensor used in TPMS, this paper introduced some kinds of wireless passive SAW sensors. The wireless passive SAW pressure and temperature sensor with single sensing unit is showed. The SAW sensor has the simple structure and small size compared with the active TPMS sensor. The passive SAW sensor will replaced the active sensor used in TPMS in the future due to its advanced features shows in this paper.

6. A novel pressure and temperature SAW microsensor

Typical applications of surface acoustic wave (SAW) sensors using MEMS technology for the measurement of temperature (Kim et al., 2004) (Bao et al., 1987) and pressure (Schimetta et al., 2000) (Oh et al., 2008) have been studied for years. Due to their advantages of wireless and averting the need for power supply at the sensor location, SAW sensors are able to be used in such moving and harsh conditions as tire pressure monitoring (Ballandras et al., 2006). In practical applications, such as tire pressure monitoring systems, it is necessary to measure both pressure and temperature simultaneously. The common solution is to use more than one sensing units to measure pressure and temperature separately, in which case, however, the whole structure of the SAW sensor is complicated for manufacturing and packaging. The preliminary design theory of a novel wireless and passive SAW microsensor, which comprises single sensing unit and is able to measure real-time pressure and temperature accurately was suggested by the authors recently (Li et al., 2008). In this letter, further investigation on this novel sensor is to be reported both in theory analysis and practical test. In the following sections, the design theory and test results for the SAW sensor will be described.

6.1 Design and theory for SAW microsensor

The SAW microsensor in this letter comprised an interdigital transducer (IDT), three reflectors, R_1, R_2, and R_3, on the top surface of a piezoelectric substrate. The schematic diagram of the sensor structure is shown in Fig. 11. The three reflectors located on the both sides of the IDT, such a design being able to minimize the energy loss of echo signal from each reflector. d_1, d_2, and d_3 are the distances between the IDT and R_1, R_2, and R_3, respectively. The double values of the traveling time differences of SAW signal, τ_{12}, τ_{13} between R_1 and R_2, R_1 and R_3 can be respectively defined as Equation (1):

$$\tau_{1i} = \frac{2 \bullet (d_i - d_1)}{v} = \frac{2 \bullet d_{1i}}{v} \quad (i = 2, 3) \tag{1}$$

where, d_{12} and d_{13} are the differences between d_2 and d_1, d_3 and d_1, respectively, v the propagation velocity of SAW signal. The phase differences φ_{12} and φ_{13} between the echo signals reflected by R_2 and R_1, R_3 and R_1 are defined as Equation (2):

$$\varphi_{1i} = \omega_0 \tau_{1i} \quad (i = 2, 3) \tag{2}$$

where ω_0 is the angular frequency of RF pulse signal.

The Part A bottom of the piezoelectric substrate was attached on the sensor package while Part B was left free to form a cantilever for pressure measurement. The dimensions of the whole piezoelectric substrate, including Parts A and B, are the function of circumstance temperature. For Part A of the substrate, τ_{12} is the function of temperature change ΔT and can be described as Equation (3) (Bao et al., 1987) [2]:

$$\tau_{12}(\Delta T) = \tau_{12}^0 (1 + \alpha \Delta T) \tag{3}$$

where α is the temperature coefficient of the SAW device substrate, τ_{12}^0 the initial value of τ_{12} under initial temperature. Combining Equations (1), (2), and (3), Equation (4) is obtained.

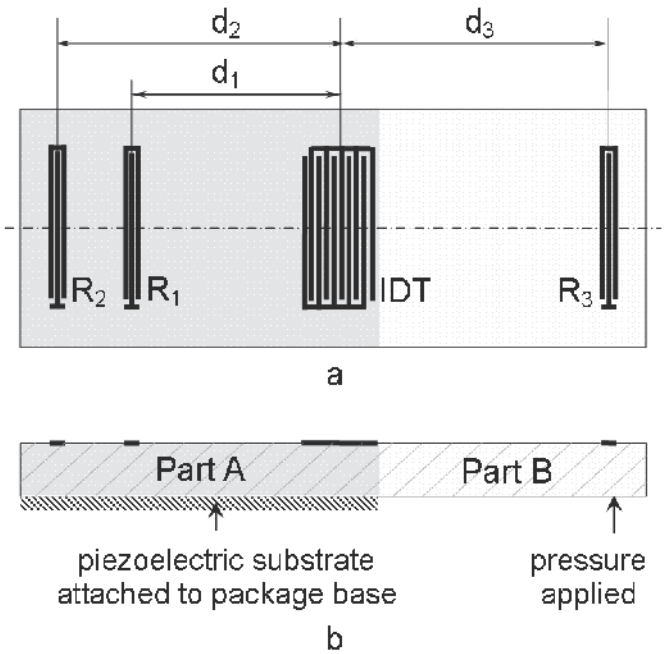


Fig. 11. (a) Vertical view, and (b) profile view of schematic diagram of the sensor structure

$$\Delta T = \frac{v}{2d_{12}^0 \alpha \omega_0} \varphi_{12} - \frac{1}{\alpha} \tag{4}$$

Here, d_{12}^0 is the initial value of d_{12} under initial temperature.
Since d_{13}, which is the difference between d_3 and d_1 in Part B, is affected by both ΔT and pressure, Equation (5) can be set up if the correlation of the effects of ΔT and pressure on τ_{13} is neglected (Li et al., 2008).

$$\tau_{13}(\varepsilon_P, \Delta T) = \tau_{13}^0 \left[1 + \varepsilon_P + \alpha \Delta T\right] \tag{5}$$

Here, ε_P is the change of d_{13} caused by the pressure, τ_{13}^0 the initial value of τ_{13} under initial temperature. Thus combining Equations (1), (2), (4), and (5), the phase shift being principally linear with applied pressure φ_P can be expressed as:

$$\varphi_P = \varphi_{13} - \left(d_{13}^0 / d_{12}^0\right) \varphi_{12} = \varphi_{13} - W \varphi_{12} \tag{6}$$

where d_{13}^0 is the initial value of d_{13} under initial temperature, W the weighted factor and equal to d_{13}^0 / d_{12}^0.

6.2 Device and tests for SAW microsensor

Y-Z cut $LiNbO_3$ was used as the substrate material of the sensor. The dimensions of the sensor die are 18 mm long, 2 mm wide, and 0.5 mm thick, respectively. The IDT and the three reflectors R_1, R_2, and R_3 were patterned onto the surface of the substrate using MEMS lift-off fabrication process. Fig. 12a is the schematic diagraph of a completely packaged

sensor. Fig. 12b is the photograph of a real microsensor without the packaging header cap, showing more structural details inside the sensor. The package, which includes a sensitive membrane and a header cap together with the package base attaching part of the substrate bottom, sealed the piezoelectric substrate in a vacuum cavity. The sensitive end of the piezoelectric cantilever contacts the membrane with negligibly small pre-force. The pressure difference between the cavity and the outside pressure can cause the deformation of the cantilever end along the vertical direction through the sensing membrane. The SAW signal frequency for this sensor is 433 MHz, corresponding to a wavelength of 8 μm. The IDT aperture is 50 times wave length, and d_1, d_2, and d_3 are 2400 μm, 4800 μm, and 7000 μm, respectively. Fig. 13 shows the different measured echo signals reflected from the correspondent reflectors of the sensor with an oscilloscope (DSA70604, Tektronix Co. Ltd., Pudong New Area, Shanghai, China). (Li et al., 2009)

The SAW microsensor with complete packaging was tested in a sealed chamber, inside which the air pressure and temperature are controllable. The pressure was measured with the pressure meter embedded in an electro-pneumatic regulator (ITV2030, 1 kPa resolution, SMC, 1 Claymore Drive #08-05/06 Orchard Towers, Singapore). A Pt100 thermal resistance connected with a digital meter (0.1 °C resolution) was used to measure the inside temperature of the chamber. The pulse signals for testing the sensor were generated and received by a vector signal generator SMJ100A and a spectrum analyzer FSP, respectively. Both were made by Rohde-Schwarz, Mühldorfstraße 15, München, Germany. The test temperature and pressure values were recorded by a time interval of 10 s.

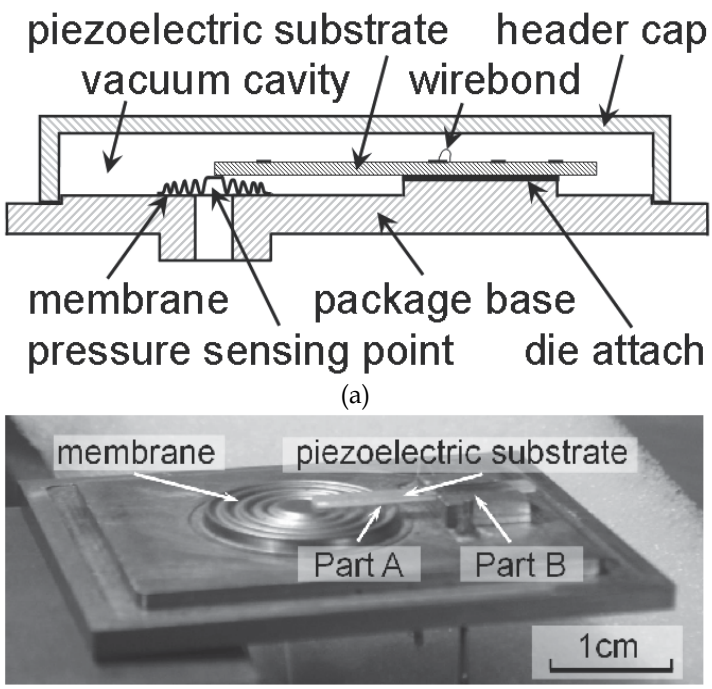


Fig. 12. (a) Schematic diagraph of a completely packaged sensor, and (b) photograph of a real microsensor without the packaging header cap

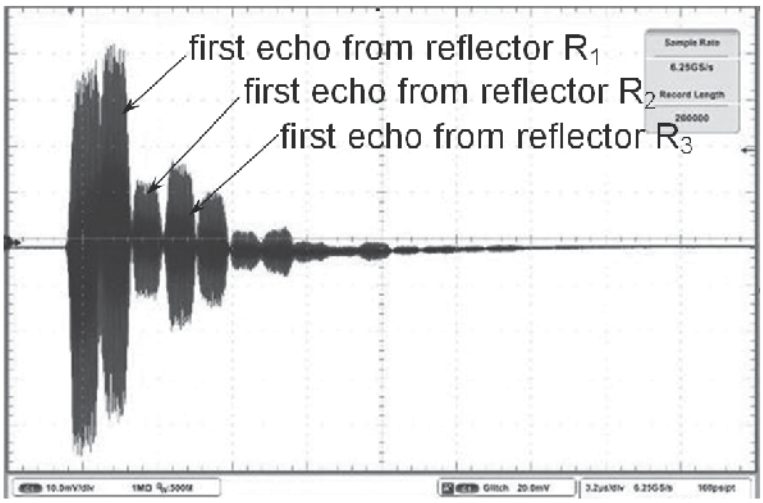


Fig. 13. Different measured echo signals from the reflectors

6.3 Results for SAW microsensor

Fig. 14a shows the measured data of phase differences φ_{12}, φ_{13} within the time range of 700 s, which are corresponding to the temperature and temperature effected pressure values, respectively. Using the measured φ_{12} by the SAW sensor and Equation (4), the calculated temperature values are compared with the direct measurement temperature data and shown in Fig. 14c. They match each other well although the calculated values have a higher temperature resolution than the direct measurement results, which was limited by the Pt 100 thermal resistance characteristics in the temperature range between 27.9 and 29.1 °C. The calculated pressure values eliminating the temperature variation effect using Equation (6) are shown in Fig. 14b, which agree the direct measured pressure data very well ranging from 0 to 150 kPa. (Li et al., 2009)

7. Conclusion

In this chapter, TPMS sensors are introduced, then a novel wireless passive SAW pressure and temperature microsensor with single sensing unit is reported. Its structural design, theoretical analysis, and test results are described. The calculated pressure and temperature values with this sensor measurement agree with the directly measured data very well.

8. References

Ballandras, S.; Lardat, R.; Penavaire, L. et al. (2006). Micro-machined, all quartz package, passive wireless SAW pressure and temperature sensor, IEEE Ultrasonics Symp., 1441-1444, 2006, Vancouver, Canada

Bao, X.; Burkhard, W.; Varadan, V. et al. (1987). SAW temperature sensor and remote reading system, *Proc. IEEE Ultrasonics Symp.*, 583-585, 00905607, Denver, USA

Buff,W. ; Klett, S.; Rusko,M. *et al*. (1998). Passive Remote Sensing for Temperature and Pressure Using SAW Resonator Devices, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol.45, No.5, 1388-1392, 08853010

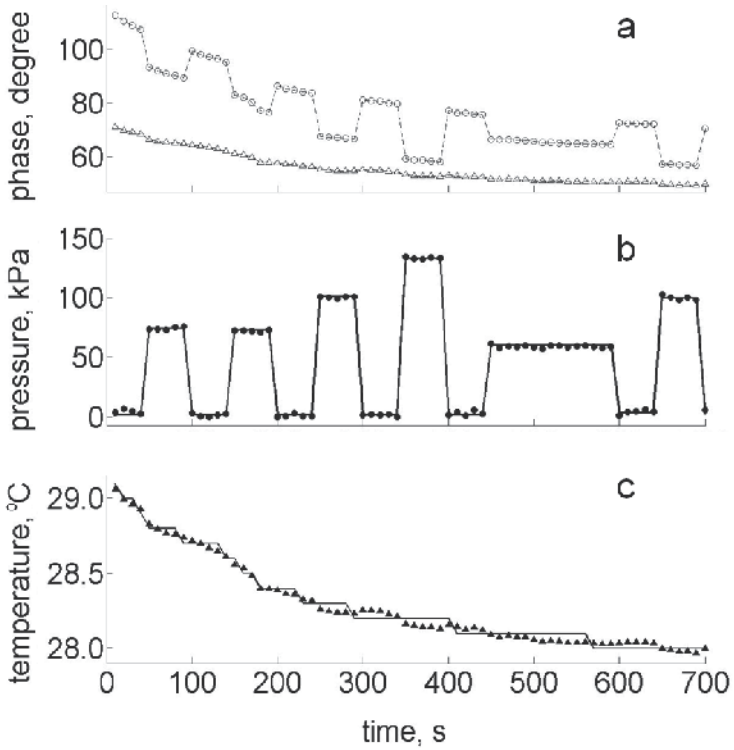


Fig. 14. (a) Measured phase differences with the SAW sensor
Hollow circle and dashed line phase difference data of φ_{13}
Hollow triangle and dashed line phase difference data of φ_{12}
(b) Comparison between calculated pressure from sensor measurement and direct measured pressure
Solid circle calculated pressure
Thick solid line direct measured pressure
(c) Comparison between calculated temperature from sensor measurement and direct measured temperature
Solid triangle calculated temperature
Thin solid line direct measured temperature

David, M. (2004). Safety Check: Wireless sensors eye tyre pressure, *EDN Europe*, No.9, 43-38

Kim, Y.; Chang, D. & Yoon, Y. (2004). Study on the optimization of a temperature sensor based on SAW delay line, *Korean Phys. Soc.*, Vol.45, No.5, 1366-1371, 0374-4884

Li, T.; Wu, Z.; Hu, H. *et al.* (2009). Pressure and temperature microsensor based on surface acoustic wave, *Electronics Letters*, Vol.45, No.6, 337-338, 0013-5194

Li T.; Zheng L.; Hu H. (2008). A novel wireless passive SAW sensor based on the delay line theory, *Proc. 3rd IEEE International Conf. Nano/Micro Engineered and Molecular Systems*, 440-443, 978-1-4244-1907-4, 2008, Sanya, China

Oh, J.; Choi, B.; Lee, S. (2008). SAW based passive sensor with passive signal conditioning using MEMS A/D converter, *Sensors and Actuators A*, Vol.141, No.2, 631-639, 0924-4247

Pohl, A.; Seifert, F. (1997). Wirelessly interrogable surface acoustic wave sensors for vehicular applications, *IEEE Transactions on Instrumentation and Measurement*, Vol.46, No.4, 1031-1037, 00189456

Reindl, L.; Ruppel, C. C.W.; Riek, K. *et al.* (1998) .A wireless AQP pressure sensor using chirped SAW delay line structures, *IEEE Ultrasonics Symposium*, Vol.1,355-358, 0780340957

Reindl, L.; Scholl, G.; Ostertag, T. *et al*. (1998). Theory and application of passive SAW radio transponders as sensors, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol.45, No.5, 1281-1292, 0885-3010

Schimetta, G. ; Dollinger, F.; Scholl,G. *et al*. (2000). Wireless pressure and temperature measurement using a SAW hybrid sensor, *IEEE Ultrasonics Symposium*, Vol.1, 445-448, 0780363655

Schimetta, G.; Dollinger, F.; Weigel, R. (2000). A wireless pressure measurement system using a SAW hybrid sensor, *IEEE Transactions on Microwave Theory and Techniques*, Vol.48, No.12, 2730-2735, 0018-9480

Scholl, G.; Schmidt,F.; Ostertag, T. *et al*. (1998). Wireless passive SAW sensor system for industrial and domestic applications, *l998 IEEE International Frequency Control Symposium*, Vol.1, 595-601, 0780343735

Steidl, R.; Pohl,A.; Reindl, L. *et al*. (1998). SAW delay lines for wirelessly requestable conventional sensors, *IEEE Ultrasonics Symposium*, No.1, 351-354, 10510117

Wang, F.; Wang, Z.; Shan, G., *et al*. (2003). Study Progress and Prospect of Smart Tire. *Tire Industry*, Vol.23, No.1, 10-15

16

Analysis and Modelling of Surface Acoustic Wave Chemical Vapour Sensors

Marija Hribšek and Dejan Tošić
Institute Goša Belgrade, University of Belgrade
Serbia

1. Introduction

Surface Acoustic Wave (SAW) sensors demonstrate superior selectivity for the detection of chemical agents. Due to their solid state design and fabrication, compatible with other modern technologies such as MIC (microwave integrated circuits), MEMS (micro-electro-mechanical-systems), CMOS, CCD (charge coupled devices) and integrated optic devices, SAW chemical sensors are extremely reliable. They have compact structure, high sensitivity, small size, outstanding stability, low cost, fast real-time response, passivity, and above all the ability to be incorporated in complex data processing systems. They can be used for *in situ* monitoring and sensing systems [Ho et al., 2003; Wohltjen & Dessy, 1979; Wohltjen, 1984; Comini, 2009] and for wireless sensing and monitoring in harsh environment [Pohl, 2000] including the detection of chemical warfare agents [Data Sheet, 2005] and land mine detection [Kannan et al., 2004]. It is interesting that a SAW-based sensor system is used as a volatile organic contamination monitoring system for the satellite and space vehicle assembly rooms in NASA. SAW sensors can distinguish organophosphates, chlorinated hydrocarbons, ketones, alcohol, aromatic hydrocarbons, saturated hydrocarbons, and water [Ho et al., 2003].

Surface acoustic waves were discovered in 1885 by Lord Rayleigh and are often named after him as Rayleigh waves [Rayleigh, 1885]. A surface acoustic wave is a type of mechanical wave motion which travels along the surface of a solid material, referred to as substrate. The amplitude of the wave decays exponentially with distance from the surface into the substrate, so that the most of the wave energy is confined to within one wavelength from the surface [Farnell, 1977; Martin et al., 1994].

A basic SAW device was originally developed in 1965 [White & Voltmer, 1965] by White and Voltmer when they found out how to launch a SAW in a piezoelectric substrate by an electrical signal. The basic SAW device consists of two interdigital transducers (IDTs) on a piezoelectric substrate such as quartz, Fig. 1.

Each IDT is a reversible transducer made of interleaved metal electrodes, which are used to convert an electrical signal to an acoustic wave and vice versa. An IDT is a bidirectional transducer: it radiates energy equally on both sides of the electrodes. Consequently, theoretical insertion loss introduced by an IDT is at least 6 dB. SAW devices work in the frequency range of 10 MHz to several GHz.

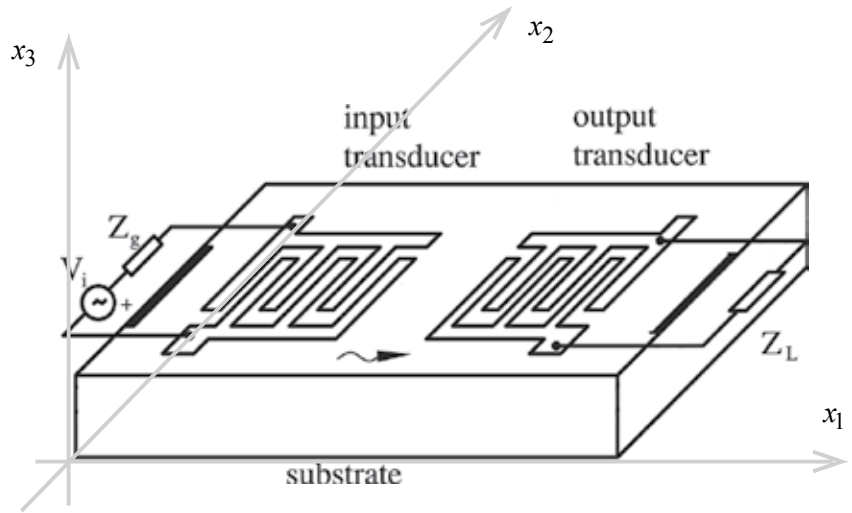


Fig. 1. The basic structure of a SAW device.

A sinusoidal voltage v of frequency f applied to the input IDT forms an electric field which through the piezoelectric effect causes a strain pattern of periodicity $2d$, where d denotes the distance between the centres of the electrodes. If the frequency f is such that $2d$ is close to the surface wave wavelength, a surface wave will be launched in two opposite directions away from the transducer. The surface wave causes the corresponding electric field in the output transducer and thus the voltage at the impedance Z_L. The magnitude of the output signal is the function of the ratio of the signal's wavelength and the distance $2d$. If the distance $2d$ is equal to the wavelength, the magnitude of the output voltage is maximal. The corresponding frequency is then called the centre or synchronous frequency of the device. The magnitude of the output voltage decays as the frequency shifts from the centre frequency. It means that a SAW device is a transversal bandpass filter with constant group delay. Therefore, it is usually called a SAW filter or delay line type of a SAW device. The phase characteristic is a function of the distances between the electrodes and the amplitude characteristic is a function of the electrodes' number and lengths. The width of the electrodes usually equals the width of the inter-electrode gaps giving the maximal conversion of electrical to mechanical signal and vice versa. The minimal electrode width obtained in industry is around 0.3 μm, which determines the highest frequency of around 3 GHz. The commonly used substrate crystals are: quartz, lithium niobate, lithium tantalate, zinc oxide and bismuth germanium oxide. They have different piezoelectric coupling coefficients and temperature sensitivities. The ST quartz is used for the most temperature stable devices. The wave velocity is a function of the substrate material and is in the range of 1500 m/s to 4800 m/s, which is 10^5 times lower than the electromagnetic wave velocity. This enables the construction of a small size delay line of a considerable delay.

In the second type of SAW devices, called SAW resonators, Fig. 2, IDTs are used only as converters of electrical to mechanical signals and vice versa, but the amplitude and phase characteristics are tailored using the reflections of the wave from either metal stripes or grooves of small depths.

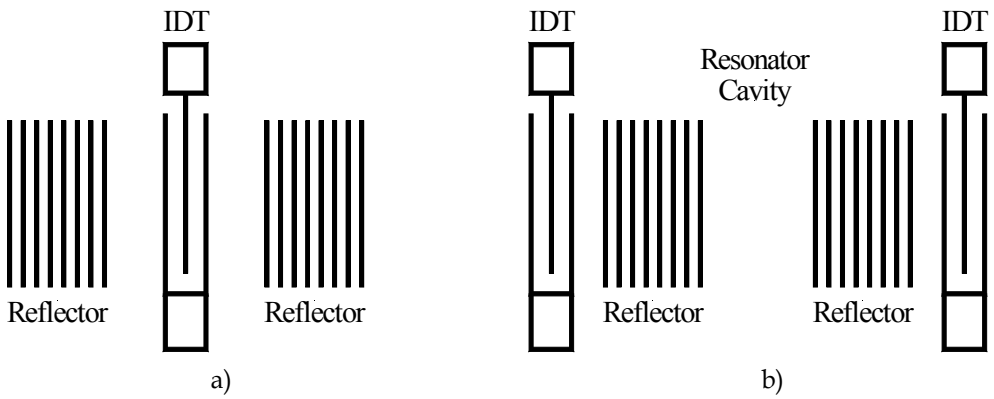


Fig. 2. a) One-port SAW resonator and b) two-port SAW resonator.

SAW resonators are made as one-port or two-port devices. In a one-port SAW resonator only one IDT, placed in the centre of the device, is used for both, input and output, transductions. The input electrical signal connected to IDT, via antenna or directly, forms an acoustical wave in the piezoelectric substrate which travels along the surface on both sides from the transducer. The wave reflects from the reflective array and travels back to the transducer, which transforms it back to the electrical signal. The attenuation of the signal is minimal if the frequency of the input signal matches the resonant frequency of the device. The resonant frequency is determined by the geometries of the transducer and reflectors, the distance between the transducer and the reflectors and the wave velocity. The wave velocity depends upon the substrate type, and the temperature. One-port resonators are used in oscillators. Two-port resonators are used as narrow bandpass filters. The dimensions of resonators are smaller than the delay line filters of the same centre frequency and bandwidth.

Beginning from around year 1970, versatile SAW devices were developed for pulse compression radars, band pass filters for the TV receivers (sets), and radio systems. The rise of mobile radio of the eighties, and particularly cellular telephones, caused an increase in the demand for SAW filters, so that they are now produced in vast number.

In the last three decades SAW devices of both types have found applications as identification tags, sensors of different physical quantities, chemical sensors, and biosensors [Pohl, 2000; Seifert at al., 1994; Hribšek et al., 2009; Hribšek et al., 2010; Mitsakakis et al., 2009]. They are used in consumer and highly professional devices and systems. SAW sensors are passive elements (they do not need power supply). The main advantage of all SAW sensors is their ability to be accessed wirelessly enabling remote monitoring in harsh environment. Wireless access is achieved simply by connecting an antenna to the input transducer.

The operation of delay line SAW sensors is based on the fact that the measurand (temperature, pressure, strain, chemical vapour etc.) affects the propagation of the SAW in the sensor in attenuation and delay. If the sensor is heated, stretched or compressed, or if it is mass loaded, the substrate's length and its elasticity constants are changed. These changes cause velocity and phase delay variations, which then proportionally change the centre frequency, attenuation and time delay of the device. The first reported use of SAW technology for a sensor application was in 1975 for pressure sensing [Cullen & Reeder, 1975; Cullen & Montress, 1980]. SAW temperature sensors have millidegree resolution, good

linearity, fast response, and low hysteresis [Pohl, 2000]. They are sealed in hermetic packages. The response time is about 0.3 s, 1000 times faster than in bulk acoustic wave (BAW) sensors. For temperatures up to 200°C lithium niobate is the ideal material for temperature sensors, because of its large temperature coefficient (TCD) of approximately 90 ppm/°C and its high electro-acoustic coupling constant. For temperatures up to 1000°C langasit substrate is used.

SAW chemical vapour sensors were invented by Wohltjen [Wohltjen & Dessy, 1979; Wohltjen, 1982]. A SAW chemical vapour sensor is made from a SAW device by placing chemically sensitive coatings (usually polymer films) on the device surface. The absorbed chemical vapours into the coating cause a change in the centre or resonant frequency of the sensor. A microcomputer can measure these changes and use them to determine the presence and concentration of chemical agents.

The SAW sensor coatings have unique physical properties which allow a reversible adsorption of chemical agents. In order to make the whole system as compact as possible, the SAW device should be incorporated in CMOS or MEMS integrated circuits [Zaki et al., 2006]. In that case piezoelectric material is placed on the top of the IC circuit, e.g. on the top of silicon or the isolating layer, usually silicon dioxide. Commonly used piezoelectric materials in classical SAW applications are ST-cut quartz and lithium niobate. Besides them ZnO, AlGaN, GaN, AlN are used [Zaki et al., 2006; Rufer et al., 2006; Assouar et al., 2000; Rufer et al., 2005; Kirsch et al., 2006; Kirsch et al., 2007; Omori et al., 2008]. Recently, multilayered substrates are used for the wave velocity increase [Ahmadi et al., 2004]. The highest velocities are achieved when the piezoelectric material is placed on the top of the diamond layer, due to its highest acoustic wave velocity [Assouar et al., 2000; Benetti et al., 2004; Benetti et al., 2005; Besmaine et al., 2008; Hakiki et al., 2005; Mortet et al., 2008; Jian et al., 2008; Shikata et al., 2005]. Several piezoelectric materials in combination with diamond/silicon substrates have been investigated theoretically and experimentally.

Theoretical calculation of the wave velocity in the multilayer structures is based on the solution of the wave equation demanding elaborate numerical computations. The use of diamond in the multilayered SAW structure has the following advantages: high frequencies up to 5 GHz, high coupling coefficients up to 1.2%, small temperature deviations, high power capability, and small device size without submicron lithography. The disadvantages of the layered SAW structures are the complex design and the problem related to the deposition of a piezoelectric layer with appropriate crystalline orientation. These facts probably have caused insufficient research on SAW sensors using diamond. Extreme chemical stability and bio-inertness [Specht et al., 2004] make diamond ideal material for sensors operating in harsh or biologic environments.

This chapter describes principles of operation, analyses and modelling of delay line chemical vapour SAW sensors.

2. Principles of chemical vapour SAW sensor operation

The basic principle of chemical vapour SAW sensors is the reversible sorption of chemical vapours by a coating which is sensitive to the vapour to be detected. A transversal, or delay line, SAW chemical sensor can be schematically presented as in Fig. 3. It consists of two IDTs and a chemically sensitive thin layer placed between them on the top surface of the piezoelectric substrate.

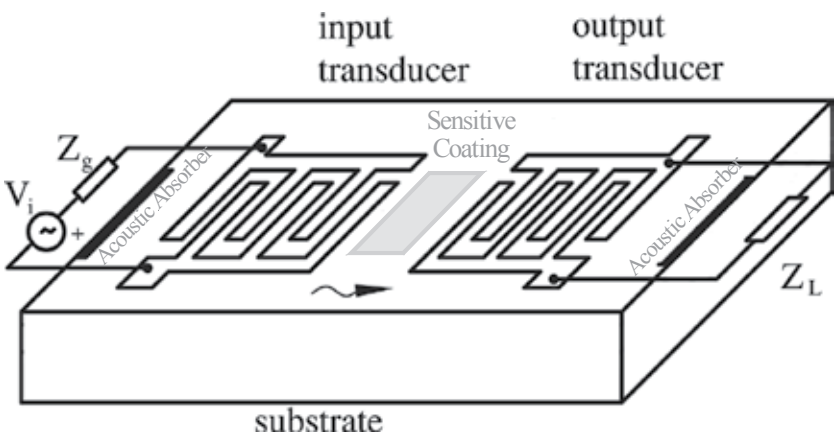


Fig. 3. The basic configuration of a chemical SAW sensor.

The surface wave is induced by an electrical signal applied to the input IDT. The output signal (voltage) is taken from the output IDT. The velocity and attenuation of the wave are sensitive to mass and viscosity of the thin layer, usually polymer film. The purpose of the thin layer is to absorb chemicals of interest. When the chemical is absorbed, the mass of the polymer increases causing a change in velocity and phase of the acoustic signal, which causes a change in amplitude and frequency of the output voltage at the load impedance Z_L. Acoustic absorbers placed on the substrate edges damp unwanted SAW energy and eliminate spurious reflections that could cause signal distortions.

The IDTs are identical with uniformly spaced electrodes of equal lengths and equal ratio of electrodes width and spacing. The number of electrodes defines the frequency bandwidth of a SAW device. The electrodes' lengths and their number, and matching networks at the electrical ports, should be chosen to match the IDT input resistance, at the centre frequency f_0, to the load resistance R_L and the generator resistance R_g. In that case, the overall minimal loss due to IDTs is 12 dB. The wavelength corresponding to the centre frequency equals $2d$ (the distance between the electrodes of the same polarity). The centre frequency and the bandwidth are determined by the IDT`s geometry and the substrate type.

The middle part of a SAW sensor, a delay line, is generally treated as lossless. However, its losses can be neglected only for lower frequencies and small delays (small distances between the transducers). The transfer function of the delay line is normally assumed unity, although this may not be true for high frequencies ($f > 0.5\,\mathrm{GHz}$) or if there are films in the propagation path [Golio, 2008]. In communications, in electrical filtering applications, the distance between the IDTs is small. Quite opposite, in chemical sensors this part is essential and must have a certain length, usually 100–200 wavelengths [Martin et al., 1994], which should be taken into account.

The frequency and the magnitude of the output voltage across the load are proportional to the mass loading of the sensing part. The output voltage in the presence of sensing material (without vapour) serves as a reference. The difference of the output voltage in the presence of vapour and the reference is proportional to the vapour concentration. Sometimes the output voltage is directly measured, but usually a SAW delay line is placed in the feedback loop of the oscillator, Fig. 4, so that the oscillation frequency is proportional to the measurand and it can be easily measured.

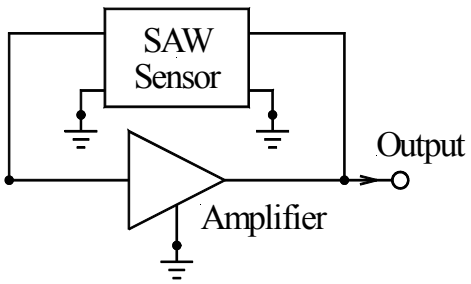


Fig. 4. SAW delay line oscillator.

Time delay τ of the SAW delay line sensor is the ratio of acoustical length L (distance between the first electrodes of the input and output transducers) and SAW velocity v. Generally, in the known sensor applications, L and v are changed due to a temperature change, mechanical stress, and strain, or because of a mass loading from a thin surface layer. In chemical vapour sensors L does not change so that the relative change of the delay due to the variation of a loading from a thin sensitive layer, with or without the vapour, denoted as measurand y, can be expressed as follows:

$$\frac{\mathrm{d}\tau}{\tau} = -\frac{\mathrm{d}v}{v\mathrm{d}y}\mathrm{d}y = \gamma_y \mathrm{d}y \tag{1}$$

where γ_y can be called the delay sensitivity with respect to y. It is determined by the orientation and type of crystalline material used to fabricate the sensor [Pohl, 2000; Živković, 2003].

The oscillations are sustained if the following conditions are met:

- amplification in the open loop is greater than 1,
- net phase in the closed loop, acoustical plus electrical, equals $2n\pi$, where n is the number of the mode, e.g.,:

$$\frac{2\pi f L}{v} + \phi_{\mathrm{A}}(f) = 2\pi n \tag{2}$$

where f is the oscillation frequency, and ϕ_{A} is the phase of the amplifier. Since the electrical delay is much smaller than acoustical, from (1) and (2) can be found:

$$\frac{\mathrm{d}f}{f} = -\frac{\mathrm{d}\tau}{\tau} = -\gamma_y \mathrm{d}y \tag{3}$$

which gives the straight influence of the measurand on the frequency. To avoid temperature influence on measuring results sensors should be made on ST-cut quartz.

In some applications the sensor is a part of a more sophisticated system. In that case two equal SAW sensors are used: one is vapour-free and serves as reference, the other one is exposed to vapour and actually performs the sensing function, Fig. 5. The two SAW sensors are embedded into electrical oscillator circuits and the frequency shift between the oscillators is proportional to the gas concentration. Using an electronic circuit called the mixer the voltage proportional to the vapour concentration is obtained from the frequency shift.

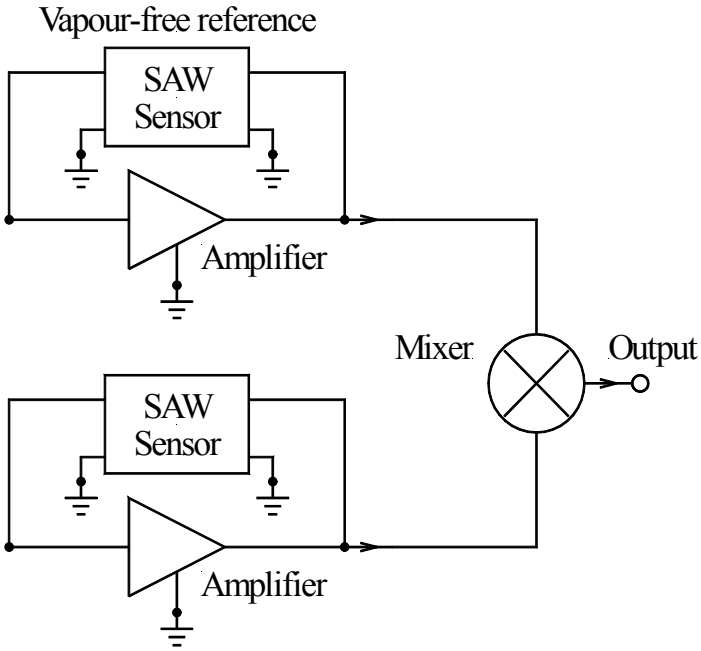


Fig. 5. Block diagram of a differential chemical SAW sensor system.

3. Analyses and modelling of delay line SAW sensors

According to equation (3) the goal of the analysis of SAW chemical sensors is the derivation of formulas which connect the frequency shift and chemical quantities (e.g., vapour concentration). The existing analysis approaches are usually: (a) the analysis based on the wave equation solutions [Wohltjen & Dessy, 1979; Martin et al., 1994], (b) the analysis based on published formulas derived from the wave equation [Grate & Klusty, 1991; Grate & Zellers, 2000], and (c) approximate analysis by means of equivalent electro-mechanical circuits [Živković et al., 2009]. In the first two approaches chemical SAW sensors have been analyzed mainly from the chemical point of view without giving any relations between chemical quantities and the geometry of the SAW sensor and matching conditions at its electrical ports. Only in the last approach the derived formulas give the straightforward connection between the chemical vapour concentration and the geometry of the SAW sensor, sensing part, and substrate properties. The most complete analyses based on the wave equation are reported in [Martin et al., 1994].

In all of these approaches the centre frequency f_0 without sensing film should be known. The centre frequency is proportional to the wave velocity v and inversely proportional to the wavelength λ_0 (which is equal to the period of IDTs $\lambda_0 = 2d$):

$$f_0 = \frac{v}{\lambda_0} \tag{4}$$

If the substrate is single layer piezoelectric crystal the centre frequency can be easily calculated from the IDT geometry and velocity of the material (can be found in the open

literature). However, if the substrate is multilayered, as is the case when the sensor is imbedded in MEMS or integrated circuits or when diamond is used, theoretical velocity determination is rather tedious task and numerical calculation methods are to be performed in each case separately [Mortet et al., 2008; Benetti et al., 2004; Adler et al., 1990].

3.1 Analyses based on wave equation

It is well known that the exact analysis using surface wave theory is very complex even in the case of a free surface and single layer piezoelectric substrate [Farnell, 1977; Martin et al., 1994; Golio, 2008; Ballantine et al., 1997]. It can be found in classical SAW books [Farnell, 1977; Farnell, 1978; Feldman & Henaff, 1989] for single piezoelectric substrate. It starts from the second Newton's law applied to particle motion, which gives a set of partial differential equations. The equations are solved for the appropriate boundary conditions and relations between mechanical and electrical quantities of a piezoelectric substrate. The Maxwell's equations for the electromagnetic field should be taken into account, as well.

Velocity determination is performed by solving wave equations. The wave equations are the partial differential equations of the form [Farnell, 1977; Farnell, 1978; Hribšek, 1982; Hribšek, 1986]:

$$\rho \frac{\partial^2 u_m}{\partial t^2} = \sum_{j=1}^{3} \frac{\partial T_{mj}}{\partial x_j}, \ (m = 1,2,3) \tag{5}$$

where

ρ is the density of the substrate,

u_m is displacement in the direction x_m,

T_{mj} is the component of the stress tensor,

x_j's are space coordinates,

t is time, and

$T_{mn} = T_{nm}$.

Surface wave is propagating in x_1 direction, and x_3 direction is normal to the surface, Fig. 1.

For the piezoelectric substrate the relations between mechanical and electrical variables can be expressed as follows:

$$T_{mn} = c_{mnpq} S_{pq} - e_{pmn} E_p, \ (m,n,p,q = 1,2,3) \tag{6}$$

$$D_m = e_{mnp} S_{np} + \varepsilon_{mn} E_n \tag{7}$$

where

e_{pmn} are the piezoelectric constants of the material (elements of the piezoelectric tensor),

c_{mnpq} are elastic constants of the material measured at the constant electric field,

ε_{mp} are dielectric constants measured at the constant mechanical conditions,

E_p is electric field,

D_m is electric displacement, and

S_{pq} is relative mechanical displacement defined as

$$S_{pq} = S_{qp} = \frac{1}{2}\left(\frac{\partial u_p}{\partial x_q} + \frac{\partial u_q}{\partial x_p}\right). \tag{8}$$

Besides that, Maxwell's equations for electromagnetic field are taken into account. Based on the fact that the electromagnetic field is slowly varying, it can be assumed to be static, e.g., that the electric field is the gradient of a scalar, but variable, potential φ:

$$E_m = -\frac{\partial \varphi}{\partial x_m}. \tag{9}$$

The lack of free charges in the substrate yields

$$\frac{\partial D_m}{\partial x_m} = 0. \tag{10}$$

Using (6)–(10), equations (5) and (10) can be transformed to a more convenient form

$$\rho \frac{\partial^2 u_m}{\partial t^2} - c_{mnpq} \frac{\partial^2 u_p}{\partial x_m \partial x_q} = e_{pmn} \frac{\partial^2 \varphi}{\partial x_p \partial x_m}, \tag{11}$$

$$\varepsilon_{mp} \frac{\partial^2 \varphi}{\partial x_m \partial x_p} = e_{mpq} \frac{\partial^2 u_p}{\partial x_m \partial x_q}. \tag{12}$$

Last equation represents the Laplace equation for anisotropic piezoelectric materials. The equations are solved for the appropriate mechanical and electrical boundary conditions at the surface $x_3 = 0$. For the unloaded surface, the Laplace equation holds for the potential above the surface and

$$T_{3n}\big|_{x_3=0} = 0, \ (n = 1,2,3) \tag{13}$$

Potential and electric displacement D_3 also satisfy the continuity equation for $x_3 = 0$, while for $x_3 = \infty$ vanish. Consequently, potential φ has the form

$$\varphi = \varphi(0)\mathrm{e}^{-kx_3}\mathrm{e}^{\mathrm{j}k(x_1 - vt)} \tag{14}$$

where $k = \omega/v$ is the wave number and v is the wave velocity.
Equations (11) and (12) can be solved only numerically using various methods. One method is to represent the solution as a sum of partial solutions given by [Farnell, 1977; Farnell, 1978]

$$u_m^{(i)} = \alpha_m^{(i)} \mathrm{e}^{\mathrm{j}kb^{(i)}x_3}\mathrm{e}^{\mathrm{j}k(x_1 - vt)}, \ \varphi^{(i)} = u_4^{(i)}, \ (m = 1,2,3,4). \tag{15}$$

Each partial solution must satisfy equations (11)–(12) and equals zero for $x_3 = \infty$. By substituting partial solutions in (11) and (12), a set of four linear algebraic equations is formed in which the coefficients are the functions of density and elastic, dielectric and piezoelectric constants of the substrate. In order to get non trivial solutions, the determinant

of the system must be zero, which gives an algebraic equation of the eighth order (degree) in $b^{(i)}$. Since the wave amplitude decays with substrate's depth, only four solutions within lower halfplane of the complex variable b are of interest. Consequently, the solution has the form

$$u_m = \left(\sum_{i=1}^{4} C_i \alpha_m^{(i)} \mathrm{e}^{\mathrm{j}kb^{(i)}x_3} \right) \mathrm{e}^{\mathrm{j}k(x_1 - vt)}, \ \varphi = u_4, \ (m = 1,2,3,4) \ . \tag{16}$$

Coefficients k, $b^{(i)}$ and $\alpha^{(i)}$ are the functions of velocity v. Weighting coefficients C_i are found to satisfy boundary conditions at $x_3 = 0$. From these conditions a set of four equations is obtained. Velocity v is calculated setting the determinant of the system to zero. Even in the case of substrate crystals with symmetry, where many of the coefficients c, e and ε are equal to zero, the explicit solution for velocity cannot be found. The solution has to be found numerically, using iterative procedures. When the velocity is calculated, the coefficients C_i can be found, e.g., solutions for the particle motion (displacements) and potential. The procedure can be used for any substrate, but the calculation time depends on the substrate type. From (16) is obvious that all variables are independent of x_2 and u_2. The wave amplitude decays exponentially with the distance from the surface, so that the most of the wave energy is confined within the depth of one wave length. Therefore, the motion in x_1 direction can be generally represented as

$$u_1 = u_1(x_3)\mathrm{e}^{\mathrm{j}\omega t - \gamma x_1}, \ \gamma = \alpha + \mathrm{j}k = \alpha + \mathrm{j}\frac{\omega}{v} \tag{17}$$

where

γ is complex propagation factor and

α is attenuation [Ballantine et al., 1997].

For multi layer substrates calculations are even more difficult. In that case, equations (11) and (12), assuming $u_2 = 0$ and that all variables are independent of x_2, expand to three two-dimensional partial differential equations with three unknowns: particle displacement components u_1 and u_3 and potential component ϕ. To find the velocity, these three equations along with two equations for normal stress and one equation for normal electrical displacement are solved in each layer with appropriate boundary conditions at the top and bottom surface and across the interfaces. To find the velocity, matrix techniques can be used [Ahmadi et al., 2004; Adler et al., 1990].

The finite element method (FME) can be also used for multilayer substrates analyses [Sankaranarayanan et al., 2005]. It is employed to calculate the effective phase velocity in multilayer structures with diamond [Hashimoto, 2000; Plessky & Koskela, 2000; Sung et al., 2009].

At McGill University PC Acoustic Wave Software was developed for the velocities calculation in multilayer substrates [Adler et al., 1990].

Derivation of the frequency shift due to the sensing film and chemical quantities can be found from (4):

$$\frac{\Delta f_0}{f_0} = \frac{\Delta v}{v} \ . \tag{18}$$

The velocity shift is not solely determined by the material constants but also by the ratio between the thickness of the piezoelectric layer and the wavelength corresponding to the centre frequency.
Applying perturbation method to the solutions of the wave equation, Tiersten and Sinha derived a formula relating velocity change to film properties for the case of an acoustically thin, elastic film [Tiersten & Sinha, 1978]. Wohltjen first applied the Tiersten formula to analyze the response of polymer-coated SAW sensors [Wohltjen, 1984].
The simplest (and the one most utilized) interaction for SAW sensor applications is the response due to changes in the mass density on the device surface. For that case, in [Ballantine et al., 1997] the relation between the changes in wave velocity, changes in wave energy density, and fractional change in mass density of the lossless medium is derived:

$$\frac{\Delta v}{v_0} = -\frac{\Delta U}{U_0} = -\frac{\Delta \rho}{\rho_0} \tag{19}$$

where v_0, U_0 and ρ_0 denote unperturbed propagation velocity, energy density and density, respectively. Using expression (19), solutions for velocities from the wave equation and grouping together all the substrate-dependent constants, result in the expression for the mass-induced change of the thin film in SAW propagation velocity in the form

$$\frac{\Delta v}{v_0} = -c_m f_0 \rho_s \tag{20}$$

where c_m is the mass sensitivity factor, and ρ_s is the density of the mass load. Coefficient c_m for quartz, lithium niobate and gallium arsenide can be calculated from the data given in [Table 3.1, Ballantine et al., 1997].
In [Martin et al., 1994] a perturbation method is also used to find the changes in the complex propagation factor (velocity and attenuation) contributed by acoustically thin and thick visco-elastic polymer films. In acoustically thin films, displacement is uniform across the film and varies only in the direction of propagation. For thick films, inertial effects cause a phase lag across the film for shear displacements. To obtain velocity changes linearly proportional to absorbed vapour concentration, it is necessary that the film remains in the acoustically thin regime [Martin et al., 1994]. The regime of the film operation can be determined from the ratio *R* of cross-film to in-plane gradients induced by the SAW [Martin et al., 1994]:

$$R = \frac{A f v_0 \rho h}{|G|} \tag{21}$$

where
ρ, *h* and *G* are the film density, thickness, and shear modulus, respectively;
A is a substrate-dependent parameter (having a value of 1.9 for ST-cut quartz) [Martin et al., 1994].
When the film coating is sufficiently thin (small *h*) and rigid (large *G*) such that $R << 1$, the film is *acoustically thin*. If the film properties are such that $R >> 1$, the film is *acoustically thick*.
When the films are elastic the intrinsic elastic moduli are real, resulting in zero attenuation changes, and the Tiersten formula [Tiersten & Sinha, 1978] for fractional velocity change, written in terms of the Lamé constants (λ, μ) [Martin et al., 1994; Ballantine et al., 1997]:

$$\frac{\Delta v}{v_0} = -\omega h\left(c_1\left(\rho - \frac{\mu}{v_0^2} \right) + c_2\rho + c_3\left(\rho - \frac{4\mu}{v_0^2}\frac{\lambda+\mu}{\lambda+2\mu} \right) \right) \tag{22}$$

Where ω is the angular frequency and c_i are the elastic constants of the material.
Wohltjen described the frequency shift Δf_s due to a thin, non-conducting film as [Wohltjen, 1984]

$$\Delta f_s = (k_1 + k_2)F^2 h\rho - k_2 F^2 h(4\mu / V_R^2)[(\lambda+\mu)/\lambda+2\mu)] \tag{23}$$

where
F is the centre frequency of a SAW device,
V_R is the wave velocity in the substrate, and
k_1, k_2 are material constants of the substrate. If only mass loading is taken into account, frequency shift is calculated using the first term in (23).
In [Grate et al., 1988; Grate et al., 1992] frequency shift Δf_v due the vapour sorbed in the film is given by

$$\Delta f_v = \frac{\Delta f_s C_v K}{\rho} \tag{24}$$

where
C_v is the concentration of the vapour in the vapour phase, and
K is the partition coefficient, which is the ratio of the concentration of the vapour in the sorbent phase, C_s, to the concentration of the vapour in the gas phase, C_v.
In [Grate & Zellers, 2000] equation (24) is modified to include the contribution of swelling-induced modulus changes to vapour sensor responses:

$$\Delta f_v = \frac{\Delta f_s C_v K}{\rho_s} + f_L \frac{C_v K}{\rho_L}\frac{\Delta f_s A_{SAW}}{\alpha} \tag{24a}$$

$$\Delta f_v = \frac{\Delta f_s C_v K}{\rho_s} + f_1 v_1 \frac{\Delta f_s A_{SAW}}{\alpha} \tag{24b}$$

3.2 Analysis based on published formulas

Typically, the published formulas which connect frequency shifts and chemical compounds quantities are applied formally, without any insight into the influence of many properties of a real SAW delay line (geometry, propagation losses, technological constraints, and production tolerances) on the frequency change. This is the reason why some researchers perform more experiments than needed, or have difficulties in explaining discrepancies between the expected and measured values [Joo et al., 2005]. The mostly used formula for the frequency shift is actually equation (20) expressed in a slightly different form. In [Balcerzak & Zhavnerko, 2007] it has the form (neglecting the changes of viscoelasticity, dielectric constant and electric conductivity of the layer)

$$\Delta f = K_s f_0^2 \Delta m \frac{1}{A} \tag{25}$$

where

K_s is a constant (for lithium niobate $K_s = 5.49 \cdot 10^{-11}\,\mathrm{s\,m^2 g^{-1}}$),

f_0 is operating frequency of the sensor,

A is the surface of chemosensitive layer, and

Δm is the change of mass bonded to the sensor.

In [Benetti et al., 2004] it is in the form

$$\Delta f = -c_m f_0^2 h' \Delta \rho_s \tag{26}$$

where

c_m is the mass sensitivity coefficient (frequency independent),

h' is the thickness of the part of the coating that incorporates gas molecules, and

$\Delta \rho_s$ is the mass density change due to absorption.

In [Ho et al., 2003] modifications of (24) are used.

3.3 Analyses based on the electro-mechanical equivalent circuit

The analysis of transversal chemical SAW sensors, based on the electro-mechanical equivalent circuit, develops in a straight forward manner explicit general relations between electrical signals, voltages or frequencies, and vapour detection estimations taking into account properties of real SAW devices [Živković et al., 2009]. The whole sensor is modelled as a two-port network consisting of three parts: (1) the input interdigital transducer, (2) the delay line that is the sensing part, and (3) the output interdigital transducer, Fig. 6. The transducers are modelled as three-port networks and the delay line as a two-port network. The closed form formula for vapour concentration estimation is derived using analogy between electrical and mechanical quantities, the properties of the surface wave and the technological process and implementation of the sensor.

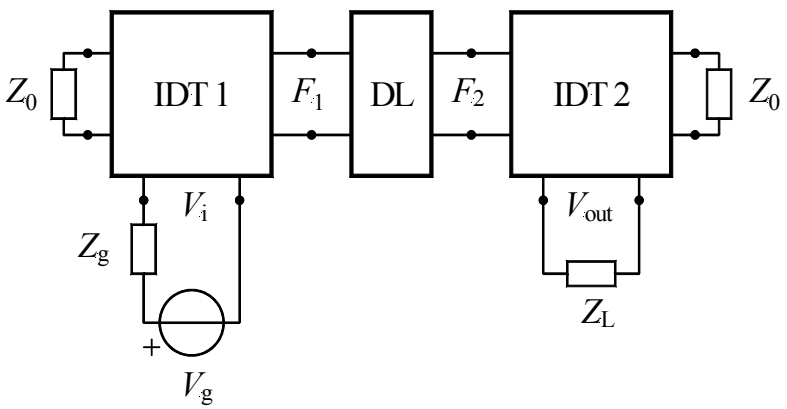


Fig. 6. The equivalent circuit of a SAW sensor.

The characteristic SAW acoustic impedance of the unloaded substrate is designated by Z_0 and the acoustic impedance due to the mass loading of the thin film is Z_m :

$$Z_0 = A \rho_s v \tag{27}$$

$$Z_{\mathrm{m}} = A_{\mathrm{m}} \rho_{\mathrm{m}} v \tag{28}$$

where

A is the substrate cross-section area through which the waves propagate,

ρ_{s} is the mass density of the piezoelectric substrate,

v is the SAW velocity in the piezoelectric substrate,

A_{m} is the cross-section area of the thin film, and

ρ_{m} is the mass density of the film.

$Z_{\mathrm{g}} = R_{\mathrm{g}}$ and $Z_{\mathrm{L}} = R_{\mathrm{L}}$ are purely resistive electrical impedances of the generator and the electrical load, respectively.

The relative variation of the centre frequency due to the mass loading is equal to the relative variation of the wave velocity, equation (18), and can be found using the equivalent circuit of a mass loaded delay line, Fig. 7.

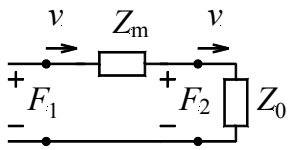


Fig. 7. The equivalent circuit of a mass loaded delay line.

By analogy between electrical and mechanical quantities, the relative variation of frequency and the relative variation of velocity, for Z_{m} much smaller than Z_0, are

$$\frac{\Delta f}{f_0} = \frac{\Delta v}{v} = \frac{-Z_{\mathrm{m}}}{Z_0} = \frac{-\rho_{\mathrm{m}} h_{\mathrm{m}}}{\rho_{\mathrm{s}} \lambda_0} K_{\mathrm{w}} \tag{29}$$

where

ρ_{m} and h_{m} are the density and thickness of the thin layer, respectively,

ρ_{s} is the density of the piezoelectric substrate, and

K_{w} is a coefficient that depends on the technological process and implementation of the sensor.

The components of the wave decay exponentially inside the substrate and the penetration is of the order of one wavelength. Therefore, in (29), instead of the substrate thickness, one wavelength λ_0 is used. From the last equation Δf can be determined as

$$\Delta f = -\frac{\rho_{\mathrm{m}} h_{\mathrm{m}}}{\rho_{\mathrm{s}} v} f_0^2 K_{\mathrm{w}} \tag{30}$$

The last equation shows that the higher sensitivity will be obtained if the centre frequency is higher, thickness and density of the film larger, and the substrate density and velocity smaller. This means that quartz ($\rho_{\mathrm{s}} = 2.62\ \mathrm{g/cm^3}$) is a better choice for the substrate than lithiumniobate ($\rho_{\mathrm{s}} = 4.7\ \mathrm{g/cm^3}$). Furthermore, if ST-cut quartz is used temperature dependence can be neglected. Using the last equation the frequency shift, Δf_{p}, due to the sensing film (without vapour) can be determined:

$$-\frac{\Delta f_{\mathrm{p}}}{f_0} = \frac{\rho_{\mathrm{p}} h_{\mathrm{p}}}{\rho_{\mathrm{s}} v} f_0 K_{\mathrm{w}} \tag{31}$$

where ρ_p and h_p are the density and thickness of the film, respectively.
Using the same reasoning, the fact that h_p is much smaller than λ_0, and the partition coefficient K, the frequency shift Δf_{vap} due to the concentration of the vapour in the vapour phase C_v (concentration in the ambient), can be calculated as:

$$\Delta f_{vap} = KC_v \frac{\Delta f_p}{\rho_p} \tag{32}$$

4. Conclusion

Analysis of chemical SAW sensors can be approached to in three ways: (1) exact analysis by solving the wave equation, (2) published formulas which connect frequency shifts and chemical compounds quantities, and (3) approximate analysis by means of equivalent electro-mechanical circuits.
The exact analysis of SAW sensors using surface wave theory is very complex even in the case of a free surface of a single layer piezoelectric substrate.
The published formulas which connect frequency shifts and chemical compounds quantities are applied formally, without any insight into the influence of many properties of a real SAW delay line (geometry, propagation losses, technological constraints, and production tolerances) on the frequency change. This is the reason why some researchers perform more experiments than needed, or have difficulties in explaining discrepancies between the expected and measured values.
The analysis based on electro-mechanical equivalent circuits of SAW sensors connects, in a straight forward manner, electrical signals and chemical vapour concentrations, taking into account important properties of real SAW devices, such as propagation losses, technological constraints, and production tolerances. The unique feature of this approach is a set of closed form analytic expressions for vapour concentration estimations. The expressions explicitly relate the vapour concentration, substrate parameters, and centre frequency. They enable insight into the influence of the sensor design parameters on the sensor performance and predict very efficiently and correctly the frequency and voltage shifts due to the vapour concentrations. The closed-form expressions can be used for the design of optimal sensors for a given vapour.

5. References

Adler, E. L.; Slaboszewicz, J. K.; Farnell, G. W. & Jen, C. K. (1990). PC software for SAW propagation in anisotropic multilayers, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. 37, No. 3, May 1990, pp. 215-223, ISSN 0885-3010.

Ahmadi, S.; Hassani, F.; Korman, C.; Rahaman, M. & Zahloul M. (2004). Characterization of Multy- and Single- layer Structure SAW Sensor, *Proceedings of the IEEE Conference on Sensors*, pp. 1129–1132, Vienna, Austria, Oct. 24–27, 2004.

Assouar, M. B.; Elmazria, O.; Jimenez, R.; Sarry, F. & Alnot, P. (2000). Modelling of SAW filter based on ZnO/diamond/Si layered structure including velocity dispersion, *Applied Surface Science*, Vol. 164, No. 1-4, Sept. 2000, pp. 200–204, ISSN 0169-4332.

Ballantine, D. S.; White, R. M.; Martin, S. J.; Ricco, A. J.; Zellers, E. T.; Frye, G. C.; Wohltjen, H. (1997). *Acoustic Wave Sensors: Theory, Design, Physico-Chemical Applications*, Academic Press, ISBN 0-12-077460-7, San Diego, USA.

Balcerzak, A. & Zhavnerko, G. (2007). Ultrasonic Chemical Sensor For Detection of Vapors of Some Volatile Organic Compounds, *Molecular and Quantum Acoustics*, Vol. 28, 2007, pp. 7-12.

Benetti, M.; Cannata, D.; D'Amico, A.; Di Pietrantonio, F.; Macagnano, A. & Verona, E. (2004). SAW Sensors on AlN/Diammond/Si Structures, *Proceedings of the IEEE Conference on Sensors*, pp. 753–756, Vienna, Austria, Oct. 24–27, 2004.

Benetti, M.; Cannatta, D.; Pietrantonio, F. D. & Verona, E. (2005). Growth of AlN piezoelectric film on diamond for high-frequency surface acoustic wave devices, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. 52, No. 10, 2005, pp. 1806–1811, ISSN 0885-3010.

Besmaine, S.; Brizoual, L. Le; Elmazria, M.; Fundenberger, J. J.; Belmahi, M. & Benyoucef, B. (2008). SAW devices based on ZnO inclined c-axis on diamond, *Diamond and Related Materials*, Vol. 17, No. 7-10, 2008, pp. 1420-1423, ISSN 0925-9635.

Comini, G.; Faglia, G. & Sberveglieri X. (2009). *Solid State Gas Sensing*, Springer Science+Business Media LLC, ISBN 0387096647, New York, USA.

Cullen, D. & Reeder, T. (1975). Measurement of SAW Velocity Versus Strain for YX and ST Quartz, *Proceedings of Ultrasonics Symposium*, pp. 519-522, Sept. 1975.

Cullen, D. & Montress, T. (1980). Progress in the Development of SAW Resonator Pressure Transducers, *Proceedings of Ultrasonics Symposium*, Vol. 2, pp. 696-701, Nov. 1980.

Data Sheet-07-2093/September (2005). MSA International, ©MSA 2005, http://www.MSAnet.com

Farnell, G. W. (1977). Elastic Surface Waves, In *Surface Wave Filters*, Matthews, H. (Ed.), pp. 1-55, John Wiley, ISBN 0471580309, New York, USA.

Farnell, G. W. (1978). Types, Properties of Surface Waves, In *Acoustic Surface Waves*, Oliner, A. A. (Ed.), Springer Verlag, ISBN 0387085750, Germany.

Feldman, M. & Henaff, J. (1989). *Surface acoustic waves for signal processing*, Artech House, ISBN 0890063087, London, UK.

Golio, M. (2008). *The RF and Microwave Handbook*, 2nd Ed., CRC Press LLC, ISBN 0849372178, Boca Raton, USA.

Grate, J. W.; Snow, A.; Ballantine, D. S.; Wohltjen, H.; Abraham, M. H.; McGill, R. A. & Sasson, P. (1988). Determination of partition coefficients from surface acoustic wave vapor sensor responses and correlation with gas-liquid chromatographic partition coefficients, *Analytical Chemistry*, Vol. 60, No. 9, May 1988, pp. 869-875, ISSN 0003-2700.

Grate, J. W. & Klusty, M. (1991). Surface Acoustic Wave Vapor Sensor Based on Resonator Devices, *NRL Memorandum report 6829*, May 23, 1991, pp. 1-38.

Grate, J. W.; Klusty, M.; McGill, R. A.; Abraham, M. H.; Whiting, G. & Andonian-Haftvan, J. (1992). The predominant role of swelling-induced modulus changes of the sorbent phase in determining the responses of polymer-coated surface acoustic-wave vapor sensors, *Analytical Chemistry*, Vol. 64, No. 6, March 1992, pp. 610-624, ISSN 0003-2700.

Grate, J. W. & Zellers, E. T. (2000). The Fractional Free Volume of the Sorbed Vapor in Modeling the Viscoelastic Contribution to Polymer-Coated Surface Acoustic Wave Vapor Sensor Responses, *Analytical Chemistry*, Vol. 72, No. 13, July 2000, pp. 2861-2868, ISSN 0003-2700.

Hakiki, El; Elmazria, M.; Assouar, O.; Mortet, M. B.; Brizoual, V. Le; Vanecek, L. & Alnot, M. (2005). ZnO AlN diamond layered structure for SAW devices combining high velocity and high electromechanical coupling coefficient, *Diamond & Related Materials*, Vol. 14, No, 3-7, 2005, pp. 1175-1178, ISSN 0925-9635.

Hashimoto, K. Y. (2000). *Surface Acoustic Wave Devices in Telecommunications: Modeling and Simulation*, Springer, ISBN 354067232X, USA.

Ho, C. K.; Lindgren, E. R.; Rawlinson, K. S.; McGrath, L. K.; Wright, J. L. (2003). Development of a Surface Acoustic Wave Sensor for In-Situ Monitoring of Volatile Organic Compounds, *Sensors*, Vol. 3, No. 7, July 2003, pp. 236-247, ISSN 1424-8220.

Hribšek M. (1982). Elementi sa površinskim akustičnim talasom i njihova primena, *Tehnika-Elektrotehnika*, Vol. 31, 1982, pp. 1725-1733. (survey paper) in Serbian

Hribšek M. (1986). Surface Acoustic Wave Devices and their applications, *Proceedings of MIEL'86*, Vol. 1, pp. 37-46, Belgrade, Serbia, 1986. (invited paper)

Hribšek, M.; Ristić, S.; Živković, Z. & Tošić, D. (2009). Modelling of SAW Biosensors, *Proceedings of the International Conference on Biomedical Electronics and Devices (BIODEVICES 2009)*, pp. 376-379, ISBN 978-989-8111-64-7, Porto, Portugal, Jan. 14-17, 2009, INSTICC Press.

Hribšek M.; Tošić D. & Radosavljević M. (2010). Surface Acoustic Wave Sensors in Mechanical Engineering, *FME Transactions*, Vol. 38, No. 1, 2010, pp. 11-18, ISSN 1451-2092.

Hribšek M. F.; Ristić S. S. & Radojković B. M. (2010). Diamond in Surface acoustic wave sensors, *Acta Physica Polonica A*, Vol. 117, No. 5, May 2010, pp. 794-798, ISSN 1898-794X.

Jian, S.; YiZhen, B.; JingChang, S.; GuoTong, D. & Xin, J. (2008). Structural and electrical properties of ZnO films on freestanding thick diamond films, *Chinese Science Bulletin*, Vol. 53, No. 19, Oct. 2008, pp. 2931-2934, ISSN 1001-6538.

Joo, B.-S.; Lee, J.-H.; Lee, E.-W.; Song, K.-D. & Lee, D.-D. (2005). Polymer Film SAW Sensors for Chemical Agent Detection, *Proceedings of the 1st International Conference on Sensing Technology*, pp. 307-310, Palmerston North, New Zealand, Nov. 21-23, 2005

Kannan, O. K.; Bhalla, R.; Kapoor, J. C.; Nimal, A. T.; Mittal, U. & Yadava, R.D.S. (2004). Detection of Landmine Signature using SAW-based Polymer-coated Chemical Sensor, *Defence Science Journal*, Vol. 54, No. 3, July, 2004, pp. 309-315, ISSN 0011748X.

Kirsch, P.; Assouar, M. B.; Elmazria, O.; Mortet, V. & Alnot, P. (2006). GHz surface acoustic wave devices based on aluminum nitride/diamond layered structure realized using electron beam lithography, *Appl. Phys. Letts.*, Vol. 88, No. 22, 2006, 223504/1-3 doi:10.1063/1.2208372

Kirsch, P.; Assouar, M. B.; Elmazria, O.; Hakiki, M. E.; Moret, V. & Alnot, P. (2007). Combination of e-beam lithography and of high velocity AlN/diamond-layered structure for SAW filters in X-band, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. 54, 2007, pp. 1486–1491, ISSN 0885-3010.

Martin, S. J.; Frye, G. C. & Senturia, S. D. (1994). Dynamics and Response of Polymer-Coated Surface Acoustic Wave Devices: Effect of Viscoelastic Properties and Film Resonance. *Analytical Chemistry*, Vol. 66, 1994, pp. 2201-2219, ISSN 0003-2700.

Mitsakakis, K. et al. (2009). SAW device integrated with microfluidics for array-type biosensing, *Microelectronics Engineering*, 2009, doi:10.1016/j.mee.2008.12.063

Mortet, V.; Williams, O. A. & Haenen, K. (2008). Diamond: a material for acoustic devices, *p hys. stat. sol. (a)*, 205, No. 5, 1009–1020, DOI 10.1002/pssa.200777502

Omori, T.; Kobayashi, A.; Takagi, Y.; Hashimoto, K. & Yamaguchi, M. (2008). Fabrication of SHF range SAW devices on AlN/Diamond-substrate, *Proceedings of Ultrasonics Symposium*, pp. 196–200, Nov. 2-5, 2008

Plessky, V. & Koskela, J. (2000). "Coupling-of-modes analysis of SAW devices,"International Journal of High Speed Electronics and Systems, vol. 10, pp. 867–947.

Pohl A. (2000). A Review of Wireless SAW Sensors. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. 47, 2000, pp. 317-332, ISSN 0885-3010.

Rayleigh, L. (1885). On waves propagated along the plane surface of an elastic solid, *Proc. London Math. Soc.*, Vol. 17, 1885, 4-11.

Rufer, L.; Torres, A.; Mir, S.; Alam, M. O.; Lalinsky, T.; Chan, Y. C. (2005). SAW chemical sensors based on AlGaN/GaN piezoelectric material system: acoustic design and packaging considerations, *Proceedings of the 7th International Conference on Electronics Materials and Packaging, EMAP 2005*, pp. 204-208, 2005, Tokyo, Japan.

Rufer, L.; Lalinský, T.; Grobelný, D.; Mir, S.; Vanko, G.; Őszi, Zs.; Mozolová, Ž. & Gregus, J. (2006). GaAs and GaN based SAW chemical sensors: acoustic part design and technology, *Proceedings of the 6th International Conference on Advanced Semiconductor Devices and Microsystems, ASDAM 2006*, pp. 165-168, 2006, Smolenice, Slovakia.

Sankaranarayanan, S.; Bhethanabotla, V. R. & Joseph, B. (2005). A 3-D Finite Element Model of Surface Acoustic Wave Sensor Response, *Proceedings of the 208th ECS Meeting, Acoustic Wave Based Sensors and Sensor Systems*, Vol 1, Issue 19, pp. 19-27, ISBN 978-156-677-493-2, Los Angeles, California, October 16-21, 2005, Curran Associates, Inc.

Seifert, F.; Bulst, W.E. & Ruppel, C. (1994). Mechanical sensors based on surface acoustic waves, *Sensors and Actuators A*, Vol. 44, No. 3, 1994, pp. 231-239, ISSN 0924-4247.

Shikata, S.; Fujii, S.; Uemura, T.; Itakura, K.; Hachigo, A.; Kitabayashi, H.; Nakahata, H. & Takada, Y. (2005). Improvements of diamond SAW device characteristic and applications to communication system, *New Diamond and Frontier Carbon Tehnology*, Vol. 15, No. 6, 2005, pp. 350-361, ISSN 1334-9931.

Specht, C. G.; Williams, O. A.; Jackman, R. B. & Schoepfer, R. (2004). Ordered growth of neurons on diamond, *Biomaterials*, Vol. 25, No. 8, Aug. 2004, pp. 4073- 4078, ISSN 0142-9612.

Sung, C.-C.; Chiang, Y.-F.; Ro, R.; Lee, R. & Wu, S. (2009). Theoretical analysis of SAW propagation characteristics in (100) oriented AlN/diamond structure, *Proceedings of the Frequency Control Symposium 2009 Joint with the 22nd European Frequency and Time forum, IEEE International*, pp. 446-449, ISSN 1075-6787, Besancon, April 2009.

Tiersten, H. F. & Sinha, B. K. (1978). A perturbation analysis of the attenuation and dispersion of surface waves, *Journal of Applied Physics*, Vol. 49, No. 1, Jan 1978, pp. 87-95, ISSN 0021-8979.

White, R. M. & Voltmer, F. W. (1965). Direct Piezoelectric Coupling to Surface Electric Waves, Applied Physical Letters, Vol. 7, No. 12, Dec. 1965, pp. 314-316, ISSN 0003-6951.

Wohltjen, H.; Dessy, R. (1979). Surface acoustic wave probe for chemical analysis. I. Introduction and instrument description, *Analytical Chemistry*, Vol. 51, No. 9, Aug. 1979, pp. 1458-1464, ISSN 0003-2700.

Wohltjen, H. (1982). US. Patent 4,312,228, Jan. 26, 1982.

Wohltjen, H. (1984). Mechanism of operation and design considerations for surface acoustic wave device vapour sensors, *Sensors and Actuators*, Vol. 5, No. 4, July 1984, pp. 307-325, ISSN 0924-4247.

Zaki, A.; Elsimary, H. & Zaghloul, M. (2006). Miniature SAW Device for RF-Wireless Applications Using MEMS Technology, *Proceedings of the 5th WSEAS International Conference on Circuits, Systems, Electronics Control and Signal Processing*, pp.10–13, ISBN ~ ISSN 1790-5117, 960-8457-55-6, Dallas, USA, Nov. 1–3, 2006, World Scientific and Engineering Academy and Society (WSEAS), Stevens Point, Wisconsin, USA.

Živković, Z. (2003). SAW sensors and their applications, *Proceedings of XLVII ETRAN*, Vol. III, pp. 424-427, Herceg Novi, Monte Negro, June 2003. (in Serbian)

Živković, Z.; Hribšek, M. & Tošić, D. (2009). Modeling of Surface Acoustic Wave Chemical Vapor Sensors, *Journal of Microelectronics, Electronic Components, Materials*, Vol. 39, No. 2, June 2009, pp.111-117, ISSN 0352-9045.

17

Laser-Based Determination of Decohesion and Fracture Strength of Interfaces and Solids by Nonlinear Stress Pulses

Peter Hess
University of Heidelberg
Germany

1. Introduction

To determine the mechanical strength of films, coatings and solids the loss of adhesion or fracture must be investigated. The failure of interfaces and bulk materials is a dominant issue in microelectronics with multilayer systems and micro-electro-mechanical-system (MEMS), nano-electro-mechanical-system (NEMS) and sensor devices, based essentially on single-crystal silicon as basic material, which is considered here in detail.
A versatile tool that is quite often used for strength analysis is the scratch tester, where a diamond stylus is drawn across the coated surface under increasing load to determine the critical load, or the stylus is used as an indenter. These methods are versatile, but besides being influenced by the properties of the system itself they also depend on several test parameters, such as scratching velocity and stylus properties, which affect the critical load. Owing to the complexity of the failure processes involved, in connection with strongly inhomogeneous deformation fields, it is generally very difficult to extract quantitative values of the cohesion or fracture strength (Lacombe, 2006). In fact, the most widely used testing methods, such as peel, pull, scratch, blister, indentation and beam-bending tests, usually involve plastic deformations, which are difficult to analyze (Wei & Hutchinson, 1998). Some of these quasi-static methods do not reach the strength limit of the strongest material systems or require intricate sample preparation.
In this review the contact-free measurement of the strength of interfaces and bulk materials will be discussed employing pulsed lasers to launch strongly nonlinear stress pulses. In fact, quantitative information on the failure strength of materials can be obtained by laser-based excitation and detection techniques. Normally a nanosecond laser pulse is used to excite either bulk or surfaces acoustic stress waves, which develop shocks during propagation. In these laser-controlled pump-probe setups a continuous-wave (cw) laser probe is used to measure the transient surface displacement or surface velocity, providing the information on the elastic stresses achieved. The laser techniques are contact-free and normally need no artificial seed crack to induce failure because stresses of 5-10 GPa can be attained in the elastic shock pulses generated.
For studying interfacial strengths usually bulk acoustic waves are used. With a laser pulse a one-dimensional (1D) compressive longitudinal wave packet is launched in a thin metal film covering the back side of the substrate. The critical failure stress of the film/substrate

interface is reached by the transformation of this compressive pulse into a tensile stress pulse upon reflection at the opposite free film surface. The critical load is extracted from the surface distortions caused by pulse reflection at the film surface (see, for example, Wang et al., 2003a).

Surface acoustic waves (SAWs) are employed to elucidate the fracture modes of solids by inducing surface-breaking cracks. The elliptically polarized surface wave pulses are generated by strongly focusing laser pulses into a line at the surface. These elastic pulses develop shock fronts during propagation along the surface by frequency-up conversion processes, if the shock formation length is smaller than the attenuation length of the wave. The transient displacement or surface velocity of the elastic pulse propagating along the surface is monitored at two locations (Lomonosov et al., 2001).

When dynamic fracture techniques with high strain rates of $\geq 10^6$ s^{-1} are used the undesirable inelastic deformation effects of quasi-static methods can be reduced considerably. Such high elastic peak stresses can be generated, for example, by high velocity impact of a flyer plate, by detonations, by particle beams or pulsed laser irradiation. The application of lasers is preferable owing to contact-free operation and the spatial control of the impulsive load by the laser pulse energy, pulse duration and the possibility of focusing the radiation.

In high-rate fracture experiments with nanosecond bulk waves usually the resistance to dynamic failure or spall strength is measured. Note that this is not an intrinsic material property because its value depends on the competition between stress growth due to wave interactions and stress relaxation due to the formation of voids. In ductile materials spallation is a rather complex kinetic process including nucleation, growth and coalescence of microvoids and microcracks. For this reason, it is not an intrinsic threshold phenomenon (Kanel, 2010). In brittle materials failure occurs by nucleation of microcracks and their further extension. Important information has been reported on the ultimate strength of layered systems involving different film/substrate combinations of metals, dielectrics and semiconductors.

In dynamic fracture experiments using nanosecond SAW pulses, similar strain rates $\geq 10^6$ s^{-1} can be realized. Owing to the strong evolution of tensile stress during pulse propagation, no reflection of the surface pulse is needed, as in the case of bulk waves. Since in single crystals plane SAWs propagate along a well-defined crystallographic plane and in a given direction, they allow the investigation of the anisotropic fracture behaviour. Extensive results for the formation of surface-breaking cracks in different geometries (crystal planes and directions) are discussed for single-crystal silicon. The method also yields the critical fracture stress for the selected geometries. The size dependence of strength at the nano-scale is also discussed.

2. Experimental

The quantitative investigation of spallation, delamination and fracture mechanics by laser ultrasonics is based on the efficient excitation of longitudinal, shear and surface acoustic waves by strong nanosecond (picosecond or femtosecond) laser pulses. Especially methods that apply bulk stress pulses to study mode-I, mode-II and mixed-mode failure in film systems have found widespread use in recent years. The investigation of crack formation in bulk materials by nonlinear SAWs is a relatively new field of nonlinear ultrasonics.

The original bulk method is based on laser excitation of longitudinal compressive waves on the side of the sample opposite to the film (see Fig. 1). The nanosecond pulse of a Nd:YAG

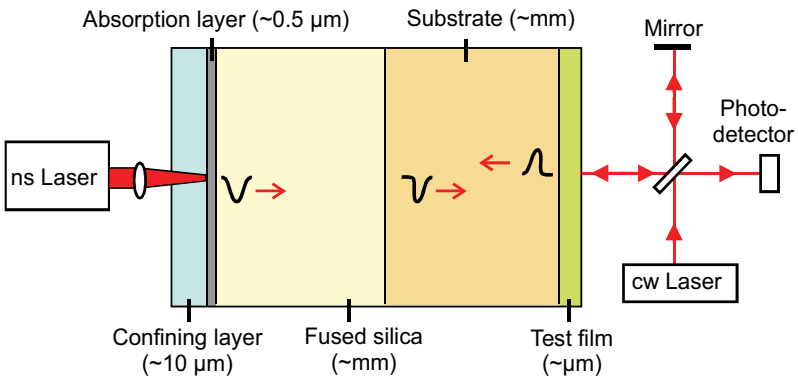


Fig. 1. Setup for tensile interfacial spallation with a pulsed laser, consisting of layers for shock formation, substrate with test film and interferometer to monitor the film surface (Gupta et al., 1990; Wang et al., 2003a).

laser is collimated at the absorbing medium to an area of about 1-3 mm diameter (Gupta et al., 1990). Usually, the elastic pulse is generated by absorption of laser radiation in a ~0.5 µm thick metal (Al) layer, which is sandwiched between the back side of the substrate and an about 10 µm thick transparent confining layer (SiO_2, waterglass). Instead of an Al film as absorbing medium a 20 µm thick layer of silicone grease, containing fine MoS_2 particles, has been employed to excite stress waves by laser breakdown (Ikeda et al., 2005). Tensile stress is generated when the compression pulse is reflected from the free film surface. At the surface the resulting stress is zero and reaches its maximum at a distance equal to half of the spatial pulse extension. This restricts the thickness of the films to be delaminated. The situation can be improved by modifying the profile of the stress pulse using an unusual nonlinear property of fused silica, which develops a rarefaction shock at the tail of the pulse for compressive stresses below 4 GPa, as shown in Fig. 1. (Wang et al., 2003a). In this case the tensile stress reaches its maximum at a distance of the width of the post-peak shock, making the method applicable for significantly thinner films. From measurements of the transient out-of-plane displacement or velocity of the free film surface at the epicenter by a laser interferometer, the interfacial strength is obtained for specular and diffuse surfaces using a cw laser as probe (Pronin & Gupta, 1993).

Besides longitudinal stress pulses also shear pulses can be obtained by using a triangular fused silica prism for partial mode conversion of the excited longitudinal compressive wave into a shear wave upon oblique incidence onto a surface, as illustrated in Fig. 2 (Wang et al., 2003b; Wang et al., 2004; Hu & Wang, 2006; Kitey et al., 2009). With an optimized setup, nearly complete conversion into high amplitude shear pulses, and therefore mode-II fracture by in-plane shear stress, can be achieved at a prism angle of $\theta = 57.7°$ (Hu & Wang, 2006). In fact, controlled mixed-mode loading and the quantitative analysis of the stresses involved is possible. It is important to note that in most practical situations thin films tend to fail under mixed-mode I+II conditions.

Controlled dynamic delamination of thin films has been achieved recently by insertion of a weak adhesion region below the film to be delaminated (Kandula et al., 2008a). While spallation experiments characterize the interface strength or critical stress for microvoid or microcrack initiation the delamination process can be more closely associated with the propagation of cracks.

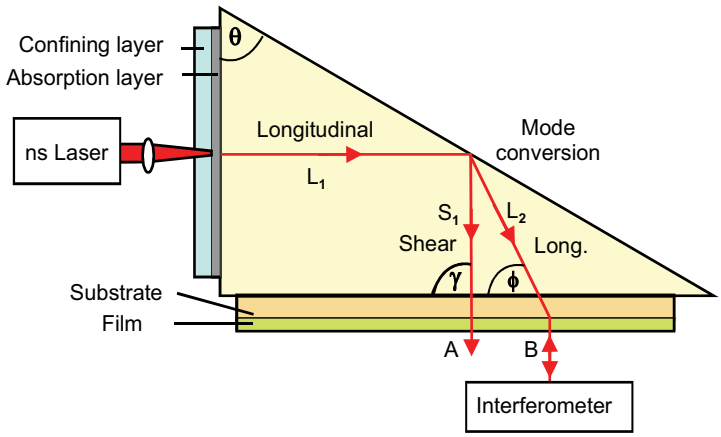


Fig. 2. Setup for shear stress delamination with a pulsed laser, fused-silica prism for mode conversion and interferometer to monitor the film surface (Hu & Wang, 2006).

In only a few studies have femtosecond lasers been employed to investigate spallation of metal targets (Tamura et al., 2001). With such ultrashort laser pulses ultrafast strain rates of $\geq 10^8$ s^{-1} may be accessible. These laser pulses with intensities in the 10^{15} W/cm^2 range launch shock pulses with a steep unloading stress profile. The effects of femtosecond laser-driven shocks using very high laser pulse energies have been described recently based on time-resolved measurement of the surface velocity by Doppler interferometry (Cuq-Lelandais et al., 2009; de Rességuier et al., 2010).

SAWs are guided waves that penetrate approximately one wavelength deep into solids. Thus, the main part of the elastic energy stays within this depth during wave propagation along the surface. Note that the elliptically polarized surface waves possess in-plane and out-of-plane displacements, and thus both a longitudinal and shear component. In the corresponding pump-probe setup a pulsed nanosecond laser is employed to launch a nanosecond SAW pulse with finite amplitude, which is sufficiently nonlinear to develop shocks during propagation (Lomonosov et al., 2001). A distinctive property of SAWs is their intrinsic tensile stress and its further development during nonlinear pulse evolution. A cw laser is used for detection of the moving surface distortions at two different surface locations (Kolomenskii et al., 1997).

Typically, a Nd:YAG laser radiating at 1.064 μm with 30–160 mJ pulse energy and 8 ns pulse duration was applied in single-pulse experiments. As depicted in Fig. 3, the explosive evaporation of a thin layer of a highly absorbing carbon suspension (ink), deposited only in the source region, is used to launch SAW pulses with sufficient amplitude for nonlinear evolution. By sharply focusing the pump laser pulse with a cylindrical lens into a narrow line source, a plane surface wave propagating in a well-defined crystallographic direction is launched. If the shock formation length is smaller than the attenuation length a propagating SAW pulse with finite amplitude develops a steep shock front. These nonlinear SAW pulses gain amplitudes of about 100–200 nm, as compared with few nanometers for linear SAWs. The shape of the pulse changes not only due to frequency-up conversion but in addition frequency-down conversion processes take place, caused by the elastic nonlinearity of the solid. The value of the absolute transient surface displacement can be detected with a stabilized Michelson interferometer. In most experiments, however, the more versatile

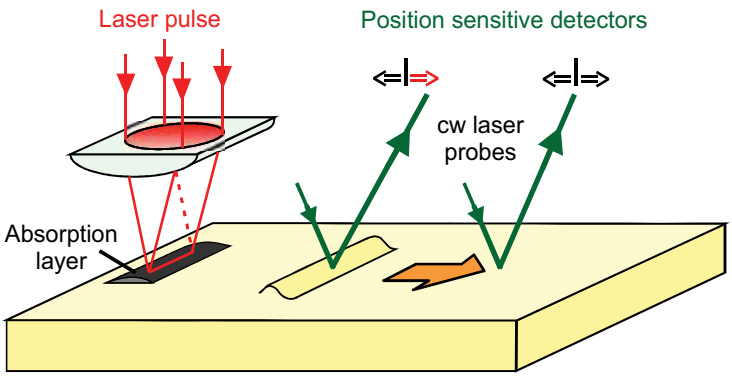


Fig. 3. Setup for exciting plane SAW pulses with shocks using pulsed laser irradiation and two-point cw laser probe-beam deflection to monitor the transient surface velocity (Lomonosov et al., 2001).

transient deflection of a cw probe-laser beam is monitored by a position-sensitive detector, to determine the surface velocity or shear displacement gradient (Lomonosov et al., 2001). In the two-point-probe scheme the SAW profile usually is registered at distances of 1–2 mm and 15–20 mm from the line source. The pulse shape measured at the first probe spot is inserted as an initial condition in the nonlinear evolution equation to simulate the nonlinear development of the SAW pulse and to verify agreement between theory and experiment at the second probe spot.

3. Interfacial decohesion by longitudinal and shear waves

3.1 Determination of the interfacial spallation and delamination strength

Up to now in most bulk experiments longitudinal pulses have been used for spatially localized spallation or delamination of films by pure tensile stresses (mode I). As discussed before, the tensile stress pulse reflected at the free film surface is responsible for the more or less complete removal (spallation) of the film predominantly in the irradiated area. From the interferometric measurement of the transient out-of-plane displacement at the film surface, the stress development in the substrate and at the interface can be inferred.

For a substrate with a single layer the evolution of the substrate stress pulse σ_{sub} and the interface stress σ_{int} are determined using the principles of wave mechanics. If the film thickness h is smaller than the spatial spread of the substrate pulse during the rise time t_{rise}, i.e., $h << c_{film} \times t_{rise}$, where c_{film} is the wave speed in the film, the following approximations can be applied to estimate the substrate and interface stresses. Note that in this situation the loading region is large compared to the actual film thickness. For a Gaussian compressive 1D stress pulse launched in the absorbing metal layer and propagating towards the substrate one finds under this condition (Wang et al., 2002).

$$\sigma_{sub}(t) = -\frac{1}{2}(\rho c)_{sub}\frac{du}{dt} \tag{1}$$

where the assumption is made that the displacement amplitude of the wave in the substrate is half that at the free surface, and u is the displacement of the free film surface. Here ρ is the density of the substrate and c the longitudinal speed of the stress wave in the substrate.

The tensile stress acting at the film/substrate interface can be estimated by assuming that the stress is given by $(\rho h)_{film}$ multiplied by the acceleration of the free surface

$$\sigma_{int}(t) = -(\rho h)_{film} \frac{d^2u}{dt^2} \tag{2}$$

The subscripts 'sub' and 'film' represent substrate and film properties, respectively. When the film thickness becomes comparable to the spatial extension of the rising part of the pulse, i.e., $h \approx c_{film} \times t_{rise}$, the following equation provides a more accurate 1D description of the stress history at the interface, because in reality the stress loading of the interface results from the superposition of the incoming compressive wave and the reflected tensile pulse

$$\sigma_{int}(t,h) = \frac{1}{2}(\rho c)_{film}\left[v(t + h / c_{film}) - v(t - h / c_{film})\right] \tag{3}$$

Here v is the measured surface velocity v = du/dt (Gupta et al., 2003). For small values of h/c_{film} this equation transforms into Equation (2), which is analogous to Newton's second law of motion, stating that the interface tensile strength is given by the mass density of the film times the outward acceleration of the centre of mass of the film (Wang et al., 2002). These 1D approximations provide physical insight into the relevant stress loading processes. Numerical simulations are needed to obtain a more accurate description of the three-dimensional evolution of the stress field.

The treatment of the more complicated mixed-mode case, where tensile and shear stresses act simultaneously, can be found in several publications (Wang et al., 2003b); Wang et al., 2004; Hu & Wang, 2006; Kitey et al., 2009). In these reports the equations have been derived that are needed to extract the interfacial adhesion strengths for mixed-mode failure and to compare these results with those for purely tensile loading.

Here the derivation is presented for an experimental arrangement similar to the one shown in Fig. 2, where the shear wave travels nearly perpendicular to the film surface ($\phi = 60°$ and $\gamma \approx 86.9°$), following Hu and Wang (2006). The stress waves S_1 and L_2 load the film interface with different mode-mixities at points A and B. At these points another mode conversion takes place, when S_1 and L_2 reach the film surface. The out-of-plane displacements $u_{\perp A}$ and $u_{\perp B}$ and the in-plane displacements $u_{||A}$ and $u_{||B}$ can be calculated as a function of L_1, S_1, and L_2 (Hu & Wang, 2006). The results indicate that the out-of-plane displacement at point B is about 2.5 times that at point A.

From the information on the displacements the substrate and interface stresses are derived on the basis of the 1-D approximation (Hu & Wang, 2006)

$$\sigma_{sub}^{L_2} = -(\rho c_\ell)_{sub} \frac{du_{L_2}}{dt} \tag{4}$$

$$\tau_{sub}^{S_1} = -(\rho c_s)_{sub} \frac{du_{S_1}}{dt} \tag{5}$$

$$\sigma_{film}^{X} = -(\rho h)_{film} \frac{d^2u_{\perp X}}{dt^2} \tag{6}$$

$$\tau_{int}^{X} = -(\rho h)_{film} \frac{d^2 u_{||X}}{dt^2} \quad (7)$$

where the normal and shear stresses in the substrate at the points X (A or B) caused by L_2 and S_1 are given by $\sigma_{sub}^{L_2}$ and $\tau_{sub}^{S_1}$, respectively (see Fig. 2). The corresponding normal and shear interface stresses are σ_{int}^{X} and τ_{int}^{X} and the out-of-plane and in-plane displacements are $u_{\perp X}$ and $u_{||X}$, respectively. The relatively large ratio of the shear to the normal interface stress of ~14 indicates nearly pure shear loading for this particular configuration of the silica prism.

3.2 Results of interface spallation and delamination experiments

Especially spallation experiments have been performed for a large variety of layered material systems. The controlled delamination of a film is more difficult to achieve, but has been reported recently (Kandula et al., 2008a). In the following results obtained for some characteristic systems are selected to illustrate the potential of this laser-based method to study pure and mixed-mode decohesion of thin films in layered systems.

Si/Si_xN_y/Au system

Mixed-mode failure was studied in this particular work using a silicon wafer of 730 nm thickness covered with a Si_xN_y passivation layer (400 nm) and an Au film of thickness of 300 nm, 600 nm, or 1200 nm reported recently (Kandula et al., 2008a). The back side of the silicon substrate was bonded to a fused silica prism equipped with an Al layer (400 nm) and a confining waterglass layer. For the pure tensile strength between the Au film and passivated silicon substrate a critical stress of 245 MPa was found. Under mixed-mode conditions, delamination was observed at about 142 MPa tensile stress and about 436 MPa shear stress. Thus, by applying the shear load the tensile strength was reduced by approximately 100 MPa. The effective stress in the mixed-mode case was about 449 MPa. An interpretation of this finding in comparison with mode-I failure is that mixed-mode decohesion consumes more energy. It is important to note that the laser spallation method clearly yields mode-resolved strength values, whereas the stress fields generated by conventional scratch, peel, pull, blister and indentation tests are difficult to analyze quantitatively due to stress inhomogeneities and plastic deformations involved in these techniques.

To illustrate the whole measurement and evaluation procedure of this laser technique, the registered photo-diode signal is presented in Fig. 4a), the corresponding normal surface displacement is shown in Fig. 4b), the substrate shear stress is displayed in Fig. 4c) and the tensile and shear stress components acting at the interface are exhibited in Fig. 4d) for a 600 nm Au film deposited on a passivated silicon substrate (Kitey et al., 2009).

Si/TaN/Cu system

In the case of very thin films, the reflected tensile pulse may overlap with the incoming compressive pulse, reducing the effective stress at the interface. In this situation it can happen that the critical fracture strength of the substrate material is first reached at a certain penetration depth of the tensile pulse into the substrate. By increasing the film thickness the incoming and reflected pulse can be separated, finally leading to film spallation. Such a behaviour has been observed for silicon covered by a bilayer of TaN/Cu. The TaN layer thickness was fixed at 20 nm, whereas the Cu layer was varied in five steps between 100 nm and 10 μm. At a Cu-layer thickness ≤1 μm, silicon fracture with an intrinsic tensile strength

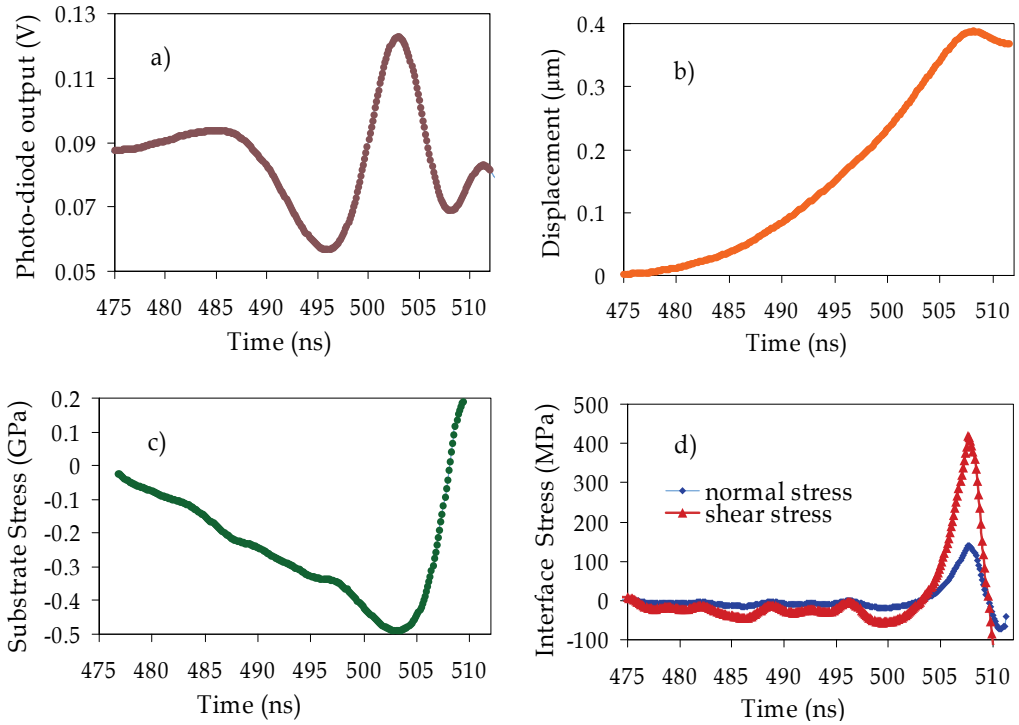


Fig. 4. Mixed-mode failure of a 600 nm Au film on passivated silicon: a) displacement fringes of incident shear wave S_1, b) time dependence of out-of-plane displacement of film surface, c) substrate shear stress τ_{sub} and d) normal and shear interface stresses (Kitey et al., 2009).

of approximately 5 GPa was observed, while at Cu-layer thicknesses of ≥5 μm the Si/TaN interface was debonded at about 1.4 GPa (Gupta et al., 2003).

Silica/W/W film system

Recently, the fracture of bulk polycrystalline tungsten and spallation of a tungsten/tungsten interface, produced by magnetron sputtering of a tungsten film, was studied (Hu et al., 2009). For polycrystalline bulk tungsten a strength of 2.7–3.1 GPa was found. Crack propagation occurred essentially along certain crystallographic orientations by coalescence of microvoids due to grain boundary decohesion. Only at extremely high strain rates did in-plane cracks not distinguish between the bulk and boundaries of grains and propagated along relatively straight paths of lengths two-to-three times the laser loading diameter. The observed spall strengths were substantially higher than the value of ~0.5 GPa reported for plate-impact shock loading at strain rates of about 10^5 s^{-1} and the stress of 1.2 GPa observed under quasi-static loading (see Hu et al., 2009). The interfacial strength of the tungsten/tungsten interface, created by sputtering, was only 875 MPa.

Si/Si_xN_y/PBO system

The interface strength of a dielectric polymer film has been studied in a multilayer system (Si/Si_xN_y/PBO) consisting of a poly(*p*-phenylene benzobisoxazole) (PBO) film (5 μm), which is used as stress buffer in microelectronics, a silicon nitride (Si_xN_y) interface layer of 30 nm

or 400 nm thickness and a silicon wafer (Kandula et al., 2008b). Stress wave propagation in this multilayer system was analyzed analytically and numerically, by neglecting the influence of the silicon nitride layer in the analysis. At strain rates of about 10^7 s^{-1} and laser fluences of 65 mJ/mm^2, compressive stresses of up to 3.5 GPa could be obtained. Such a stress is sufficient to fracture bulk silicon in certain configurations as observed already before (Wang et al., 2002). As expected, the failure of the film interface was observed at much lower laser fluences and varied strongly with the preparation, treatment and thickness of the PBO layers, yielding an upper tensile interface stress of about 0.35 GPa.

Si/neuron cell system

First experimental results and finite element simulations on the extension of the laser-induced bulk stress wave technique to the investigation of biological samples such as cell/substrate adhesion have been reported (Hu et al., 2006). In this pioneering work the noncontact detachment of neuron cells from a silicon substrate was studied. Since the time scale of the experiment is in the nanosecond range cells remain essentially undisturbed before their detachment, which is not the case with other techniques. While adhesion could be characterized only in terms of the critical Nd:YAG laser fluence, it can be expected that the method will be able to quantify the adhesion strength in the near future. The principal detachment mechanism predicted by the simulations performed is strain-driven failure resulting from the cell's tendency to flatten and elongate along the substrate (Miller et al., 2010).

4. Fracture of anisotropic crystals by surface acoustic wave pulses

4.1 Determination of the bulk fracture strength

With SAWs, strong nonlinearities and very high strains in the range of 0.01 can be realized much more easily than with bulk waves (Lomonosov et al., 2001; Lomonosov & Hess, 2002). As mentioned before, SAWs are guided waves that only penetrate approximately one wavelength deep into the solid. This particular property reduces diffraction losses as compared with acoustic bulk waves. In addition, frequency-up conversion concentrates the energy in an even smaller depth from the surface. For certain crystal geometries the displacements of SAWs are confined to the sagittal plane, defined by the in-plane propagation direction x_1 and the surface normal x_3. Thus, x_2 is normal to the sagittal plane.

To extract quantitative values of fracture strengths from experiments with laser-induced SAWs, a theoretical description of shock formation in a SAW pulse with finite amplitude during its propagation in a nonlinear elastic medium is required. A suitable nonlinear evolution equation that also takes into account dispersion of SAWs has been developed to describe solitary surface pulses in layered systems (Lomonosov et al., 2002; Eckl et al., 2004; Hess & Lomonosov, 2010). In systems without a length scale, such as single crystals, the dispersion term is not needed because SAWs are not dispersive. Therefore, in silicon, the profiles of the recorded SAW pulses were simulated by solving the following dispersionless nonlinear evolution equation

$$i\frac{\partial}{\partial\tau}B_n = nq_0\left[\sum_{0<n'<'n} F(n'/n)B_{n'}B_{n-n'} + 2\sum_{n'>n}(n/n')F^*(n/n')B_{n'}B^*_{n'-n}\right] \tag{8}$$

where B_n is the n-th harmonic of the signal, τ the stretched coordinate along the direction of wave propagation, q_0 the fundamental wave number and F(x) a dimensionless function. This

function describes the efficiency of frequency conversion and depends on the ratio of the second-order to third-order elastic constants of the selected geometry. For example, F(1/2) describes the efficiency of second-harmonic generation. Comparison with experiments showed that this equation provides a quantitative description of nonlinear SAW evolution (Lomonosov & Hess, 2002; Lehmann et al., 2003).

Experimentally, the SAW pulse is measured at two surface spots by laser-probe-beam deflection, one 1-2 mm from the source and the other at a distance of 15–20 mm (see Fig. 3). The calibration procedure exploits the predictor-corrector method for the iterative solution of the evolution equation, which connects the Fourier components of the transient profiles measured at the first and the second probe spots. Since the distance between the two probe spots was fixed, the observed changes depend only on the initial magnitude of the absolute strain. The aim was to determine the calibration factor 'a', with the dimension [1/volt], in the equation $u_{31} = a \times U(t)$, where u_{31} is the surface velocity or shear displacement gradient and U(t) is the signal measured at the first probe spot. The solution with correct calibration factor should describe the profile registered at the second probe spot and allows one to estimate the absolute surface strain at any other location, e.g., where a surface crack can be seen. The spectrum of the initial laser-excited transient was limited to about 200 MHz, mainly due to the laser pulse duration of 8 ns. The purpose was to measure the surface slope at a position close to the source, where frequency components in the gigahertz range are still negligible. As can be clearly seen in Fig. 5, the sharp spikes developed at larger propagation distances could no longer be recorded with the experimental setup. Since in a nonlinear medium like a silicon crystal both frequency-up conversion and frequency-down conversion processes take place, a lengthening of the pulse profile occurs simultaneously with shock formation. This effect is proportional to its magnitude, and therefore the pulse length can be used as a sensitive measure of the nonlinear increase of strain. In particular, when the shock fronts become steeper this quantity can be determined quite accurately (Lomonosov & Hess, 2002; Kozhushko & Hess 2007).

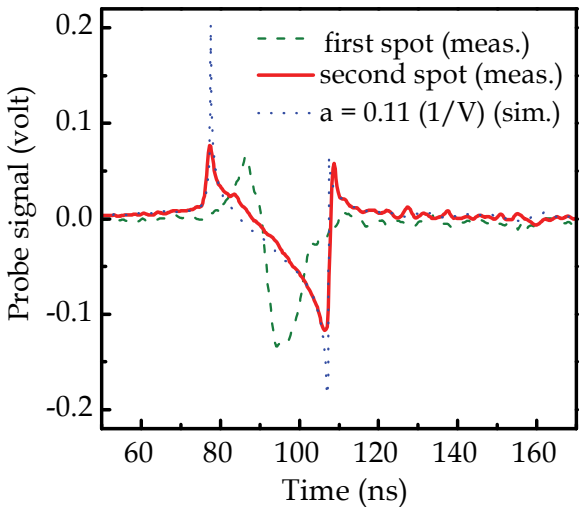


Fig. 5. Typical pulse shapes measured at the first and second probe spots in silicon. Comparison of the latter experimental profile with the predicted shape with spikes explains the calibration procedure of fitting the length of the pulse (Kozhushko & Hess, 2007).

4.2 Results for mode-resolved fracture strength of silicon

Up to now there is no generally accepted microscopic theory of brittle fracture of materials, because only simulations are possible on the molecular level. Certainly, dynamic fracture consists of two stages, namely nucleation and subsequent propagation of the crack tip. In the experiments considered here, fracture was induced by intrinsic surface nucleation with SAWs propagating along defined geometries. For some special geometries the shocked SAW pulse introduced not only a single crack but a whole field of about 50–100 μm long cracks by repetitively fracturing the crystal after a certain additional propagation distance along the surface that was sufficient to restore the shocks.

Previous fracture experiments indicate that the {111} plane is the weakest cleavage plane in silicon. Failure usually occurred perpendicular to the SAW propagation direction and extended along one of the Si{111} cleavage planes into the bulk. There are three orthogonal pairs of stress components defining three fracture modes, namely tensile or opening σ_{11}, in-plane shear or sliding σ_{31} and out-of-plane shear or tearing σ_{21}, briefly called fracture modes I, II and III, respectively.

By assuming that the {111} plane is the weakest cleavage plane of silicon, geometries were chosen where the intersection line of the {111} cleavage plane with the free surface was normal to the wave vector of the plane SAW pulse. The four basic cleavage planes provide a set of possible orientations. We studied the geometries Si(112)[111], Si(111)[112], Si(223)[334] and Si(221)[114], which are a subset of the general set of geometries (m m n)[n n 2m], where the particle displacements are confined to the sagittal plane, and therefore only the σ_{11} opening stress component has a non-zero value at the surface (Kozhushko & Hess, 2008). Note that in the coordinate system associated with tilted cleavage planes the initial σ_{11} stress can be represented by simultaneously acting orthogonal components, which are associated with a tensile mode and an in-plane shearing mode. The orientation of the family of {111} cleavage planes, which are normal to the sagittal plane, is displayed in Fig. 6. In all these cases, the initial σ_{11} opening stress can be represented by two orthogonal components with their ratio defined by the tilt angle of the cleavage plane with respect to the surface normal.

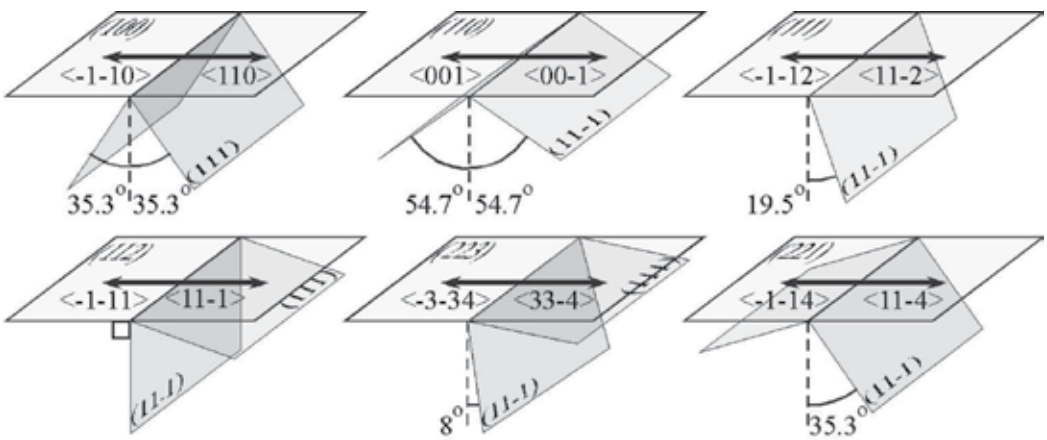


Fig. 6. Crystallographic configurations of the Si{111} cleavage planes normal to the sagittal section, e.g., for the subset of the (m m n)[n n 2m] geometries (Kozhushko & Hess, 2007).

In the following discussion results are presented for the low-index planes Si(112), Si(111), Si(223), Si(221) and Si(110) and SAW propagation in selected directions, described in more detail previously (Lomonosov & Hess, 2002; Kozhushko et al., 2007; Kozhushko & Hess, 2007; Kozhushko & Hess, 2008; Kozhushko & Hess, 2010)

Silicon (112) plane

Initiation of impulsive fracture by nonlinear SAW pulses in the Si(112)< $\overline{1}\overline{1}1$ > geometry revealed that SAW pulses propagating in the < $\overline{1}\overline{1}1$ > direction induced fracture at significantly lower laser pulse energies, and thus at lower SAW strains, than the mirror-symmetric wave propagating in the opposite <11 $\overline{1}$ > direction. This surprising effect is a consequence of differences in the elastic nonlinearity of the two propagation directions.

The easy-cracking configuration was used for fracture experiments with low laser pulse energies of 30–40 mJ. An optical microscope image of the induced crack field of a typical fractured surface is presented in Fig. 7. The vertical line at the right-hand side is the imprint of the laser-generated line source. The position of the first probe spot was approximately 0.5 mm from the source. With further propagation the finite SAW pulse developed the critical stress needed for fracture. At a distance of about 1 mm from the source the first crack can be seen. For crack nucleation and formation of the crack faces a certain amount of energy is needed. The resulting loss in pulse energy mainly reduces the high frequency part of the SAW pulse spectrum. The crack field extending further to the left-hand side is the result of repetitive fracture processes, occurring due to repetitive recovery of the shock fronts during propagation after each fracture event.

On the surface the cracks extended into the < $1\overline{1}0$ > direction, perpendicular to the SAW propagation direction and sagittal plane, with a length of up to 50 μm, controlled by the length of the SAW pulse in the nanosecond range. As expected, failure occurred along the intersection line of the surface with the {11 $\overline{1}$ } cleavage plane (see Fig. 6). The resulting peak value of the σ_{11} stress at the surface is associated with the tensile strength of the material for nucleation of cracks at the surface. A series of experiments yielded about 4.5 GPa for the critical opening stress of silicon at the surface in this particular geometry. Note that here only normal stress acts on the {11 $\overline{1}$ } cleavage plane, which is perpendicular to the surface for this particular geometry, and consequently the nucleation of cracks can be considered as a pure mode-I process (Kozhushko & Hess, 2007).

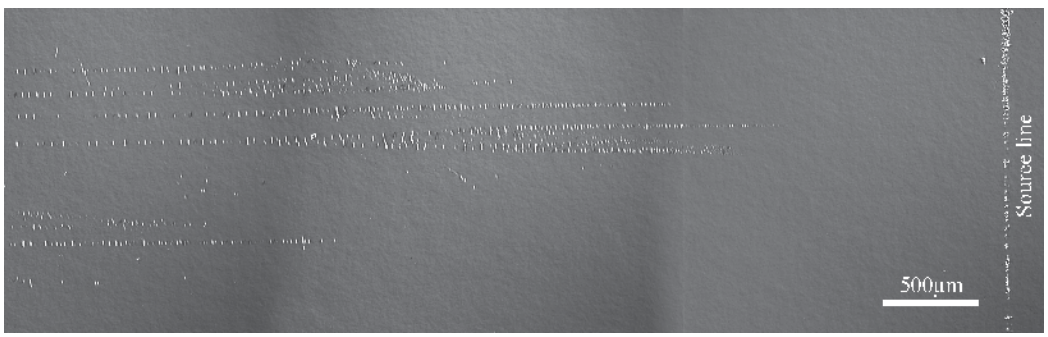


Fig. 7. Optical microscope image of the Si(112) surface after propagation of a single nonlinear SAW pulse in the < $\overline{1}\overline{1}1$ > direction from the source on the right side to the left (Kozhushko & Hess, 2007).

Silicon (111) plane

For this geometry we also made the observation that counterpropagating nonlinear SAW pulses, moving in opposite directions, e.g., in the $< \bar{1}1\,2>$ and $<11\,\bar{2}>$ directions on the Si(111) plane, develop completely different nonlinear pulse shapes. In the easy-cracking geometry Si(111) $< \bar{1}\bar{1}2 >$ the tensile fracture strength was in the range of about 4 GPa. The surface-nucleated cracks propagated into the bulk along the $\{11\,\bar{1}\}$ cleavage plane, which is inclined by 19.5° to the normal of the free surface (see Fig. 6). According to the boundary conditions for SAWs, only the tensile opening stress σ_{11} is nonzero at the surface in the initial coordinate system.

This tensile stress of $\sigma_{11} = 4$ GPa can be represented by a set of orthogonal components in the coordinate system associated with the tilted cleavage plane $\{11\bar{1}\}$. The stresses in the new coordinate system are calculated by applying the transformation rule $\sigma_{ij} = A_{ik}A_{jl}\sigma_{kl}$, where A_{ik} is the corresponding rotation matrix around x_2. As in the initial coordinate system only σ_{11} has a non-zero value at the surface, we find the tensile stresses $\sigma_{11}^T(t) = \cos^2\varphi\sigma_{11}(t)$ and $\sigma_{33}^T(t) = \sin^2\varphi\sigma_{11}(t)$, and the shearing stress components $\sigma_{13}^T(t) = \sigma_{31}^T = -(1/2)\sin 2\varphi\sigma_{11}(t)$, where x_1^T is normal to the $\{11\bar{1}\}$ cleavage plane and φ is the angle of rotation around x_2. In the following estimate the time dependence will be omitted. The amplitudes of the calculated stress components were reduced according to the transformation law. The mean value of the predicted σ_{11} stress at the first fracture point is 4.0 GPa. The value of σ_{11}^T is equal to 3.6 GPa and can be considered as an estimate of the fracture strength of silicon in this special geometry (Kozhushko et al., 2007; Lomonosov & Hess, 2008).

In fact, a combination of mode I (tensile opening) and mode II (in-plane shearing or sliding) processes is expected to control this fracture geometry. The resulting stress components for a biaxial fracture mechanism in the tilted coordinate system are $\sigma_{11}^T = 3.6$ GPa and $\sigma_{31}^T = -1.3$ GPa. In addition, a smaller contribution from the component $\sigma_{33}^T = 0.4$ GPa has to be taken into account in a rigorous treatment. Fig. 8 illustrates the 'biaxial' fracture components with respect to the $\{11\,\bar{1}\}$ cleavage plane for this geometry (Kozhushko & Hess, 2010).

Silicon (223) plane

A very small tilt of the cleavage plane from the surface normal of 8° (see Fig. 6) generates a σ_{31}^T component acting as sliding mode of fracture along the $\{11\,\bar{1}\}$ cleavage plane with a value of $\sigma_{31}^T = -0.14\,\sigma_{11}$. Since fracture is a dynamic process, even such a relatively small shear stress may play an important role during crack extension. The other components are $\sigma_{11}^T = 0.98\,\sigma_{11}$ and $\sigma_{33}^T = 0.02\,\sigma_{11}$. A series of such experiments resulted in a critical fracture strength of only $\sigma_{11} \approx 3.0$ GPa in this geometry, the lowest critical opening stress found for the investigated geometries (Kozhushko & Hess, 2010).

Silicon (221) plane

In this fracture geometry the tilt of the cleavage plane is 35.3° to the surface normal (see Fig. 6). A laser pulse energy of about 70 mJ had to be applied to achieve fracture in the easy cracking direction, namely the Si(221)$< \bar{1}\bar{1}\,4>$ geometry. As the stress acting normal to the $\{11\bar{1}\}$ cleavage plane is reduced to 2/3 of the initial σ_{11} stress at the surface, all components can be easily obtained as described above as $\sigma_{11}^T = 0.67\,\sigma_{11}$, $\sigma_{31}^T = -0.48\,\sigma_{11}$ and $\sigma_{33}^T = 0.33\,\sigma_{11}$. It is noteworthy that nearly half of the initial tensile stress of $\sigma_{11} = 3.5$ GPa is transformed to an in-plane-shearing action in this case.

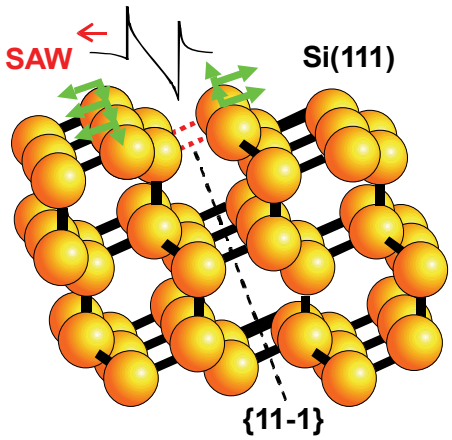


Fig. 8. Scheme of the crack nucleation process for the easy cracking geometry Si(111)< $\bar{1}\bar{1}$ 2> with biaxial crack components and propagation along the {11$\bar{1}$} cleavage plane, tilted by 19.5° to the surface normal.

The first crack was nucleated ~1 mm from the line source. At a distance of ~3 mm from the sources several fracture tracks generated a number of cracks located at larger distances from the source. In contrast to the previous geometries, these cracks consisted of a line along < 1$\bar{1}$0 > and two branches with an angle of about 70°. This angle is very close to the angle between the intersection lines of the {1$\bar{1}$1} and {$\bar{1}$11} cleavage planes with the surface, providing an example of crack bifurcation after nucleation. Numerical estimates show that these planes undergo fracture induced by the following stress components: $\sigma_{11}^T = 0.67\ \sigma_{11}$, $\sigma_{22}^T = 0.31\ \sigma_{11}$, $\sigma_{21}^T = 0.45\ \sigma_{11}$, $\sigma_{31}^T = 0.13\ \sigma_{11}$, $\sigma_{32}^T = 0.09\ \sigma_{11}$ and $\sigma_{33}^T = 0.03\ \sigma_{11}$. The opening stress component σ_{11}^T has the same value as the opening stress component of the {11$\bar{1}$} plane. These components bifurcate and draw cracks along the initial cleavage plane. Note that initial failure occurs along the line normal to the direction of the wave vector of the SAW pulse. This supports our basic assumption that tension of chemical bonds normal to the cleavage plane is strongly involved in the process of crack nucleation, while other stress components also influence the mechanical strength and furthermore may induce branching after nucleation or draw crack tips along other cleavage planes (Kozhushko & Hess, 2007; Kozhushko & Hess, 2010).

Silicon (110) plane

The geometry Si(110)< 1$\bar{1}$1 > was chosen because one plane of the {111}-cleavage-plane family is normal to the surface and the direction of the wave vector. The SAW solution indicates that the value of the out-of-plane shearing component σ_{21}, stretching the material normal to the sagittal plane, is not zero since in this geometry particle motion is no longer restricted to this plane.

It was demonstrated that a steep shock front is generated in the Si(110)< 1$\bar{1}$1 > geometry by transient SAW pulse evolution. Calculations of the stress field resulted in a positive σ_{11} peak, which means that the acting forces stretch bonds normal to the cleavage plane. Moreover, there are also displacements of particles along the x_2-coordinate axis, normal to the sagittal plane, which produce non-zero stress at the surface. The out-of-plane shearing component σ_{21} can be associated with fracture mode III and the σ_{22} stress stretches the material normal

to the sagittal plane. These contributions, however, are relatively small, since the analysis yields $\sigma_{11}/\sigma_{21} \approx 15$ and $\sigma_{11}/\sigma_{22} \approx 7$ for the peak stress components at the surface.
Dynamic fracture under the predominant action of biaxial stresses in the Si(110)< $1\bar{1}1$ > geometry was studied to further examine a more complicated mixed case. There is no difference for nonlinear pulse evolution of counterpropagating SAWs in this particular geometry. Fracture could be achieved only by increasing the laser pulse energy up to about 150 mJ and no extensive crack field could be observed. The average value found for the initial critical tensile stress σ_{11} was about 7 GPa. The transformation of the coordinate system of the SAW solution to the fracture geometry gives the following estimates of the peak stress components for the second plane normal to the free surface, namely { $\bar{1}11$ }, in comparison with the initial { $1\bar{1}1$ } cleavage plane: $\sigma_{11}^T = 0.2\ \sigma_{11}$, $\sigma_{21}^T = 0.22\ \sigma_{11}$ and $\sigma_{22}^T = 0.95\ \sigma_{11}$. These components are strong enough to branch the nucleated crack (Kozhushko & Hess, 2007; Kozhushko & Hess, 2010).

4.3 Comparison of fracture strength for different silicon geometries

The examples presented show that in the selected geometries SAWs generate dominating tensile stress at the surface, which is responsible for the nucleation of surface-breaking cracks. Several multi-mode fracture processes were characterized in anisotropic silicon. The technique provides values of the tensile stress between 3 and 7 GPa for the low-index-plane geometries responsible for nucleation and the corresponding tensile and shear stress components governing crack propagation into the bulk in those systems, where the weakest {111} cleavage plane deviates from the surface normal (see Table 1). The observed stresses of several gigapascals agree with the bulk strength of about 3-5 GPa estimated for undefined silicon geometries by longitudinal stress pulses (Wang et al., 2002; Gupta et al., 2003).

Geometry	σ_{11} (GPa)	σ_{11}^T (GPa)	σ_{31}^T (GPa)	σ_{33}^T (GPa)	σ_{22}^T (GPa)
Si(112)<-1-11>	4.5 GPa				
Si(223)<-3-34>	3.0 GPa	2.9 GPa	–0.42 GPa		
Si(111)<-1-12>	4.0 GPa	3.6 GPa	–1.3 GPa	0.4 GPa	
Si(221)<-1-14>	3.5 GPa	2.3 GPa	0.46 GPa	0.11 GPa	1.1 GPa
Si(110)<1-11>	7.0 GPa	1.4 GPa			6.7 GPa

Table 1. Cleavage geometries with tensile stress at the surface and tensile and shear stress components for the family of {111} cleavage planes deviating from the surface normal.

4.4 Fracture behaviour of silicon in mesoscopic and nanoscopic systems

With the extension of crystalline silicon devices and sensors to smaller and smaller sizes the dependence of the mechanical strength on the system size becomes an important issue. In applications of MEMS and NEMS devices, for example, the mechanical stability is essential for their manipulation, functionalization and integration into complexer systems. In general it is expected that the strength increases with decreasing size of the system due to the smaller number of crystal defects such as voids, microcracks or dislocations. In the early work on the fracture strength of silicon whiskers with diameters at the micrometer scale (~1–20 μm) tensile fracture strengths of 2–8 GPa were found. This is in the same range as the values measured here for well-defined test geometries. This may be interpreted by the assumption that similar fracture geometries and failure mechanisms were involved in these

processes. In more recent experiments using nanowires with diameters of 700 to 100 nm the strength increased from 0.03 to 2–4 GPa (Gordon et al., 2009). For the mechanical properties of self-welded [111] single-crystal silicon nanowire bridges, grown between two silicon posts, the maximum bending stress increased from 300 to 830 MPa for a wire diameter decreasing from 200 to 140 nm, depending on the loading conditions (Tabib-Azar et al., 2005). This means that at the micrometer scale the mechanical strength of the best silicon materials is comparable with the strength of wafers at the millimeter scale. This is consistent with the observation that size effects do not play a role on the elastic behaviour of silicon nanowires with a diameter >100 nm (Sohn et al., 2010).

In recent years, several techniques such as the chemical vapor deposition (CVD) vapor liquid solid (VLS) or CVD-VLS method have been developed to grow nanowires with diameters down to the few nanometer range. Currently, however, it is very difficult to extract general conclusions from this pioneering work, since contradictory results have been reported for the size effects of mechanical properties. It seems that the strength of silicon can increase to about 12 GPa, as the nanowire diameter decreases to 100–200 nm (Hoffmann et al., 2006) and 15–60 nm (Zhu et al., 2009) in wires grown along the [111] direction. This value comes already near to the theoretical strength for tensile cleavage of silicon along the {111} plane of 22 GPa obtained by ab initio calculations for an ideal silicon lattice (Roundy & Cohen, 2010). A similar value of 21 GPa has been reported by Dubois et al. in 2006. In Fig. 9 the strength values measured for macroscopic, mesoscopic and nanoscopic silicon systems are compared with ab initio theory of an ideal silicon crystal. It is interesting to note that for silicon the difference between the highest measured and ideal strength is only a factor of two, while it is 1-2 orders of magnitude for diamond (Hess, 2009).

Nanowires with a diameter below 20 nm can grow in the [111], [110] and [112] directions. For [110]-oriented nanowires with diameter <60 nm ductile failure has been observed (Han et al., 2007), while for [111]-oriented wires with diameters of 100–200 nm brittle failure occurred without plastic deformation (Heidelberg et al., 2006). These findings point to changes in the fracture behaviour at the nano-scale, which seem to be connected with the increasing surface/volume ratio and a smaller influence of defects.

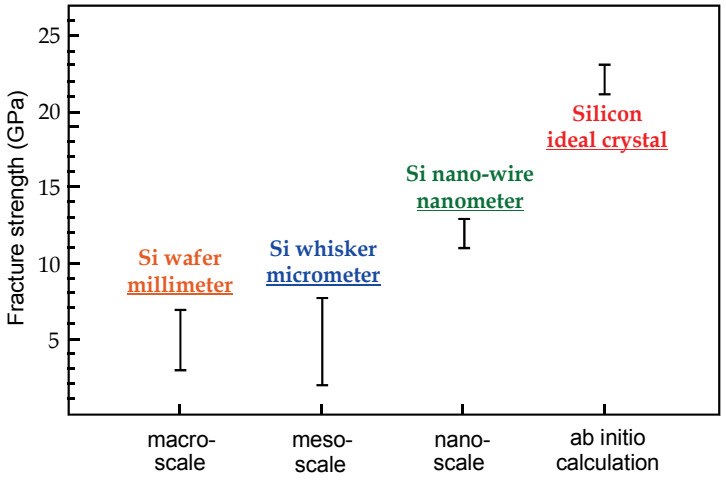


Fig. 9. Comparison of theoretical and experimental strength values for ideal and real silicon systems with decreasing size.

4.5 Molecular dynamics simulation of molecular silicon fracture mechanisms

Ab initio calculations are quite accurate but up to now provided only the brittle fracture strength of a few selected low-index failure configurations of silicon without detailed mechanistic information on the dynamics of the rearrangement of bonds. It is important to note that brittle fracture is a complicated multi-scale phenomenon involving nanoscopic and mesoscopic length scales. Atomic-scale crack growth proceeds via individual bond breaking events of variously oriented bonds depending on crack speed, which controls the individual propagation steps of fracture. These fundamental processes can be studied by simulations combining classical potentials and quantum mechanics, describing the stress fields and chemical rearrangements at the crack tip. For silicon such simulations have been performed giving new insight into the various possibilities of crack nucleation and extension. In the well studied case of cracking along the Si{111} cleavage plane, for example, it is possible to discriminate between clean continuous propagation of a crack along {111} by breaking six-member rings and discontinuous fracture by the formation of five- and seven-member rings in a recontruction process, as illustrated in Fig.10. In the bulk, plastic deformation along the dislocation glide is prohibited if the Peierls stress for the movement of nucleated dislocations is too high, as assumed for low temperatures (Kermode et al., 2008).

Recent simulations of the fracture mechanism in silicon nanowires by the modified embedded atom method (MEAM) potential indicate that cleavage is initiated by nucleation of a surface microcrack, while shear failure is initiated by the nucleation of a dislocation at the surface (Kang & Cai, 2010). Contrary to the situation in the bulk, failure seems to be controlled by the nucleation of dislocations and not by the dislocation mobility in these nano

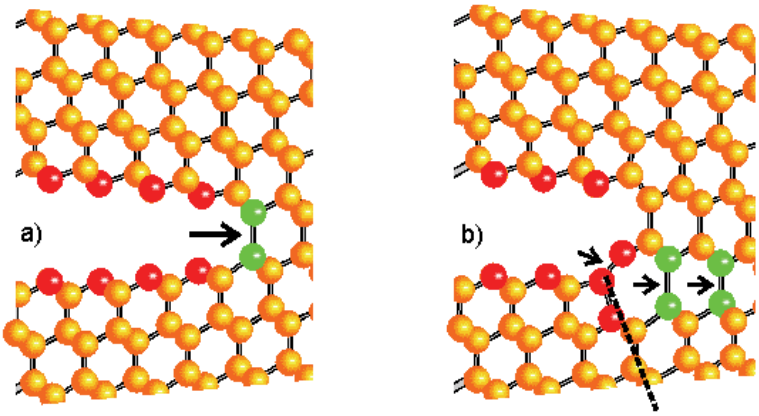


Fig. 10. Illustration of two different cracking modes of silicon: a) clean continuous crack propagation along {111} by breaking six-member rings and b) discontinuous fracture by the formation of five- and seven-member rings. The dashed line indicates the dislocation glide.

systems. It is interesting to note that nanowires with a diameter below 4 nm fail by shear processes at any temperature. For nanowires with a diameter >4 nm these simulations predict a fracture stress of 13 GPa for nanowires grown along the [110] direction and 15 GPa for the [111] direction, at a strain rate of 5×10^{8} s^{-1}. This indicates that for nanowires with a diameter in the nanometer range the fracture strength of [110] nanowires may be lower than that of the [111] nanowires, contrary to the behaviour of bulk silicon. One has to bear in

mind, however that the theoretical strain rates are many orders of magnitude higher than those of conventional fracture techniques, whereas for laser-induced fracture the difference in the strain rate is approximately two orders of magnitude. In fact, the higher strain rates used in the simulations could be the reason for the higher strength, as compared with the experimental value of about 12 GPa. While it is still difficult to compare molecular-dynamics simulations with experiments performed at the same time scale fracture of nanowires allows a comparison with theory at least on the same length scale.

5. Conclusions

Laser-induced stress pulses provide a novel efficient tool to elucidate the mode-specific spallation and delamination of films. Here the shock-pulse method measures the interface strength. In anisotropic crystals intrinsic nucleation at the surface by uni-axial and multi-axial fracture can be studied. The defined loading geometries realized with bulk and surface waves allow a straightforward interpretation in comparison with conventional methods. Therefore, the laser methods provide important insight into the relevant pure mode but also multi-mode failure processes. In fact, the critical stress or failure strength of spallation, delamination, fracture and cleavage can be determined in layered and anisotropic materials. For high-quality materials such as single-crystal silicon the measured critical stress components may be compared with *ab initio* calculations, performed for well-defined low-index geometries, to judge the mechanical quality of real materials. Owing to the high strain rates ($\sim 10^6\ s^{-1}$) involved, the influence of plastic deformations is reduced. This usually allows a more sensitive and accurate analysis of the mechanical strength, which is of increasing interest in thin-film technology and sensor device fabrication. The mechanical strength of nano-scale systems such as nanowires is currently under intense investigation. First interesting results on the mechanical response of MEMS structures have been reported (Kimberley et al., 2008). Important mechanistic insight into the atomic-scale processes during crack nucleation and propagation is provided by molecular dynamics simulations.

6. References

Cuq-Lelandais, J.-P.; Boustie, M.; Berthe, L.; de Rességuier, T.; Combis, P.; Colombier, J. P.; Nivard, M. & Claverie, A. (2009). Spallation generated by femtosecond laser driven shocks in thin metallic targets: *J. Phys. D: Appl. Phys.*, 42, pp. 065402-1–10

Dubois, S. M.-M.; Rignanese, G.-M.; Pardoen, T. & Charlier, J.-C. (2006). Ideal strength of silicon: An ab initio study: *Phys. Rev. B*, 74, pp. 235203-1–7

Eckl, C.; Kovalev, A. S.; Mayer, A. P.; Lomonosov, A. M. & Hess, P. (2004). Solitary surface acoustic waves: *Phys. Rev. E*, 70, pp. 046604-1–15

Gordon, M. J.; Baron, T.; Dhalluin, F.; Gentile, P. & Ferret, P. (2009). Size effects in mechanical deformation and fracture of cantilevered silicon nanowires: *Nano Lett.*, 9, pp. 525–529

Gupta, V.; Argon, A. S.; Cornie, J. A. & Parks, D. M. (1990). Measurement of interface strength by laser-pulse-induced spallation: *Mater. Sci. Eng. A*, 126, pp. 105–117

Gupta, V.; Kireev, V.; Tian, J.; Yoshida, H. & Akahoshi, H. (2003). Glass-modified stress waves for adhesion measurement of ultrathin films for device applications: *J. Mech. Phys. Solids*, 51, pp. 1395–1412

Han, X.; Zheng, K.; Zhang, Y.; Zhang, X.; Zhang, Z. & Wang, Z. L. (2007). Low- temperature in situ large-strain plasticity of silicon nanowires: *Adv. Mater.*, 19, 2112–2118

Heidelberg, A.; Ngo, L. T.; Wu, B.; Phillips, M. A.; Sharma, S.; Kamins, T. I.; Sader, J. E. & Boland, J. J. (2006). A generalized description of the elastic properties of nanowires: *Nano Lett.*, 6, pp. 1101–1106

Hess, P. (2009). Determinetion of linear and nonlinear mechanical properties of diamond by laser-based acoustic waves: *Diamond Relat. Mater.*, 18, pp. 186-190

Hess, P. & Lomonosov, A. M. (2010). Solitary surface acoustic waves and bulk solitons in nanosecond and picosecond laser ultrasonics: *Ultrasonics*, 50, pp. 167–171

Hoffmann, S.; Utke, I.; Moser, B.; Michler, J.; Christiansen, S. H.; Schmidt, V.; Senz, S.; Werner, P.; Gösele, U. & Ballif, C. (2006). Measurement of the bending strength of vapor-liquid-solid grown silicon nanowires: *Nano Lett.*, 6, pp. 622–625

Hu, L. & Wang, J. (2006). Pure-shear failure of thin films by laser-induced shear waves: *Exp. Mech.*, 46, pp. 637–645

Hu, L.; Zhang, X.; Miller, P.; Ozkan, M.; Ozkan, C. & Wang, J. (2006). Cell adhesion measurement by laser-induced stress waves: *J. Appl. Phys.*, 100, pp. 084701-1–5

Hu, L.; Miller, P. & Wang, J. (2009). High strain-rate spallation and fracture of tungsten by laser-induced stress waves: *Mater. Sci. Eng. A*, 504, pp. 73–80

Ikeda, R.; Cho, H.; Sawabe, A. & Takemoto, M. (2005). Laser spallation method to measure strength against mode-I decohesion of CVD diamond films: *Diamond Relat. Mater.*, 14, pp. 631–636

Kandula, S. S. V.; Tran, P.; Geubelle, P. H. & Sottos, N. R. (2008a). Dynamic delamination of patterned thin films: *Appl. Phys. Lett.*, 93, pp. 261902-1–3

Kandula, S. S. V.; Hartfield, C. D.; Geubelle, P. H. & Sottos, N. R. (2008b). Adhesion strength measurement of polymer dielectric interfaces using laser spallation technique: *Thin Solid Films*, 516, pp. 7627–7635

Kanel, G. I. (2010). Spall fracture methodological aspects, mechanisms and governing factors: *Int. J. Fract.* DOI: 10.1007/s10704-009-9438-0

Kang, K. & Cai, W. (2010). Size and temperature effects on the fracture mechanisms of silicon nanowires: Molecular dynamics simulations: *Int. J. Plasticity*, DOI: 10.1016/j.ijplas. 2010 .02.001

Kermode, J. R.; Albaret, T.; Sherman, D.; Bernstein, N.; Gumbsch, P.; Payne, M. C.; Csányi, G. & De Vita, A. (2008). Low-speed fracture instabilities in a brittle crystal: *Nature*, 455, pp. 1224–1227

Kimberley, J.; Chasiotis, I. & Lambros, J. (2008). Failure of microelectromechanical systems subjected to impulsive load: *Int. J. of Solids and Struct.*, 45, pp. 497–512

Kitey, R.; Geubelle, P. H. & Sottos, N. R. (2009). Mixed-mode interfacial adhesive strength of a thin film on an anisotropic substrate: *J. Mech. Phys. Solids*, 57, pp. 51–66

Kolomenskii, A. A.; Lomonosov, A. M.; Kuschnereit, R.; Hess, P. & Gusev, V. E. (1997). Laser generation and detection of strongly nonlinear elastic surface pulses: *Phys. Rev. Lett.*, 79, pp. 1325–1328

Kozhushko, V. V. & Hess, P. (2007). Anisotropy of the strength of Si studied by a laser-based contact-free method: *Phys. Rev. B*, 76, pp. 144105-1–11

Kozhushko, V. V. & Hess, P. (2008). Nonlinear surface acoustic waves: silicon strength in phonon-focusing directions: *Ultrasonics*, 48, pp. 488–491

Kozhushko, V. V. & Hess, P. (2010). Comparison of mode-resolved fracture strength of silicon with mixed-mode failure of diamond crystals: *Eng. Fract. Mech.*, 77, pp. 193–200

Kozhushko, V. V.; Lomonosov, A. M. & Hess, P. (2007). Intrinsic strength of silicon crystals in pure- and combined-mode fracture without precrack, *Phys. Rev. Lett.*, 98, pp. 195505-1–4

Lacombe, R. H. (2006). *Adhesion Measurement Methods: Theory and Practice,* CRC Press, ISBN 0-8247-5361-5, Boca Raton, FL

Lehmann, G.; Lomonosov, A. M.; Hess, P. & Gumbsch, P. (2003). Impulsive fracture of fused quartz and silicon crystals by nonlinear surface acoustic waves: *J. Appl. Phys.*, 94, pp. 2907–2914

Lomonosov, A. M. & Hess, P. (2002). Impulsive fracture of silicon by elastic surface pulses with shocks: *Phys. Rev. Lett.*, 89, pp. 095501-1–4

Lomonosov, A. M., Hess, P. & Mayer, A. P. (2002). Observation of solitary elastic surface pulses: *Phys. Rev. Lett*., 88, pp. 076104-1–4

Lomonosov, A. M.; Mayer, A. P. & Hess, P. (2001). Laser-based surface acoustic waves in materials science, In: *Modern Acoustical Techniques for the Measurement of Mechanical Properties*, Levy, M.; Bass, H. E. & Stern R. (Eds.), pp. 65-134, Academic, ISBN 0-12-475786-6, San Diego, CA

Lomonosov, A. M. & Hess, P. (2008). Nonlinear surface acoustic waves: Realization of solitary pulses and fracture: *Ultrasonics*, 48, pp. 482-487

Miller, P.; Hu, L. & Wang, J. (2010). Finite element simulation of cell-substrate decohesion by laser-induced stress waves: *Mech. Behav. Biomed. Mater.*, 3, pp. 268–277

Pronin, A. & Gupta, V. (1993). Interferometry on diffuse surfaces in high-velocity measurements: *Rev. Sci. Instrum*., 64, pp. 2233–2236

De Rességuier, T.; Cuq-Lelandais, J.-P.; Boustie, M.; Lescoute, E. & Berthe, L. (2010). Wave propagation and dynamic fracture in laser shock-loaded solid materials, In: *Wave Propagation in Materials for Modern Applications*, Petrin, A. (Ed.), pp. 419-436, INTECH, ISBN 978-953-7619-65-7, Croatia

Roundy, D. & Cohen, M. L. (2001). Ideal strength of diamond, Si, and Ge: *Phys. Rev. B*, 64, pp. 212103-1–3

Sohn, Y.-S.; Park, J.; Yoon, G.; Song, J.; Jee S.-W.; Lee, J.-H.; Na, S.; Kwon, T. & Eom, K. (2010). Mechanical properties of silicon nanowires: *Nanoscale Res. Lett*., 5, pp. 211-216

Tabib-Azar, M.; Nassirou, M.; Wang, R.; Sharma, S.; Kamins, T. I.; Islam, M. S. & Williams, R. S. (2005). Mechanical properties of self-welded silicon nanobridges: *Appl. Phys. Lett*., 87, pp. 113102-1–3

Tamura, H.; Kohama, T.; Kondo, K. & Yoshida M. (2001). Femtosecond-laser-induced spallation in aluminum: *J. Appl. Phys.*, 89, pp. 3520–3522

Wang, J.; Weaver, R. L. & Sottos, N. R. (2002). A parametric study of laser induced thin film spallation: *Exp. Mech*., 42, pp. 74–83

Wang, J.; Weaver, R. L. & Sottos, N. R. (2003a). Laser-induced decompression shock development in fused silica: *J. Appl. Phys.*, 93, pp. 9529–9536

Wang, J.; Sottos, N. R. & Weaver, R. L. (2003b). Mixed-mode failure of thin films using laser-generated shear waves: *Exp. Mech.*, 43, pp. 323–330

Wang, J.; Sottos, N. R. & Weaver, R. L. (2004). Tensile and mixed-mode strength of a thin film-substrate interface under laser-induced pulse loading: *J. Mech. Phys. Solids*, 52, pp. 999–1022

Wei, Y. & Hutchinson, J. W. (1998). Interface strength, work of adhesion and plasticity in the peel test*: Int. J. Fract.,* 93, pp. 315–333

Zhu, Y.; Xu, F.; Qin, Q.; Fung, W. Y. & Lu, W. (2009). Mechanical properties of vapor–liquid–solid synthesized silicon nanowires: *Nano Lett*., 9, pp. 3934–3939

18

Ultrasonics: A Technique of Material Characterization

Dharmendra Kumar Pandey[1] and Shri Pandey[2]
[1]*Department of Physics, P.P.N. (P.G.) college, Kanpur-208 001, U.P.,*
[2]*Department of Physics, University of Allahabad, Allahabad- 211 002, U.P., India*

1. Introduction

The material science and characterization is a field concerned with inventing new materials and improving previously known materials by developing a deeper understanding of properties under different physical conditions. The properties of materials depend upon their composition, structure, synthesis and processing. Many properties of materials depend strongly on the structure, even if the composition of the material remains same. This is why the structure-property or microstructure property relationships in materials are extremely important.

On the basis of different physical properties, the materials are classified mainly into five categories: (a) metals and alloys, (b) semi-metals and semiconductors, (c) ceramics, glasses and glass-ceramics, (d) polymers, and (e) composite materials. Functional classification of materials includes aerospace, biomedical, electronic, energy and environmental, magnetic, and optical (photonic) materials. The structural classification of materials are of two types as (a) crystalline (single crystal and polycrystalline), and (b) amorphous.

The selection of a material and the potential to be manufactured economically and safely into useful product is a complicated process. It requires the complete knowledge of constituent material not only after production but also in processing. Increased competition and need of higher productivity and better products from material producing industries are creating more stringent requirements for process and quality control. This demands the characterization of materials. The topic material characterization essentially includes the evaluation of elastic behaviour, material microstructure and morphological features, associated mechanical properties etc. The destructive, semi-destructive and non-destructive testing (DT & NDT) techniques are available for the complete characterization. These characterization techniques are the basic tool for the quality control and quality assurance of the material or component or product.

Ultrasonics, which is a sub category of acoustics deals with acoustics beyond the audio limit. The application of ultrasonics falls into two categories as high frequency- low intensity and low frequency – high intensity. The low intensity application carries the purpose of simply transmitting energy through the medium in order to obtain the information about the medium or to convey information through the medium. High intensity application deliberately affects the propagation medium or its contents. So, the low intensity and high

intensity application of ultrasonic wave belongs in non-destructive and destructive techniques of characterization respectively.
The quantities, ultrasonic velocity and attenuation are the important parameters, which are required for the ultrasonic non-destructive technique of material characterization. The ultrasonic velocity is related to the elastic constants and density of material. Hence, it gives the information about the mechanical, anisotropic and elastic properties of medium through it passes. It is also important in low temperature physics because it is involved in the evaluation of Debye average velocity and Debye temperature. Ultrasonic velocity in nanofluid depends on the concentration of nano-particles of material dispersed in polymer matrix, thus it is not only important at bulk scale but also at nanoscale. When the ultrasonic wave propagates through the medium, its some part of energy is attenuated through the different mechanism like thermal loss, scattering, absorption, electron-phonon interaction, phonon-phonon interaction, and magnon-phonon interaction etc., called as ultrasonic attenuation. The coefficient of ultrasonic attenuation correlates several physical properties like elastic constants, guruneisen parameter, thermal conductivity, thermal relaxation time, acoustic coupling constant, thermal energy density, specific heat, particle size, density, Debye average velocity, and concentration etc. Thus, the material can be characterized with the knowledge of ultrasonic parameters under different physical conditions.
Normally, the ultrasonic NDT of material characterization are used for the determination of (a) elastic constants (Shear modulus, Bulk modulus, Young modulus and lame modulus), (b) microstructure (grain size, texture, density etc.), (c) discontinuity (porosity, creep damage, fatigue damage etc.), and mechanical properties (tensile strength, shear strength, hardness etc.). The new work in this field also provides the characterization of advanced and smart materials like GMR etc. Now a day, the synthesis and characterization of nanomaterials and nanofluids are also in touch of ultrasonic NDT&E.
In this chapter, ultrasonic material property characterization has been considered. Initially, it covers information about the ultrasonic wave, its mode of propagation and characteristic properties. After this, a brief study of ultrasonic velocity and attenuation in solid has been discussed, which covers the theoretical evaluation and experimental measurements of these ultrasonic parameters. Later on, the characterization of different material (metals, alloys, platinum group metals, nanomaterials, nanofluid, semiconductor etc) has been discussed on the basis of these ultrasonic quantities and related parameters.

2. Ultrasonic wave

As a sub category of acoustics, ultrasonics deals with the acoustics above the human hearing range (the audio frequency limit) of 20 kHz. Unlike audible sound waves, the ultrasonic waves are not sensed by human ear due to the limitations on the reception of vibrations of high frequency and energies by the membrane. Ultrasonic wave exhibits all the characteristic properties of sound. Ultrasonic vibrations travel in the form of wave, similar to the way light travels. However, unlike light waves, which can travel in vacuum, ultrasonic wave requires elastic medium such as a liquid or a solid. The wavelength of this wave changes from one medium to another medium due to the elastic properties and induced particle vibrations in the medium. This wave can be reflected off with very small surfaces due to having much shorter wavelength. It is the property that makes ultrasound useful for the non-destructive characterization/testing of materials. The knowledge of generation/detection of ultrasonic wave and its characteristics is important for its precise and suitable application.

2.1 Sources of ultrasonic wave

The ultrasonic wave (UW) can be generated with the mechanical, electrostatic, electrodynamic, electromagnetic, magnetostrictive effect, piezoelectric effect, and laser methods.

Mechanical method or Galton Whistle method is an initial method for the generation of ultrasonic wave. This uses mechanical shock or friction for the generation of wave in frequency range of 100 kHz to 1 MHz. A high frequency of ultrasonic wave (10 to 200 MHz) can be generated using electrostatic method. The magneto inductive effect is used in electrodynamic method for the production of ultrasound. The mechanical deformation in ferromagnetic material in presence of magnetic field is called as magnetostriction. This phenomenon is most pronounced in metals such as nickel, iron, cobalt and their alloys. Magnetostriction effect is used for generation of ultrasonic wave in magnetostrictive effect method. Most common method for generation of ultrasound is the Piezoelectric effect method. In this method, inverse Piezoelectric effect is used for generation of UW. When a laser light incident on the surface of suitable material, its some portion of energy is absorbed at the surface with in the skin depth and rest get reflected. The absorbed energy produces tangential stress and then bulk strain through transient surface heating; as a result UW is produced in concerned medium.

2.2 Transducers for ultrasonic wave

The device that converts one form of energy to another form is called as transducer. An ultrasonic transducer converts electrical energy to mechanical energy, in the form of sound, and vice versa. The main components are the active element, backing, and wear plate (Fig.1).

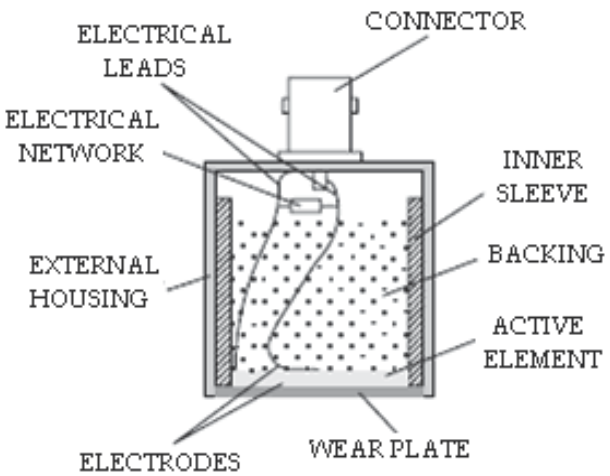


Fig. 1. Basic figure of an ultrasonic transducer

a. The Active Element

The active element, which is piezo or ferroelectric material, converts electrical energy such as an excitation pulse from a flaw detector into ultrasonic energy. The most commonly used materials are polarized ceramics which can be cut in a variety of manners to produce different

wave modes. New materials such as piezo polymers and composites are also being employed for applications where they provide benefit to transducer and system performance.

b. Backing

The backing is usually a highly attenuative, high density material that is used to control the vibration of the transducer by absorbing the energy radiating from the back face of the active element. When the acoustic impedance of the backing matches with the acoustic impedance of the active element, the result will be a heavily damped transducer that displays good range resolution but may be lower in signal amplitude. If there is a mismatch in acoustic impedance between the element and the backing, more sound energy will be reflected forward into the test material. The end result is a transducer that is lower in resolution due to a longer waveform duration, but may be higher in signal amplitude or greater in sensitivity.

c. Wear Plate

The basic purpose of the transducer wear plate is to protect the transducer element from the testing environment. In the case of contact transducers, the wear plate must be a durable and corrosion resistant material in order to withstand the wear caused by use on materials such as steel. For immersion, angle beam, and delay line transducers, the wear plate has the additional purpose of serving as an acoustic transformer between the high acoustic impedance of the active element and the water.

Now a days, following type of transducers are in use for different applications.

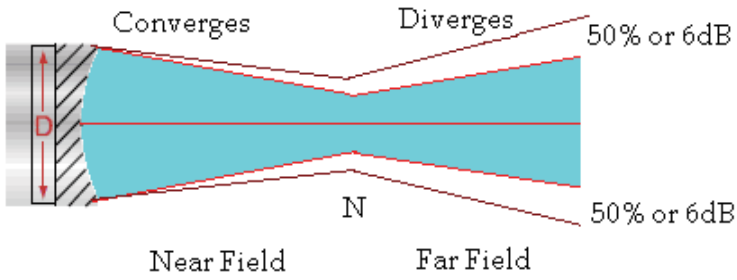
1. Normal beam or single element or delay line transducer
2. Dual element transducer
3. Angle beam transducer
4. Immersion transducer
5. Mechanical focous transducer
6. Electronic time delay focouing or array transducer
7. Capacitive transducer

In most of applications piezoelectric transducers are used for generating and receiving the ultrasonic waves.

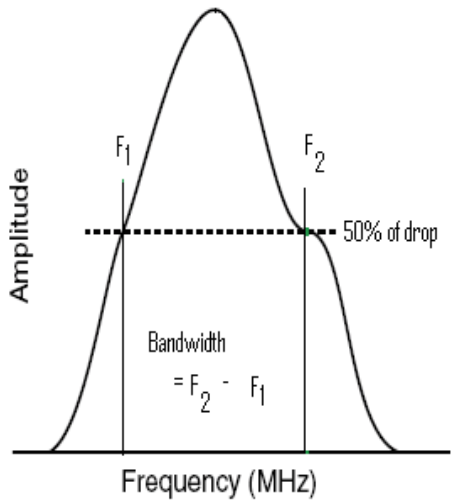
2.3 Characteristics of ultrasonic wave

For the appropriate choice of ultrasonic wave with suitable frequency and intensity, the knowledge of some essential parameters related to transducer is important. The characteristic parameters of ultrasonic wave are:

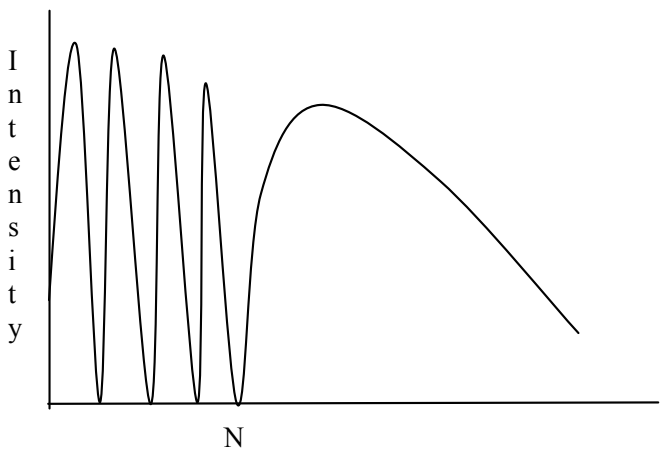
1. Sound Field (Near field and far field): The sound field of a transducer is divided in two two zones; the near field region or Fresnel zone and far field region or Fraunhofer zone. In the near field region the ultrasonic beam converges and in the far field it diverges. The near field is the region directly in front of transducer where echo amplitudes goes through a series of maxima and minima and ends at the last maximum, at the distance N ($N = D^2v/4C = D^2/4\lambda$; N: near field distance, D: Element diameter, v: frequency, c: material sound velocity, and λ: wavelength) from the transducer (Fig.2). The intensity variation along and across the axial distance up to near field region is approximately constant and after which it decreases. The beam boundary defines the limits of the beam to the point where the disturbance ceases to exist or falls below the threshold value. The beam intensity at the boundary is reduces to one half (6dB) of the intensity at the beam axis (Fig.2).



(a)



(b)



(c)

Fig. 2. (a) sound field of transducer, (b) amplitude verses frequency of UW, (c) intensity of UW verses axial distance from transducer.

2. Focal zone: The starting and ending points of the focal zone on axis of transducer are located where pulse echo signal amplitude drops to -6dB of the amplitude at the focal point. If Z_B and Z_E are the beginning and end of the focal zone from the transducer then, then focal zone will be difference of them. The length of focal zone (F_Z) is equal to $NS_F^2[2/(1+0.5S_F)]$; where S_F: Normalized focal length=F/N, F: focal length, N: near field distance.
3. Beam diameter: It is a parameter, which defines the transducers sensitivity. Smaller the beam diameter, the greater amount of energy is reflected by the flaw. At -6dB drop of intensity, the beam diameter (BD) at the focus is equal to $0.2568DS_F$ or 1.02FC/vD. For the flat transducer, normalized focal length have value one. The Fig.3 represents the clear picture of focal zone and beam diameter.

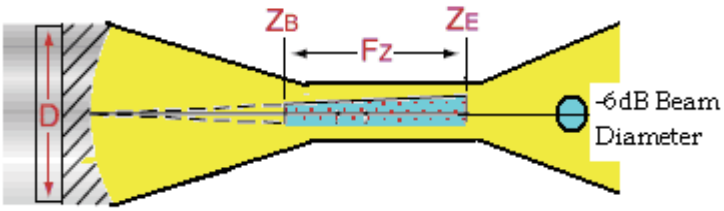


Fig. 3. Focal zone of transducer and beam diameter

4. Beam spread or half angle: The spreading of ultrasonic beam always take place as the wave travel from the transducer. In the near field, the beam has a complex shape that narrows, while in the far field it diverges. The divergence angle or beam spread angle (θ) is equal to $\sin^{-1}(K\lambda/D)$ or $\sin^{-1}(KC/\nu D)$. Where K is a constant which depends on shape of transducer, edge of beam and method used to determine the beam spread.It is clear that beam spread from a transducer can be reduced by selecting a transducer with higher frequency or larger element diameter or both. Fig.4 shows a simplistic understanding of beam spread angle.

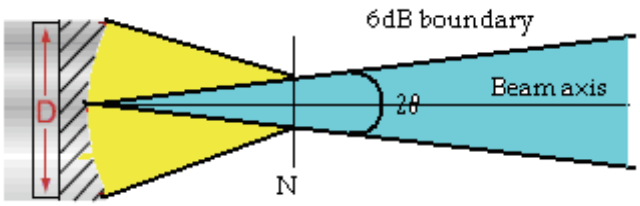


Fig. 4. Ultrasonic beam divergence and angle of divergence

2.4 Detection of ultrasonic wave

There are various methods for the detection of UW. The methods are based on the principle of piezoelectric, electrostatic and magnetostriction effects. The classical methods like mechanical and optical methods are also used for the detection of UW. Normally the devices based on piezoelectric effect are used commercially for the detection of UW, these devices comes in electrical method of detection.

The prime division of detection in electrical method are the Interferometer or continuous wave (CW) and Pulse technique (PT) methods. In CW method, the UW generated by the

source is passed through the concerned medium/specimen, which is reflected from the reflecting plate. The reflecting plate is adjusted towards the source such that the current in the oscillator of the source changes periodically in maxima or minima. The maximum in current corresponds to the half of wavelength interval due to the formation of standing wave between source and plate. This method is preferred in low frequency region for the measurements of ultrasonic parameters in liquids.

The Pulse technique is utilized for detection or measurement of transit time in both liquids and solids. It uses piezoelectric transducer with and without delay lines for the production and detection of UW. In this method short duration electric pulses generates the UW with the broadband piezoelectric transducer. The generated longitudinal or shear wave are transmitted to the specimen. The reflected wave or echo by the medium are detected by the transducer on the principle of direct piezoelectric effect and echo pattern is obtained. Using this pattern, the exact transit time needed for a signal to travel between the front and back surface of the specimen or concerned medium is determined, that is used for determination of ultrasonic velocity and attenuation. The different pulse techniques for precise measurements or detection of UW are Sing around, Pulse superposition, Pulse echo overlap, Cross-correlation, Phase slope and Pulse transmission method. Hydrophones are also piezoelectric transducer that generates electrical signal when subjected to pressure change or UW under water. It can detect UW in air, but will be less sensitive due to its design as having a good acoustic impedance match with water.

3. Material characterization techniques (NDT & DT)

The two major classification of material characterization technique are non-destructive testing (NDT) and destructive testing (DT). Under destructive technique (such as: tensile testing, creep testing, impact testing, torsion testing, hardness testing etc.) of characterization the tested material or product can not be used again. The destruction of test object usually makes this type of test more costly. Non-destructive testing technique is a specific procedure whereby the service ability of materials or components is not impaired by testing process. The various methods like visual testing, liquid penetrant testing, magnetic particle testing, eddy current testing, radiographic testing, ultrasonic testing, leak testing, thermography and neutron radiography are the NDT technique of material characterization. Among the various non-destructive testing and evalution (NDT&E) plays a key role in material characterization.Ultrasonic properties provide important diagnostic for microstructural properties as well as deformation processes in a material, controlling material behaviour based on the physical mechanism to predict future performance of the materials.

4. Classification of ultrasonic application and testing

The ultrasonic testing involves both the low intensity and high intensity ultrasonic wave for the characterization, that belongs in non-destructive and destructive techniques of characterization respectively. Uses of high intensity and low frequency ultrasonic wave includes medical therapy and surgery, atomization of liquids, machining of materials, cleaning and wielding of plastics and metals, disruption of biological cells, and homogenization of materials. The low intensity and high frequency ultrasonic waves are applied for medical diagnosis, acoustical holography, material characterization etc. The low

intensity ultrasoud measurements provides a good diagnosis of material property and process control in industrial apllication (Alers, 1965; Green,1973; Lowrance, 1975; Renolds, 1978; Teagle, 1983; Smith, 1987; Varry, 1987; Thompson,1996; Jayakumar, 1998; Kumar, 2001; Raj, 2003; Roth, 2003; Blodgett, 2005).

5. Ultrasonic NDT as a material characterization

There are four mode of propagation by which an ultrasonic wave can propagate in a medium, as: longitudinal or compressnal wave, transverse or shear wave, surface or Rayleigh wave and plate or lamb wave. The most common methods of ultrasonic examination utilize the longitudinal waves or shear waves.

Ultrasonic velocity or attenuation are the parameters that correlate to structural inhomogenities or flaw size atomistic (interstitials), elastic parameters, precipitates, dislocations, ordering of molecules in liquid crystals, phase transformations, porosity and cracks, concentration of different components of alloys or mixed crystal system, vacancies in lattice sites, size of the nanoparticles in nano-structured materials, electrical resistivity, specific heat, thermal conductivity and other thermophysical properties of the materials depending upon the different physical conditions like temperature, pressure, crystallographic orientation, magnetization etc. Thus, ultrasonic study of a material provides information about elastic constants, microstructure, discountinuty, and mechanical properties under different condition.

5.1 Ultrasonic velocity

On the basis of mode of propagation there are four types of ultrasonic velocities, as longitudinal, shear, surface and lamb wave velocity. Longitudinal and shear wave velocities are more important for the material characterization because they are well related to elastic constants and density. However, it is independent of frequency of wave and dimension of the given material. The mechanical behaviour and anisotropic properties of the material can be well defined on the knowledge of ultrasonic velocity. The mathematical formulations and measurement techniques for ultrasonic velocity are detailed in following heads.

5.1 A Ultrasonic velocity, related parameters and its theoretical evaluation

The mechanical properties of the solids differ from those of fluids in two important respects. Firstly, greater binding forces exist between their constituent atoms so that they support shear stress. Secondly, anisotropy may occur, especially in single crystal, in which the atoms form regular lattice. The velocity of ultrasonic wave of any kind can be determined from the elastic moduli (Y: Young's modulus, G: modulus of rigidity, and σ: poisson's ratio) and density (d) of the material. The logngitudinal and shear wave velocities (V_L and V_S) can be determined with following expressions.

$$V_L = \left[\frac{Y(1-\sigma)}{d(1+\sigma)(1-2\sigma)}\right]^{1/2} \text{ and } V_S = \left[\frac{Y}{2d(1+\sigma)}\right]^{1/2} = \left[\frac{G}{d}\right]^{1/2} \quad (1)$$

In terms of lame's moduli (λ and μ) , the ultrasonic velocities can be expressed as;

$$V_L = \left[\frac{\lambda + 2\mu}{d}\right]^{1/2} \text{ and } V_S = \left[\frac{\mu}{d}\right]^{1/2} \quad (2)$$

The stress strain relationships for anisotropic crystals vary with the direction. Thus velocity of ultrasonic wave varies with the direction of propagation of wave and mode of polarization. There are three type of ultrasonic velocity (one longitudinal and two shear wave) for each direction of propagation of wave in cubic (Mason, 1958; Singhal 2003) and hexagonal structured materials (Mason,1969; Alers,1958; Rosen,1970; Yadawa,2009). The expressions for the velocities are given in Table (1) and Table (2). In Tables 1-2, the V_1 is longitudinal and V_2 & V_3 are the shear wave velocities of ultrasonic wave. The C_{11}, C_{12}, C_{44}, C_{33} and C_{66} are the second order elastic constants.

The Debye theory of specific heat has proven its usefulness because it is a single –parameter theory which describes the observation remarkebly well. Its one parameter, Debye temperature (T_D) need not to be determined by any heat capacity measurements but can be calculated from the elastic moduli. Once this parameter has been determined from the elastic moduli, the Debye theory specifies the lattice contribution to the specific heat only to an accuracy of about 10 or 20% over most of temperature range. Because of this, the theoratical model assumes the solid to be an elastic continuum in which all sound waves travel at the same velocity independent of their wavelength. This model is satisfactory only in the limit of long wavelengths or low temperatures. The expression for the T_D can be given as:

$$T_D = \frac{\hbar \, V_D \, (6\pi^2 \, n_a)^{1/3}}{K_B} \quad (3)$$

Here, $\hbar$ is quantum of action and is equal to Planck's constant divided by 2π ; K_B is Boltzmann Constant; n_a is atom concentration. This Debye average velocity is important

Direction of propagation	Direction of polarization	Type of wave	Velocity expression	Velocity notation
100	100	Long.	$(C_{11}/d)^{1/2}$	$V_1=V_L$
	010	Shear	$(C_{44}/d)^{1/2}$	$V_2=V_{S1}$
	001	Shear	$(C_{44}/d)^{1/2}$	$V_3=V_{S2}$
110	110	Long.	$((C_{11}+C_{12}+2C_{44})/2d)^{1/2}$	$V_1=V_L$
	001	Shear	$(C_{44}/d)^{1/2}$	$V_2=V_{S1}$
	$1\bar{1}0$	Shear	$((C_{11}-C_{12})/2d)^{1/2}$	$V_3=V_{S2}$
111	111	Long.	$((C_{11}+2C_{12}+4C_{44})/3d)^{1/2}$	$V_1=V_L$
	Any direction in 111 plane	Shear	$((C_{11}-C_{12}+C_{44})/3d)^{1/2}$	$V_2= V_3$ $V_{S1}=V_{S2}$

Table 1. Ultrasonic velocities for cubic structured materials

Direction of propagation	Direction of polarization	Type of wave	Velocity expression	Velocity notation
001 (Along unique axis or z-axis)	001	Long.	$\left(C_{33}/d\right)^{1/2}$	$V_1=V_L$
	Any direction in 001 plane	Shear	$\left(C_{44}/d\right)^{1/2}$	$V_2=V_3$ $V_{S1}=V_{S2}$
100 (or any other direction perpendicular to 001)	100	Long.	$\left(C_{11}/2d\right)^{1/2}$	$V_1=V_L$
	001	Shear	$\left(C_{44}/d\right)^{1/2}$	$V_2=V_{S1}$
	010	Shear	$\left((C_{11}-C_{12})/2d\right)^{1/2}$	$V_3=V_{S2}$
At angle θ with the unique axis of the crystal		Long.	$[\{C_{33}Cos^2\theta+C_{11}Sin^2\theta+C_{44}+\{[C_{11}Sin^2\theta-C_{33}Cos^2\theta+C_{44}(Cos^2\theta-Sin^2\theta)]^2+4\,Cos^2\theta\,Sin^2\theta\,(C_{13}+C_{44})^2\}^{1/2}\}/2d]^{1/2}$	$V_1=V_L$
		Shear	$[\{C_{33}Cos^2\theta+C_{11}Sin^2\theta+C_{44}-\{[C_{11}Sin^2\theta-C_{33}Cos^2\theta+C_{44}(Cos^2\theta-Sin^2\theta)]^2+4\,Cos^2\theta\,Sin^2\theta\,(C_{13}+C_{44})^2\}^{1/2}\}/2d]^{1/2}$	$V_2=V_{S1}$
		Shear	$[\{C_{44}Cos^2\theta+C_{66}Sin^2\theta\}/\,d]^{1/2}$	$V_3=V_{S2}$

Table 2. Ultrasonic velocities for hexagonal structured materials

parameter in the low temperature physics because it is related to elastic constants through ultrasonic velocities. The Debye average velocity (V_D) in the materials is calculated using the following equation (Oligschleger, 1996).

$$V_D=\left(\frac{1}{3}\sum_{i=1}^{3}\int\frac{1}{V_i^3}\frac{d\Omega}{4\pi}\right)^{-1/3} \tag{4}$$

Here the integration is over all directions and summation is over the type of ultrasonic velocities. Along the [100], [111] (for cubic crystal) and [001] (for hexagonal structured crystals) direction of propagation of wave, the equation (4) reduces as:

$$V_D=\left[\frac{1}{3}\left(\frac{1}{V_1^3}+\frac{2}{V_2^3}\right)\right]^{-1/3} \tag{4a}$$

and along the [110] (for cubic) and any angle with the unique axis of hexagonal structured crystal, direction of propagation, the equation (4) reduces as:

$$V_D = \left[\frac{1}{3} \left(\frac{1}{V_1^3} + \frac{1}{V_2^3} + \frac{1}{V_3^3} \right) \right]^{-1/3} \quad (4b)$$

On the knowlewdge of elastic constants, the theoretical evaluation of ultrasonic velocity and Debye average velocity in cubic and hexagonal structured materials can be done with help of expressions written in Table (1), Table (2) and equations (4a)-(4b). There are several theories (Ghate,1965; Mori,1978; Rao,1974; Yadav AK, 2008) for the calculation of elastic constants. The elastic constants depend on the lattice parameters of structured materials. The elastic constants and elastic moduli can be calculated with the knowledge of lattice parameters.

5.1 B Measurement techniques of ultrasonic velocity

The study of the propagation of ultrasonic waves in materials determines the elastic constants, which provides better understanding of the behaviour of the engineering materials. The elastic constants of material are related with the fundamental solid state phenomenon such as specific heat, Debye temperature and Grüneisen parameters. The elastic constants in the materials can be determined by measuring the velocity of longitudinal and shear waves. Elastic constants are related to interatomic forces, co-ordination changes etc., and also with the impact shock, fracture, porosity, crystal growth and microstructural factors (grain shape, grain boundaries, texture and precipitates etc.). So, the study of ultrasonic velocity is useful not only for characterization of the structured materials, engineering materials, porous materials, composites, glasses, glass ceramics but also bioactive glasses, nanomaterials, nanofluids etc.

Interferometer or continuous wave method and pulse technique are the general electrical method for the measurement of ultrasonic velocity. In CW method, the wavelength of wave in the test material is measured, which in turn provides the ultrasonic velocity with relation $V = \nu \lambda$. While in the Pulse technique, transit time (t: the time needed for a signal to travel between the front and back surface of the specimen or concerned medium) is measured with the help of echo pattern. If x is thickness of the material then ultrasonic velocity becomes equal to 2x/t.

For precise measurement, the Pulse technique has been improved in the form of following techniques (Papadakis, 1976, Raj, 2004).

a. Sing around
b. Pulse superposition method
c. Pulse echo overlap method
d. Cross-correlation method and
e. Phase slop method
f. Pulse transmission method

The pulse echo-overlap, pulse transmission and pulse superposition techniques are widely used techniques due to their absolute accuracy and precision respectively. Now a day, computer controlled devices of pulse echo overlap and pulse superposition techniques are being used. Resonance ultrasound spectroscopy and Laser interferometry are the recent techniques for the measurement of ultrasonic velocity in thin film, crystal, textured alloy etc.

5.1 C Application of ultrasonic velocity

Ultrasonic velocity has a wide range of application in the field of material characterization. Yet it is useful for the characterization or study of all the three phase of matter but here we

concentrate only its application to solid materials. It is used in the study of following properties of materials.

1. Elastic constants: The elastic moduli of a material are important for the understanding of mechanical behaviour. If V_L and V_S are the measured ultrasonic velocities of longitudinal and shear wave then longitudinal modulus (L), Shear modulus (G), Bulk modulus (B), Poisson's ratio (σ),Young modulus (Y) and lame's modulus (λ and μ) can be obtained with the following expression.

$$\left.\begin{aligned} L &= V_L^2 d \\ \mu = G &= V_S^2 d \\ B &= L - (4/3)G \\ \sigma &= \frac{L - 2G}{2(L - G)} \\ Y &= (1+\sigma)2G \\ \lambda &= (V_L^2 - 2V_S^2)d \end{aligned}\right\} \tag{5}$$

We can also find the stiffness constants or second and forth order elastic constants with the velocity. Using Table (1)-(2), one can find the second order elastic constants along different crystallographic direction for cubic and hexagonal structured materials. If we have ultrasonic velocity under different physical condition like temperature, pressure, composition of materials etc. then we can predict the mechanical behaviour of material in different physical condition. The anisotropy of material can be explained with the knowledge of anisotropy factor $A=[2C_{44}/(C_{11}-C_{12})]$. Knowledge of pressure derivatives of the elastic constants of a structured material can be used for the evaluation of Grüneisen parameter (γ). The Grüneisen parameter is used to describe anharmonic properties of solids. The quasi harmonic model is usually the starting point for the evaluation of mode gammas γ_i which is defined as $\gamma_i = -[d(\ln \omega_i)/d(\ln V)]$, where ω_i is a normal mode frequency of crystal lattice and V is the volume of the crystal. The values of γ_i for low frequency acoustic modes in a given material can be obtained with the pressure derivates of elastic constants of that material. Finally the Grüneisen parameter is obtained with the average of γ_i as shown in the following expression.

$$\gamma = \left[\sum_{i=1}^{3N} C_i \, \gamma_i \Big/ \sum_{i=1}^{3N} C_i \right] \tag{6}$$

Different workers (Mason, 1965; Brugger, 1964; Anil, 2005; Yadav, 2007; Yadawa, 2009; Yadav AK,2008) have studied this property of the different structured materials like isotropic, cubic, rhombohedral and hexagonal structured materials.

2. Debye temperature and Debye average velocity: These parameters are essential for the understanding of lattice vibration and low temperature properties of the material. These parameters can be found directly with the velocity values using equation (4) for the cubic and hexagonal structured materials. A detail study of Debye temperature, velocity and related theories of different structured materials can be seen elsewhere (Alers, 1965).

3. Porosity: The porosity of the porous material can be examined with the knowledge of elastic moduli and Poisson's ratio as a function of pore volume fraction. These parameters can be evaluated with help of measured velocity and density. A simple expression of Young modulus and shear modulus for a porous material can be written as,

$$\left.\begin{aligned} Y &= Y_0 \exp(-ap-bp^2) \\ G &= G_0 \exp(-ap-bp^2) \end{aligned}\right\} \quad (7)$$

Here Y_0 and G_0 are the modulus of material without pore; a,b and c are the constants; p is pore volume fraction which is equal to $\{1-(d/d_0)\}$ and; d is the bulk density determined experimentally from mass and volume while d_0 is the theoretical density determined from XRD.

The elastic moduli and Poisson ratio measured ultrasonically are compared with the theoretical treatment for the characterization. The elastic moduli of porous material are not only the function of porosity but also the pore structure and its orientation. The pore structure depends on the fabrication parameters like compaction pressure, sintering temperature and time. If the pores are similar in shape and distributed in homogeneous pattern then a good justification of mechanical property can be obtained with this study.

5. Grain size: There is no unique relation of average grain size with the ultrasonic velocity. The following typical graph (Fig. 5) shows a functional relation among velocity (V), grain size (D) and wave number (k). This has three distinct regions viz. decreasing, increasing and oscillating regions. Both the I and II region are useful for the determination of grain size determination, whereas region III is not suitable.

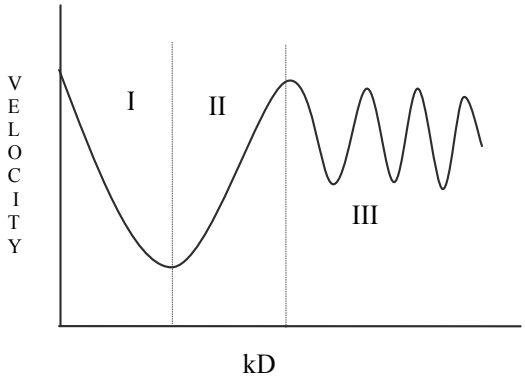


Fig. 5. Ultrasonic velocity as a function of kD

The obtained grain size with this study has good justification with grain size measured with metallography. The important advantage of using ultrasonic velocity measurements for the grain size determination is the accuracy in which ultrasonic transit time could be determined through electronic instrumentation. The different workers (Palanichamy, 1995) have studied this property for polycrystalline material with the study of ultrasonic velocity.

6. Anisotropic behaviour of compositional material: The intermetallic compound and alloys are formed by the mixing of two or more materials. These compounds have different mechanical properties depending on their composition. The different mechanical properties like tensile strength, yield strength, hardness (Fig.6) and fracture toughness at different

composition (Fig. 7), direction/orientation (Fig.8) and temperature can be determined by the measurement of ultrasonic velocity which is useful for quality control and assurance in material producing industries (Krautkramer, 1993; Raj, 2004; Yadav & Singh 2001; Singh & Pandey, 2009, Yadav AK, 2008).

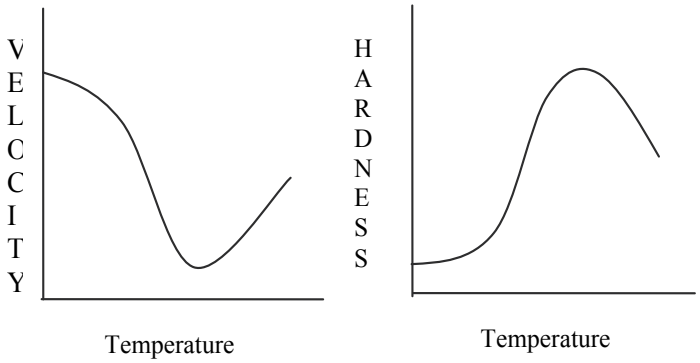


Fig. 6. Variation of velocity or hardness with temperature for some mixed materials

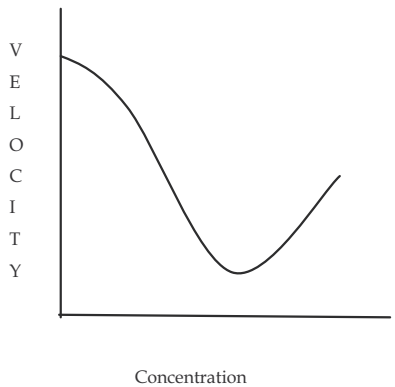


Fig. 7. Variation of velocity with concentration in some glasses

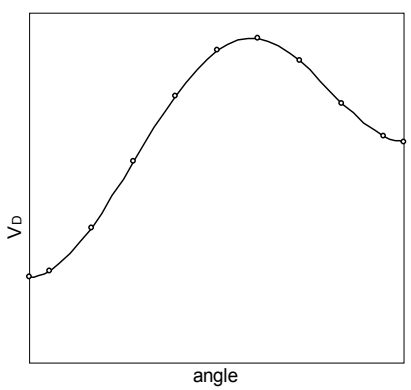


Fig. 8. Variation of V_D with the angle from the unique axis of hexagonal structured crystal

7. Recrystallisation: The three annealing process that amend the cold work microstructure are recovery, recrystallisation and grain growth. Among these processes, recrystallisation is the microstructural process by which new strain free grains form from the deformed microstructure. Depending on the material, recrystallisation is often accompanied by the other microstructural changes like decomposition of solid solution, precipitation of second phases, phase transformation etc. The hardness testing and optical metallography are the common techniques to the study the annealing behaviour of metals and alloys. A graph of longitudinal and shear wave velocity with annealing time (Fig.9) provides a more genuine understanding of recrystallisation process.

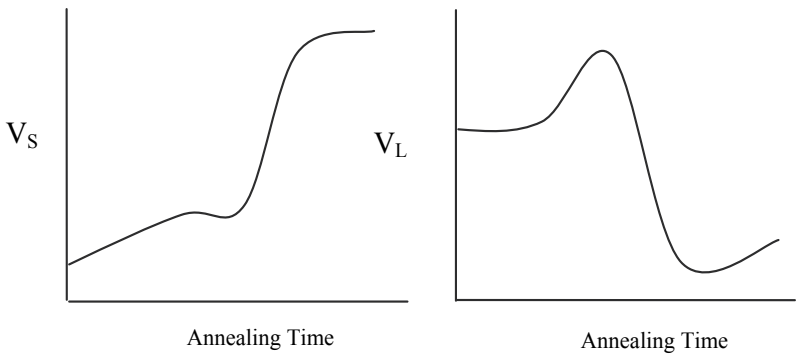


Fig. 9. Variation of V_L or V_S with annealing time

The variation of shear wave velocity represents a slight increase in recovery region followed by a rapid increase in the recrystallisation region and saturation in the completion of recrystallisation region. The slight increase in the velocity in the process of recovery is attributed to the reduction in distortion of lattice caused by the reduction in point defect due to their annihilation. The increase in velocity during recrystallisation is credited to the change in the intensity of lattice planes. The variation in longitudinal velocity have the just opposite trend to that of shear wave velocity which is credited to the change in texture and the dependence of velocity directions of polarisation and propagation of wave. The variation of velocity ratio (V_L/V_S) with annealing time shows a clear picture of recrystallisation regime (Fig. 10).

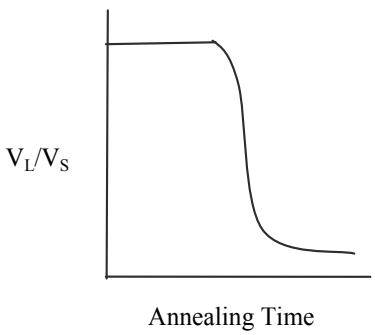


Fig. 10. Variation of V_L/V_S with annealing time

The selection of ratio avoids the specimen thickness measurement and enhances the accuracy. In short we can say that the velocity measurement provides the accurate prediction of on set and completion times of recrystallisation.

8. Precipitation: For the desired strength of material or component, the precipitation is a process like recrystallisation. It is a metallurgical process for the improvement of strength of material. The strength of improvement depends on spacing, size, shape and distribution of precipitated particles. A measurement of longitudinal ultrasonic wave velocity with ageing time provides precise value of Young modulus at different ageing temperature (Bhattacharya, 1994; Raj, 2004). With the knowledge Young modulus, the strength of material at different time of ageing can be predicted. Thus ultrasonic evaluation may be handy tool to study the precipitation reaction involving interstitial elements because this mechanism is associated with large change in the lattice strain.

9. Age of concrete: There are several attempts that have been made to find the elastic moduli, tensile strength, yield strength, hardness, fracture toughness and brittleness of different materials (Lynnworth, 1977; Krautkramer,1977). Similarly the age of concrete material can be determined with knowledge of crush strength that can be found with the ultrasonic velocity. A graph of pulse velocity of ultrasonic wave and crush with age of concrete is shown in Fig 11.

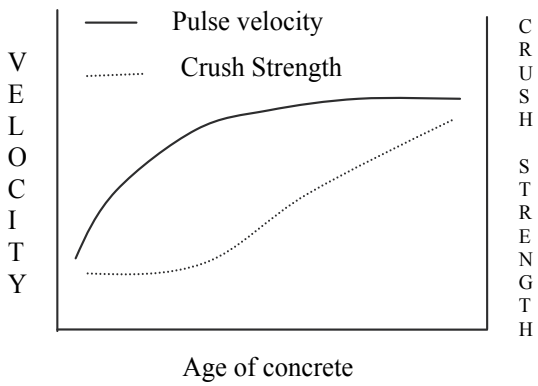


Fig. 11. Variation of velocity and crush strength with age of concrete

10. Cold work and texture: The texture of compounds can be understood with the knowledge of ultrasonic velocity. The expression of texture designates an elastic anisotropy due to the non-random distribution of crystalline directions of the single crystals in the polycrystalline aggregates. On the contrary, the isotropic, untextured solid is characterized by a totally random distribution of the grains. A study on texture gives insight into the materials plastic properties. Ultrasonic velocity measurements provide the state of texture in the bulk. For this purpose, ultrasonic velocity with cross correlation method {V_{IJ}; where I (direction of propagation) or J (direction of polarization) =1,2,3; 1:rolling, 2: transverse, 3:normal) }or Rayleigh wave velocity in transverse direction is measured as function of cold work (Raj,2004). Accordingly, three longitudinal (V_{11}, V_{22}, V_{33}) and six shears (V_{12}, V_{21}, V_{23}, V_{32}, V_{31} and V_{13}) wave velocities are measured. The velocities are found to be identical when the direction of propagation and direction of polarization are interchanged. Yet the measured velocities of longitudinal and shear wave propagating perpendicular to rolling

direction are important for estimation of cold work with good precision but V_{33} and V_{32} are found to be more suitable due to being easier in measurement. With the following relation, we can estimate the degree of cold work with help of velocity ratio (V_{33} / V_{32}).

$$V_{33} / V_{32} = 0.00527 \text{ (\% cold work) } -1.83 \text{ ; } \{\text{Correlation coefficient} = 0.9941\} \quad (8)$$

The following graph (Fig. 12) represents the variation of velocity ratio with cold work.

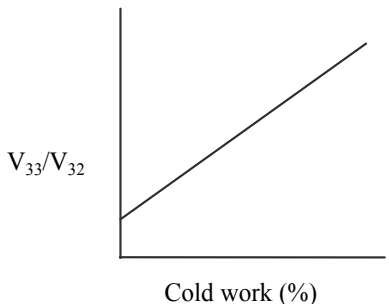


Fig. 12. Variation of velocity ratio with cold work

The Rayleigh wave velocity in transverse direction decreases with cold work and is linear in nature. A scatter in measurement is mainly attributed to the local variation in the degree of deformation, particularly close to surface caused by scattering. Both the methods are appropriate for the evaluation of cold work percentage in stainless steel. Thus measurement of bulk and surface Rayleigh wave velocities on cold rolled plates provide a tool to monitor the percentage of cold work during rolling operation.

5.2 Ultrasonic attenuation

The intensity of ultrasonic wave decreases with the distance from source during the propagation through the medium due to loss of energy. These losses are due to diffraction, scattering and absorption mechanisms, which take place in the medium. The change in the physical properties and microstructure of the medium is attributed to absorption while shape and macroscopic structure is concerned to the diffraction and scattering. The absorption of ultrasonic energy by the medium may be due to dislocation damping (loss due to imperfection), electron-phonon interaction, phonon-phonon interaction, magnon-phonon interaction, thermoelastic losses, and bardoni relaxation. Scattering loss of energy is countable in case of polycrystalline solids which have grain boundaries, cracks, precipitates, inclusions etc. The diffraction losses are concerned with the geometrical and coupling losses, that are little or not concerned with the material properties. Thus in single crystalline material, the phenomenon responsible to absorption of wave is mainly concerned with attenuation. An addition of scattering loss to the absorption is required for knowledge of attenuation in polycrystalline materials. So, the rate of ultrasonic energy decay by the medium is called as ultrasonic attenuation.

The ultrasonic intensity/energy/amplitude decreases exponentially with the source. If I_X is the intensity at particular distance x from source to the medium inside then:

$$I_X = I_0 e^{-\alpha X} \tag{9}$$

where α is attenuation or absorption coefficient. If I_{X_1} and I_{X_2} are the intensities of ultrasonic waves at x_1 and x_2 distance then from equation (9) one can write the following expressions.

$$I_{X_1} = I_0 e^{-\alpha X_1} \tag{10}$$

$$I_{X_2} = I_0 e^{-\alpha X_2} \tag{11}$$

On solving the equations (10) and (11), one can easily obtain the following expression of ultrasonic attenuation.

$$\alpha = \frac{1}{(x_2 - x_1)} \log_e \frac{I_{X_1}}{I_{X_2}} \tag{12}$$

The ultrasonic attenuation or absorption coefficient (α) at a particular temperature and frequency can be evaluated using equation (12). In pulse echo-technique the (X_2-X_1) is equal to twice of thickness of medium because in this technique wave have to travel twice distance caused by reflection, while is equal to medium thickness in case of pulse transmission technique. Attenuation coefficient is defined as attenuation per unit length or time. i.e. The α is measured in the unit of Np cm^{-1} or Np t^{-1}. The expression of α in terms of decibel (dB) unit are written in following form.

$$\alpha = \frac{1}{(x_2 - x_1)} 20 \log_{10} \frac{I_{X_1}}{I_{X_2}}; \quad \text{in unit of dB/cm} \tag{13a}$$

$$\alpha = \frac{V}{(x_2 - x_1)} 20 \log_{10} \frac{I_{X_1}}{I_{X_2}}; \quad \text{in unit of dB/}\mu\text{s} \tag{13b}$$

5.2 A Source of ultrasonic attenuation

The attenuation of ultrasonic wave in solids may be attributed to a number of different causes, each of which is characteristic of the physical properties of the medium concerned. Although the exact nature of the cause of the attenuation may not always be properly understood. However, an attempt is made here to classify the various possible causes of attenuation that are as.

a. Loss due to thermoelastic relaxation
b. Attenuation due to electron phonon interaction
c. Attenuation due to phonon phonon interaction
d. Attenuation due to magnon-phonon interaction
e. Losses due to lattice imperfections
f. Grain boundary losses
g. Loss due Bardoni relaxation and internal friction

A brief of these losses can be under stood by the following ways.

a. Loss due to thermoelastic relaxation

A polycrystalline solid may be isotropic because of the random orientation of the constituent grains although the individual grains may themselves be anisotropic. Thus, when a given stress is applied to this kind of solid there will be variation of strain from one grain to another. A compression stress causes a rise in temperature in each crystallite. But because of the inhomogeneity of the resultant strain, the temperature distribution is not uniform one. Thus, during the compression half of an acoustic cycle, heat will flow from a grain that has suffered the greater strain, which is consequently at high temperature, to one that has suffered a lesser strain, which as a result is at lower temperature. A reversal in the direction of heat flow takes place during the expansion half of a cycle. The process is clearly a relaxation process. Therefore, when an ultrasonic wave propagates in a crystal, there is a relaxing flow of thermal energy from compressed (hot region) towards the expanded (cool region) regions associated with the wave. This thermal conduction between two regions of the wave causes thermoelastic attenuation. The loss is prominent for which the thermal expansion coefficient and the thermal conductivity is high and it is not so important in case of insulating or semi-conducting crystals due to less free electrons. The thermoelastic loss $(\alpha)_{Th}$ for longitudinal wave can be evaluated by the Mason expression (Bhatia, 1967; Mason, 1950, 1965) .

$$\alpha_{Th.} = \frac{\omega^2 < \gamma_i^j >^2 KT}{2dV_L^5} \tag{14a}$$

$$(\alpha / f^2)_{Th.} = \frac{4\pi^2 < \gamma_i^j >^2 KT}{2dV_L^5} \tag{14b}$$

where ω and V_L are the angular frequency and longitudinal velocity of ultrasonic wave. d, K and T are the density, thermal conductivity and temperature of the material. γ_i^j is the Grüneisen number, which is the direct consequence of the higher order elastic constants (Mason, 1965; Yadawa 2009). In the case of shear wave propagation, no thermoelastic loss occurs because of no any compression & rarefaction and also for the shear wave, average of the Grüneisen number is zero.

b. Attenuation due to electron-phonon interaction

Debye theory of specific heat shows that energy exchanges occur in metals between free electrons and the vibrating lattice and also predicts that the lattice vibrations are quantized in the same way as electromagnetic vibrations, each quantum being termed as phonon.Ultrasonic absorption due to electron-phonon interaction occurs at low temperatures because at low temperatures mean free path of electron is as compared to wavelength of acoustic phonon. Thus a high probability of interaction occurs between free electrons and acoustic phonons. The fermi energy level is same along all directions for an electron gas in state of equilibrium, i.e. the fermi surface is spherical in shape. When the electron gas is compressed uniformly, the fermi surface remains spherical. The passage of longitudinal ultrasonic wave through the electron gas gives rise to a sudden compression (or rarefaction) in the direction of the wave and the electron velocity components in that direction react immediately, as a result fermi surface becomes ellipsoidal. To restore the spherical distribution, collision between electron and lattice occur. This is a relaxational phenomenon because the continuous varying phase of ultrasonic wave upsets this distribution.

In a new approach we may understood that the energy of the electrons in the normal state is carried to and from the lattice vibrations by means of viscous medium, i.e. by transfer of momenta. Thus the mechanism is also called as electron-viscosity mechanism. The ultrasonic attenuation caused by the energy loss due to shear and compressional viscosities of electron gas for longitudinal $(\alpha)_{Long}$ and shear waves $(\alpha)_{Shear}$ are given as (Bhatia, 1967; Mason, 1950, 1965,66):

$$(\alpha)_{Long} = \frac{\omega^2}{2dV_L^3}\left(\frac{4}{3}\eta_e + \chi\right) \tag{15a}$$

$$(\alpha)_{Shear} = \frac{\omega^2}{2dV_S^3}\eta_e \tag{15b}$$

where η_e and χ represent the electronic shear and compressional viscosities of electron gas.

c. Attenuation due to phonon–phonon interaction

The energy quanta of mechanical wave is called as phonon. With the passage of ultrasound waves (acoustic phonons), the equilibrium distribution of thermal phonons in solid is disturbed. The re-establishment of the equilibrium of thermal phonons are maintained by relaxation process. The process is entropy producing, which results absorption. The concept of modulated thermal phonons provides following expression for the absorption coefficient of ultrasonic wave due to phonon–phonon interaction in solids $(\alpha)_{Akh}$ (Bhatia, 1967; Mason, 1950, 1958, 1964, 1965; Yadav & Singh 2001; Yadawa, 2009) .

$$\alpha_{Akh} = \alpha_{PP} = \frac{\omega^2\tau\ \Delta C}{2dV^3(1+\omega^2\tau^2)} \tag{16a}$$

Where τ is the thermal relaxation time (the time required for the re-establishment of the thermal phonons) and V is longitudinal or shear wave velocity. ΔC is change in elastic modulli caused by stress (by passage of ultrasonic wave) and is given as:

$$\Delta C = 3E_0 < (\gamma_i^j)^2 > - < \gamma_i^j >^2 C_v T \tag{16b}$$

Here E_0 is the thermal energy density. ΔC is related with the acoustic coupling constant (D), which is the measure of acoustic energy converted to thermal energy due to relaxation process and is given by the following expression:

$$D = \frac{3\Delta C}{E_0} = 9 < (\gamma_i^j)^2 > 3 < \gamma_i^j >^2 \frac{C_v T}{E_0} \tag{16c}$$

Using equation (16c), the equation (16a) takes the following form under condition $\omega\tau << 1$.

$$\alpha_{Akh} = \alpha_{PP} = \frac{\omega^2\tau\ E_0 D}{6\ dV^3} \tag{16d}$$

d. Attenuation due to magnon-phonon interaction

Ferromagnetic and ferroelectric materials are composed of 'domains' which are elementary regions characterized by a unique magnetic or electric polarization. These domains are

aligned along a number of directions, but generally oriented along the polarization vector that is known as direction of easy magnetization (or electrification). These usually follow the direction of the principal crystallographic axis. Cubic crystal of a ferromagnetic material has six directions of easy magnetization lying in positive or negative pairs along the three perpendicular co-ordinate axes. Thus two neighbouring domains are aligned at 90^0 or 180^0. Because of the magnetostriction effect, assuming that the magnetostructive strain coefficient is positive (or negative), there is an increase (or decrease) in the length of domains in the direction of polarization. Which results an increase or decrease in elastic constants depending on sign of the magnetostructive coefficient. The magnitude of change depends on applied stress. The phenomenon is called as ΔE effect. Thus when a cyclic stress such as produced by ultrasonic wave, is applied to a ferromagnetic or ferroelectric material, the domain wall displaced as a result of ΔE effect that follows the hysterisis loop. Thus there is dissipation of ultrasonic energy. The loss per half cycle per unit volume is being given by area of hysterisis loop.

The another cause of the attenuation in ferromagnetic material is due to production of micro-eddy current produced in domain walls by the periodic variation of magnetic flux density. A simple consideration of the ultrasonic attenuation in ferromagnetic material is due to magnetoelastic coupling i.e attenuation is caused by interaction between magnetic energy in form of spin waves (magnon- energy quanta of spin waves) and ultrasonic energy (phonon). Thus it is called as ultrasonic attenuation due to magnon-phonon interaction.

e. Losses due to lattice imperfections

Any departure from regularity in the lattice structure for a crystalline solid is regarded as an imperfection, includes point defects such as lattice vacancies and presence of impurity atom and dislocation etc. Imperfections enhance the absorption of ultrasonic wave. Attenuation due to dislocation can occur in more than one way e.g. attenuation due to edge or screw dislocation, which is due to forced vibration in imperfect crystal i.e. due to interaction of ultrasonic energy (phonon) and vibrational energy of impurity atom or dislocation (phonon). Dislocation drag is a parameter for which the phonon-phonon interaction can produce an appreciable effect on the motion of linear imperfections in the lattice through drag phenomenon. The thermal loss due to such motion can be computed by multiplying the following drag coefficients by the square of the dislocation velocity (Yadav & Pandey, 2005).

$$B_{screw} = 0.071\varepsilon \tag{17a}$$

$$B_{edge} = \frac{0.053\varepsilon}{(1-\sigma^2)} + \frac{0.0079}{(1-\sigma^2)}\left(\frac{G}{B}\right)\chi \tag{17b}$$

Where $\chi = \varepsilon_L - (4\varepsilon_S / 3)$; $\varepsilon_L = E_0 D_L \tau_L / 3$, $\varepsilon_S = E_0 D_S \tau_S / 3$, $B = (C_{11} + 2C_{12}) / 3$, $G = (C_{11} - C_{12} + C_{44}) / 3$ and $\sigma = C_{12} / (C_{11} + C_{12})$. Here G, ε, σ, B and χ are the shear modulus, phonon viscosity, Poisson's ratio, bulk modulus and hydrostatic compressional viscosity respectively. ε_L & ε_S, D_L & D_S and τ_L & τ_S are phonon viscosity, acoustic coupling constant and thermal relaxation time for longitudinal and shear wave. C_{11}, C_{12} and C_{44} are the second order elastic constants for cubic metals.

f. Grain boundary losses

The grain boundary losses occur due to random orientation of the anisotropic grains in polycrystalline solid. At each grain boundary there is discontinuity of elastic modulus.

Therefore when ultrasonic wave of small wavelength compared to grain size propagates in such solid, regular reflections occur at grain boundaries, causes loss. The loss depends on the degree of the anisotropy of the crystallites, mean grain diameter and wavelength of wave. When the grain size is comparable to wavelength of wave then the ultrasonic attenuation caused by elastic hysterisis at grain boundary and scattering is frequency dependent and can be related as:

$$\alpha = B_1 f + B_2 f^4 \tag{18}$$

Where B_1 and B_2 are constants for the given material.

i. Loss due Bardoni relaxation and internal friction: The attenuation maximum at low temperature in some metarials like (Pb, Cu, Ag and Al) whose position on temperature scale is a function of the frequency of measurement is called as Bardoni peaks (Bhatia, 1967). These peaks are very small but when the crystal is strained by one or two percent, the peaks appear very prominantaly. These peaks are relaxational peaks. This relaxation is due to dislocation which are in the minimum energy position and are moved over the Peierls energy barrier by thermal agitation. A freshly strained material have its dislocations lying along minimum energy regions. A dislocation line between two pining points could be displaced by thermal agitation, and that the small stress would bias the potential wells and cause a change in the number of residing in the side wells, thus producing a relaxation effect. A typical graph showing Bardoni peaks under unstrained and strained condition is shown in Fig.13.

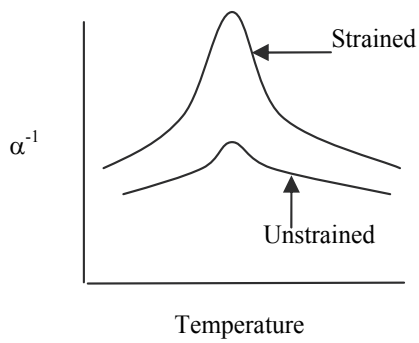


Fig. 13. Attenuation peaks at low temperature under unstrained and strained condition of materials

As the temperature increases there is an exponential increase in loss occuring at high temoperatures. It is observed for a number of polycrystalline material which is due to grain boundary relaxation effect. Such peaks are absent for the single crystals. There is also attenuation peaks on temperature scale for a number of material due to internal friction. This has been ascribed to the drag of dislocation as they are pulled through a concentration of vacancies. The internal friction peaks are caused due to damping effect of dragging the dislocations along vacancies or it can be assumed to be associated with the breakway of dislocations from their pinning points caused by thermal vibrations of the dislocation. This loss is independent of frequency and is greatly enhanced by the amount of cold work. The position of peaks appear to be independent of impurity content of the material. The loss due to internal friction can be related to frequency with following equation.

$$\alpha^{-1} = \frac{\Delta E}{E} \frac{(f / f_R)}{1 + (f / f_R)^2} \tag{19a}$$

Where $(\Delta E / E)$ is the relaxation strength, f and f_R are the frequency and relaxation frequency respectively. f_R is related to the activation energy (H).

$$f_R = f_0 \, e^{-H/RT} \tag{19b}$$

Here f_0 is the frequency with which the unit causing the relaxation attacks the energy and T is the temperature. For the frequencies f greater than f_R, the equation (19a) takes the following form.

$$\alpha^{-1} = \frac{\Delta E}{E} \frac{f_0}{f} e^{-H/RT} \tag{20}$$

On the basis of above theories of ultrasonic attenuation, it is clear that if hypothetical crystal under study is perfect, not ferromagnetic or ferroelectric then only three factors are predominantly responsible for ultrasonic attenuation that are attenuation due to thermoelastic relaxation, electron-phonon interaction and phonon-phonon interaction.

$$\alpha_{Total} = \alpha_{Th} + \alpha_{ep} + \alpha_{pp} \tag{21}$$

For nanosized metallic crystals the dislocation drag parameter gives informative results that can be used for the analysis of nanostructured materials. The electron-phonon interaction is prominant only at low temperatures while phonon-phonon interaction is effective at high temperatures. The total attenuation in magnetic material at high temperature is sharply affected with phonon-phonon and magnon-phonon interactions not only at bulk scale but also at nanoscale. When metal nano particles are dispersed in suitable polymer, then it is called as nanofluid. If the particles are of magnetic material then it is called as ferrofluid.
The total ultrasonic attenuation in ferrofluid on the temperature scale can be written as:

$$\alpha_{Total} = \alpha_V + \alpha_{MP} + \alpha_{pp} \tag{22}$$

where α_V:absorption due to viscous medium, α_{MP}: absorption due to interaction between acoustic phonon and magnon (energy quanta of spin wave associated with dis- persed particles) and α_{PP}: absorption due to interaction between acoustic phonon and dispersed crystal lattice phonon.

5.2 B Measurement techniques of ultrasonic attenuation

Similar to velocity measurement, the pulse technique and continuous wave method are being used for the measurement of ultrasonic attenuation now a day. On the basis of measurement procedure, the pulse technique is mainly classified in pulse transmission technique, pulse-echo-technique and pulse echo overlap technique. Following is a short view of pulse echo and pulse transmission techniques for the measurement of attenuation.
In the pulse-echo technique (PET) of ultrasonic testing, an ultrasound transducer generates an ultrasonic pulse and receives its echo. The ultrasonic transducer functions as both transmitter and receiver in one unit. The block diagram is shown in Fig 14. Most ultrasonic transducer units use an electronic pulse to generate a corresponding sound pulse, using the

piezoelectric effect. A short, high voltage electric pulse (less than 20 Ns in duration, 100-200 V in amplitude) excites a piezoelectric crystal, to generate an ultrasound pulse.

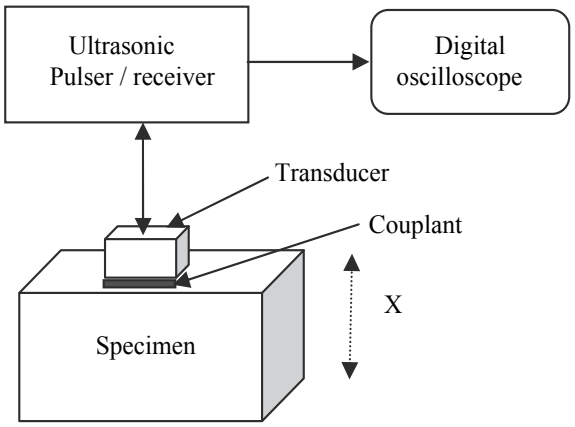


Fig. 14. Block diagram of PET

The transducer broadcasts the ultrasonic pulse at the surface of the specimen. The ultrasonic pulse travels through the specimen and reflects off the opposite face. The transducer receives the reflected echoes. The ultrasound pulse keeps bouncing off the opposite faces of the specimen, attenuating with time. The attenuation coefficient can be determined by measuring the amplitudes of the echoes from the time domain trace using the following equation.

$$\alpha = \frac{1}{2X} 20\log_{10}\left[\frac{1}{(m-n)}\frac{I_n}{I_m}\right] \qquad \text{; in unit of dB/cm} \tag{23}$$

where I_m and I_n are the maximum amplitude (voltage) of the mth and nth pulse echoes respectively. X is the specimen thickness. Normally the first and second back wall echo are used that is m=2 and n=1. The accuracy of the transit time and attenuation in this technique depend on the selection of peak amplitude of echoes and its height respectively. The Overall accuracy in the transit time in this method is the order of nanosecond.

In the Pulse transmission technique (PTT), there is separate transducer and receiver for producing and receiving the signal, that are attached on the both side of specimen through suitable couplant via wave guides (Fig.15).

This technique can be used for the both velocity and attenuation measurement. For the velocity measurement, the transit times (t_1 and t_2) are determined in the in the absence and presence of the sample between waveguides. The difference of these transit times ($\Delta t = t_2 - t_1$) provides the actual transit time for sample. If sample thickness is X then ultrasonic velocity in the sample becomes equal to $X / \Delta t$. Similarly If $I_w(f)$ refers to the amplitude of the received signal with the waveguides only and $I_s(f)$ refers to the amplitude of the received signal when the sample is inserted between the wave guides then the attenuation of the ultrasonic waves in the sample is measured using the following relation.

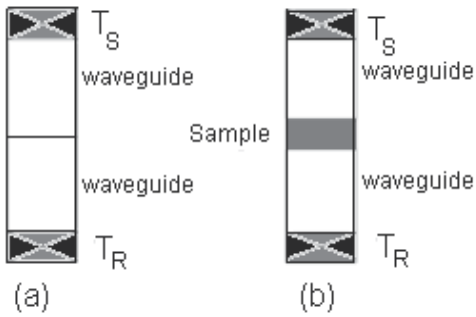


Fig. 15. Arrangement of transducer/receiver, waveguide and sample in PTT

$$\alpha = \frac{1}{X}\left[\ln T_C + \ln\frac{I_W(f)}{I_S(f)}\right] \tag{24a}$$

Here T_C is combined transmission coefficient at the sample and waveguide interface, that can be calculated with the following relation.

$$T_C = \frac{4\,Z_W Z_S}{(Z_W + Z_S)^2} \tag{24b}$$

Where Z_W and Z_S are the acoustic impedances of the waveguide and sample respectively. The exact value of attenuation in the material can not be measured from the direct measurements. It can be obtained only by the conventional attenuation method. The measured attenuation posses all loses introduced by couplant, diffraction, non-parallel specimen surfaces etc. The true value of attenuation can be obtained only when all these losses are accounted separately and subtracted from the experimental obtained value of attenuation.

5.2 C Properties characterized with ultrasonic attenuation

The ultrasonic attenuation coefficient is well correlated to several physical parameters and properties of the material. The following diagram (Fig.14) represents a view of their dependence.

Being a broad relation with material properties, the several properties of the material can be defined like grain size, yield strength, ductile to brittle transition temperature, Neel temperature, deviation number, behaviour of mechanical and magnetic properties with temperature and composition etc. The phenomenon responsible for attenuation can also be understood with the knowledge of ultrasonic attenuation. Yet there are several work have been made for the characterization of material on the basis of velocity and attenuation but here we will discuss the velocity attenuation in some structured materials like fcc, bcc, hcp, heaxagonal, NaCl / CsCl type structured materials etc.

6. Ultrasonic attenuation and velocity in different materials

Ultrasonic attenuation, velocity and their related parameters can be used to give insight into materials microstructures and associated physical properties. Behaviour of ultrasonic

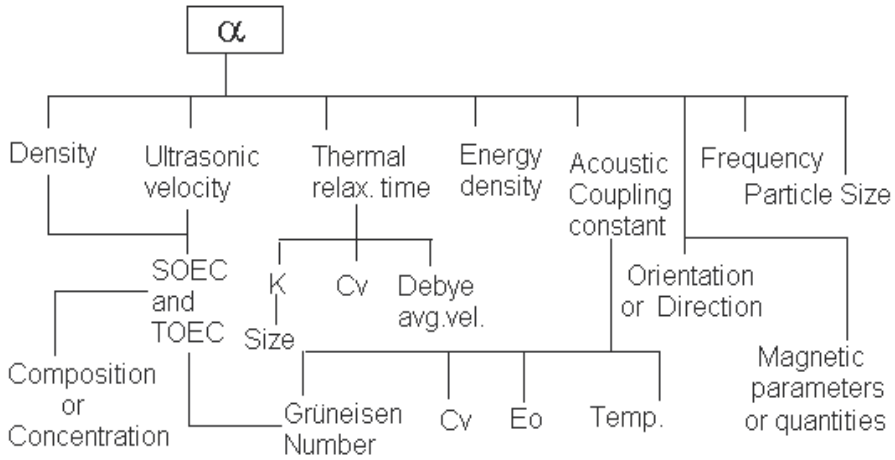


Fig. 14. Dependence of attenuation coefficient on several parameters of the material.

attenuation and velocity as a function of physical parameters related to different physical condition is used to characterize the material during the processing as well as after production. Ultrasonic can be used for the characterization of metal, rare-earth metal, semimetal, semiconductor, alloy, intermetallic, dielectric, glass, glass-ceramic, superconductor for the determination of their characteristic properties at different physical conditions like temperatures, pressure, field crystallographic direction, electric and magnetic field. Ultrasonic can also be used for the preparation and investigation of nanomaterials. Thus it is an efficient tool for the diagnosis of the material not only in bulk scale but also in nanoscale. Such interpretation is important for the quality control and assurance of the material for the industries. On the basis of structure, the materials can be divided into two classes mainly as crystalline (single crystal and polycrystalline) and amorphous. The crystalline material can have different structures like fcc, bcc, hcp, hexagonal, NaCl / CsCl type, trigonal, orthorhombic, tetragonal, monoclinic, triclinic etc. The ultrasonic study of some structured materials is written below.

Monochalcogenides of the rare-earth elements (ReX, with Re=rare-earth element Re=La, Ce, Pr, Nd, Sm, Eu, Tm and X=S, Se and Te) comprise a large class of materials that crystallize in simple NaCl-type structure. ReX exhibits interesting electrical, optical and magnetic properties. The thulium monochalcogenides TmX (X=S, Se and Te) have NaCl-type structure. Tm compounds exhibit Van Vleck paramgnetism at low temperatures owing to crystal-field singlet ground states. TmS, TmSe and TmTe are golden metal, red brown coloured intermediate valance system and silver blue semiconductor respectively. These materials are technologically important having many applications ranging from catalysis to microelectronics. Ultrasonic attenuation and other associated parameters like ultrasonic velocities, acoustic coupling constants etc. along <100>, <110> and <111> directions in the temperature range 100-300K have been studuied elswhere (Singh, Pandey & Yadawa, 2009). The order of thermal relaxation time for TmTe, TmS and TmSe are found of the order of 10^{-11}sec, 10^{-12}sec and 10^{-12}-10^{-13}sec respectively. This justifies that TmS, TmSe and TmTe have metallic, intermettallic and semiconducting behaviour. Total attenuation in these materials follows the expression $\alpha = \sum_{n=0}^{n=2} \alpha_n \, T^n$. The value of α_n depends on specific heat per unit

volume, energy density, thermal relaxation time, thermal conductivity, elastic constants and density.

The lowest attenuation is found in TmSe. This infers that this material has excellent purity and ductility in comparison to the TmS and TmTe. Thus on the basis of ultrasonic attenuation, the classification of materials can be made, i.e. it is either metallic, intermediate valence, semiconductor or dielectrics. Praseodymium and lanthanum monochalcogenides (PrS, PrSe, PrTe, LaS, LaSe, LaTe) are the materials which are used as a core material for carbon arcs in the motion picture industry for studio lighting projection. The ultrasonic study of these materials (Yadav & Singh, 2001, 2003) shows that the variation of ultrasonic attenuation with temperature in these are same as for thulium monochalcogenides. In the all monochalcogenides, the ultrasonic velocity increases with temperature due increases in the elastic constnats. The low temperature ultrasonic study of in intermetallic compound GdP, GdAs and GdSb (Yadav & Singh, 2001) shows that the temperature variation of the longitudinal ultrasonic attenuation is predominantly affected with the electrical resistivity and provides the information about the Neel temperature. The high temperature and directional ultrasonic study of SnTe, EuSe and CdO semi-conducting materials (Singh & Yadav, 2002) implies that the thermal conductivity is the governing parameter to the ultrasonic attenuation in SnTe, EuSe and CdO materials. The ultrasonic study of B1 structured CeS, CeSe, CeTe, NdS, NdSe and NdTe along different crystallographic directions at room temperature (Singh, 2009) implies NdS is more ductile and stable material in comparison to other chalcogenides systems (CeS, CeSe, CeTe, NdSe, NdTe, LaS, LaSe, LaTe, PrS, PrSe and PrTe) and rock salt-type LiF single crystal due to its lowest value of attenuation.

Aluminides are generally the most famous group of intermetallic compounds. Intermetallic compounds containing aluminium such as NiAl, offer new opportunities for developing low density, high strength structural alloys which might be used at temperatures higher than possible with conventional titanium and nickel-base alloys. Once developed, the intermetallic alloys and their composites will enable the design and production of higher performance, lighter (high thrust-to-weight ratio) engines for future military aircraft and supersonic commercial transport. Strong bonding between aluminium and nickel, which persists at high temperatures, can provide high strength at elevated temperatures such that the specific strength of intermetallics could be competitive with superalloys and ceramics. However, the high strength is usually associated with poor ductility. With respect to ductility, intermetallics fall between metals and ceramics. Intermetallics are not as brittle as ceramics because the bonding in intermetallics is predominantly metallic, compared to ionic or covalent bonding of ceramics. Nickel aluminide (NiAl) has been the subject of many development programs. The β-phase NiAl (50 at % Ni, 50 at % Al, with a CsCl, B_2 crystal structure), is very different from the γ'-phase Ni_3Al (75 at % Ni, 25 at % Al, $L1_2$ crystal structure) with respect to physical and mechanical properties. NiAl has four key advantage: its density of ≈5.95 g/cm^3 is approximately two thirds the density of state-of-the-art nickel-base superalloys; its thermal conductivity is four to eight times that of nickel-base superalloys (depending on composition and temperature); it has excellent oxidation resistance. In both the polycrystalline and single crystal forms, NiAl is brittle at room temperature in most cases and ductile at high temperatures.The elastic and ultrasonic study of β-phase NiAl at high temperature has been done elsewhere (Yadav & Pandey, 2006). A comparison of second order elastic constant Ni and Al pure metals at ≈300K with the values

of NiAl implies that the elastic anisotropy of NiAl is lower than the value of Ni and Al. A low value of anisotropy favors instability. Thus the intermetallic compound NiAl is unstable in comparison to pure metal Ni and Al. The anisotropy (A) of NiAl is found to increases with temperature. Thus intermetallic NiAl is stable at very high temperatures (300-1400K). The longitudinal ultrasonic attenuation in NiAl is found to decreases very fastly from 300 to 700K and slowly from 700K to 900K; then it recieves an increases gradually from ≈900 to 1400K. The ductile to brittle transition temperature (DBTT) is only 625 to 700K. Yield strength of NiAl decreases from ≈900K. This is predicted as the ultrasonic attenuation increases gradually from this temperature. Thus the structural stability, abrupt change in ductility at ≈DBTT, disordering at ≈DBTT could be predicted on the basis of temperature variation of the elastic constants and the ultrasonic attenuation in of NiAl.

Intermetallic compounds have received extensive attention in recent years because of technical promise as high temperature structural materials. The study of intermetallics has attracted the attention of the scientific world because of their anisotropic properties. It has also been found that fine application in advanced power engineering. Since many intermetallic compounds of different crystal structure have been found in alloy systems, the basic reason for their stability has drawn a great deal. The materials AgMg, CuZr, AuMg, AuTi, AuMn, AuZn and AuCd have a CsCl-type structure (B2 structure). The study of ultrasonic velocities, attenuation, Grüneisen parameters, non-linearity parameter (acoustic coupling constant), Debye temperature and thermal relaxation time at different crystallographic directions at room temperature can seen elsewhere (Singh & Pandey, 2009). The study of elastic constants shows that by introducing Mg, Zr, Ti, Mn, Zn, Cd in noble metals, the elastic behaviour slightly decreases due to loose interaction of impurity atoms with noble metal atoms. The Debye temperature (T_D) for Ag, Au, Cu, Cd, Zn and Mg are 226K, 289K, 224K, 214K, 272K and 400K respectively. The study also imples that the Debye temperature for mixed compounds with Ag or Cu lies between Debye temperatures of constituent materials while Debye temperature for mixed compounds with Au lies below than the constituent materials. The decrease or increase in Debye temperature indicates increase or decrease in acoustic contribution to the low temperature specific heat. The average sound velocities in these intermetallic compounds are not only larger than the noble metals but also with the Cs/Rb halides, which is due to low density of these compounds. Ultrasonic velocity in these materials decreases with their molecular weight. The velocity of these compounds is useful for determination of their anisotropic properties.The attenuation in these intermetallic compounds are mainly governed by phonon-phonon interaction and is greater than the Cs/Rb-halides and is less than the pure noble metals. The attenuation in these intermetallics are affected with combined effect of thermal conductivity, specific heat, average sound velocity and acoustic coupling constant. For CsCl-type structure, the deviation number Δ N exists from 1 to 3; Δ N denotes the difference of column number of noble metals and the secondary element in the helical periodic table. Δ N value for AgMg, CuZr, AuMg, AuTi, AuMn, AuZn and AuCd are 1, 3, 1, 3, 2, 1 and 1 respectively. The compounds AgMg, AuMg, AuZn and AuCd for which Δ N=1, have larger conductivity. The thermal conductivity are high for lower valued Δ N compounds. Since $(\alpha / f^2) \propto \tau \propto$ k thus one can write $(\alpha / f^2) \propto 1/\Delta N$. The ultrasonic attenuation in these intermetallics justify the above prediction. Thus, it may be concluded that in B2 structured intermetallic compounds the nature of ultrasonic attenuation can be determined by the deviation number.

Vanadium, Niobium and Tantalum are the transition elements of 5thB group in b.c.c. phase with high melting points and exhibit variable valency. Niobium and Tantalum are highly unreactive metals. Vanadium is seldom used on its own, but it is used in alloys of metals and acts like an important catalyst in oxidation reactions. Niobium is used in chromium nickel stainless steel, because it is unreactive and not rejected by human body. Tantalum is used for making metal plates, screws and wires for replacing badly fractured bones. The ultrasonic attenuation in these metal (Singh, Pandey, Yadawa & Yadav, 2009) decreases with the temperature and becomes negligible upto 40K, while in other normal metal this comes upto temperature 80K. This indicates that the electron-phonon interaction is possible upto 40K in these metals. Similar to other metals, the attenuation is dominated by the electrical resistance at low temperature in these metals.

The group III nitrides have unique properties such as wide direct band gap, high thermal conductivity, high thermal stability, high volume resistivity and high dielectric constant which make them the most serious candidates for high power and high frequency electronic and deep ultraviolet (UV) opto-electronic devices. The GaN, AlN and InN are hexagonal wurtzite structured Semiconducting materials. The temperature and orientation dependent ultrasonic study (Yadav & Pandey 2006; Pandey, Singh & Yadav 2007; Pandey & Yadav 2009) confirms that the AlN has minimum attenuation coefficient in comparison to GaN and InN. The temperature variation of attenuation coefficient for GaN has maximum at 400 K. The ultrasonic attenuation behaviour of AlN is just opposite to that found for GaN. Both studies indicate that the AlN is more stable and pure at high temperatures as it has low attenuation at each temperature than for GaN and the characteristic temperature for both is 400 K. It may also be predicted that at 400 K the material AlN has its purest and most ductile state as the ultrasonic attenuation in temperature range 300–800 K has a minimum at 400 K. The thermal conductivity/thermal relaxation and velocity/second order elastic constants are dominating factors to the ultrasonic attenuation before and after the temperature 400K respectively. The ultrasonic attenuation in GaN is affected by velocity and thermal energy density before 400K while after it the affecting factor is thermal relaxation time and acoustic coupling constant. In group III nitrides, the phonon-phonon interaction is the responsible mechanism for the total ultrasonic attenuation. The direction dependent ultrasonic study at room temperature of hexagonal structured rare-earth metals (Gd, Tb, Dy, Ho, Er and Tm), platinum group metal (Os and Ru), laves-phase compounds ($TiCr_2$, $ZrCr_2$ and $HfCr_2$) and fission products precipitated in nuclear fuel (Mo-Ru-Rh-Pd alloys) are reported in literature (Yadawa etal. 2009; Yadav AK etal. 2008; Pandey & Yadawa etal. 2007, 2009). The variation ultrasonic velocities with the angle from the unique axis of crystalline material are similar for all hexagonal structured material and are predominantly affected with the combined of second order elastic constants, while velocity magnitude differs due different elastic properties. Thermal relaxation times of these compounds are the order of 10^{-12}s which shows that the re-establishment of phonon distribution in equilibrium is obtained in 10^{-12}s after the passes of ultrasonic beam. The study shows that platinum group metals and rare-earth metals are durable and stable in their alloy form. The hexagonal structured materials have high elastic constant and low attenuation in comparison to fcc, bcc, NaCl/CsCl type structured material.

The ultrasonic study of fcc structured Pd and Pt, bcc structured Ta and hexagonal wurtize structured Zns at nanoscale indicates that the size dependent attenuation is dominated by

the thermal relaxation time/thermal conductivity (Yadav & Pandey, 2005; Pandey, Yadawa & Yadav 2007). The size variation of the thermal relaxation time for fcc/bcc follow the relation $\tau = \tau_0(1 - e^{-x/\lambda})$ while for hexagonal structure the expression is $\tau = \tau_0 e^{x/\lambda}$; here x: particle size, τ_0 and λ: constants. The attenuation and dislocation drag coefficient at nanoscale are larger than the normal scale.

When these nanoparticles are incorporated in suitable matrix (e.g. polymers) then nanofluids are formed. If the particles are of magnetic material then it is called as ferrofluid. The ultrasonic study of ferrofluid/nanofluid (Biwa, 2004; Singh DK etal 2009; Taketomi, 1986; Skumiel, 2000, 2003, 2004; Temkin, 1998; Gomez Alvarez, 2002) justifies the fact the velocity depends on the concentration of incorporated materials into the matrix and is independent of particle size in low frequency regime. At high frequency, both the particle size and concentration of nanoparticles are the affecting factor to the ultrasonic velocity. Ultrasonic attenuation in nanofluid is function of particle size, particle volume fraction and frequency. Commonly, the temperature dependence of ultrasonic velocity (V) for liquids is written as $V = V_0 + V_1 T$ (V_0 is ultrasonic velocity at initial temperature (273K), V_1 is absolute temperature coefficient of velocity and T is temperature difference between experimental and initial temperature). But the appropriate expression of velocity in nanofluid/Ferrofluid might be written as: $V = V_0 + V_1 T - V_2 T^2$. The third non linear term in velocity expression is caused by non-linear change in bulk modulus/density of solution/composite system with temperature. The ultrasonic study of Cr_2O_3 implies that the temperature variation of ultrasonic velocity in nano/ferrofluid mainly depends on the concentration of dispersed particles and the temperature variation of ultrasonic absorption provides direct information about Neel temperature of the ferrofluid. The Neel temperature of anti-ferromagnetic material increases at nanoscale. In the ferrofluid, absorption is mainly governed by viscous loss and magnon-phonon interaction below transition temperature while above it, the phonon-phonon interaction plays dominant role. The study of sound attenuation coefficient of magnetic fluid under an external magnetic fluid implies that the anisotropy in sound propagation is attributed to the two motions of the clusters of the ferrous colloidal particles in the fluid as rotational and translational (Taketomi 1986; Skumiel 2000, 2003, 2004).

The absorption study in glasses show a peak in attenuation at low temperature with change in temperature in simple glasses like silica, GeO_2, B_2O_3, As_2O_3 etc and in multicomponent glasses (Manghnami1,1974; Jackle, 1976). The change in velocity and attenuation in glasses are attributed to the structural change in glass network. The structural change are attributed to impurities or grain boundaries or anharmonicities of lattice. Several ultrasonic studies have been made to determine the elastic constants of the glasses such as alkali earth aluminosilicate, sodium borate, sodium borosilicate, and soda lime borosilicate glasses (Bhatti 1989; Rajendran, 2002). The transition temperature in high T_C superconductors can be obtained with the ultrasonic study (Bardeen, 1957). In this study the exponential decay of ultrasonic attenuation below T_C was used to obtain the energy gap in case of conventional mettalic superconductors, while change in ultrasonic velocity was used to explain transition temperatures for the type II superconductors.

The Grain size can be determined with the study of ultrasonic attenuation or relative attenuation (Papadakis, 1976). Normally the variation of ultrasonic attenuation or relative attenuation with average grain size or frequency follow the Fig. 15.

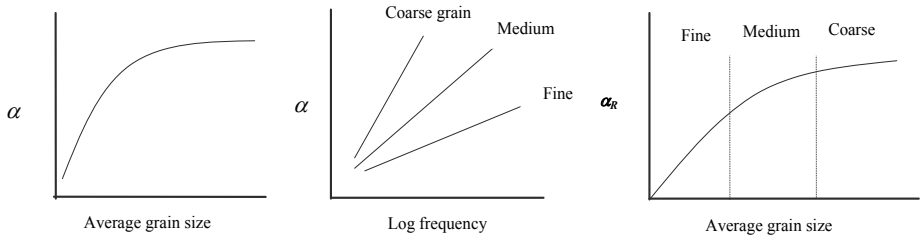


Fig. 15. variation of ultrasonic attenuation (α) or relative attenuation (α_R)with average grain size or frequency.

In this we can say the ultrasonics provide a big tool in the field of material characterization.

7. Summary

The present chapter deals the basics of ultrasonic wave generation and its detection. After that the theoretical and experimental techniques for the determination of ultrasonic properties have been discussed. The formulation of direction dependent ultrasonic velocity and its experimental measurement techniques are detailed for the understanding of mechanical properties in solids. The Different mechanisms responsible for the ultrasonic attenuation in solid material are explained to recognize the several properties of materials. Later on, the study of ultrasonic parameters in different structured materials like fcc, bcc, hcp, hexagonal, glasses, superconductors, nano-materials etc. and in nanofluids/ferrofluid is carried out. The whole study provides a short view of ultrasonic wave and its application to material characterization.

8. Acknowledgements

The authors are thankful to Prof. S. K. Kor, Prof. R. R. Yadav, Prof. R. K.Tiwari, Dr. Giridhar Mishra, Mr. D. K. Singh, Physics Department, University of Allahabad, Dr. Shripal, PPN College, Kanpur and Dr. Devraj Singh, Dr. P. K. Yadawa, Department of Applied Physics, Amity School of Engineering And Technology, New Delhi for their continuous support and valuable discussion during the preparation. One author Dr. D. K. Pandey is highly thankful to his wife Mrs. Archana Pandey for her breathless moral support up to completion of this project.

9. References

Alers, G. A. & Neighbours, J. R. (1958). The elastic constants of zinc between 4.2 and 670K. *J Phys Chem Solid* , 7., 58–64.

Alers, G. A. (1965). Use of sound velocity measurements in determining the Debye temperature of solids. In *Physical Acoustics IIIB.* W. P. Mason (Ed.). Academic Press Inc., New York, 1-42

Anil, T. V.; Menon, C. S.; Krishna, K. S. & Jayachandran, K.P. (2005). Low-temperature lattice thermal expansion of the II–VI semiconductor ZnSe. *Journal of Physics and chemistry of Solids,* 66., 11–14

Bardeen, J.; Cooper, L. N. & Schrieffer, J. R. (1957). Microscopic Theory of Superconductivty. *Physical Review*, 106., 162-164

Biwa, S.; Watanabe, Y.; Motogi, S. & Ohno, N. (2004). Analysis of ultrasonic attenuation in particle reinforced plastics by differential scheme. *Ultrasonics*, *43*., 5-12

Blodgett, D. W. & Baldwin, K. C. (2005). Laser-based ultrasonics: applications at APL. *Johns Hopkins APL Technical Digest,* 26., 1., 36-45

Bhatti, S. S. & Singh, A. P. (1989). Attenuation and velocity of ultrasonic waves in strontium borate glasses and their elastic properties. *Acoustica,* 68., 181-182

Bhatia, A. B. (1967). *Ultrasound Absorption: An introduction to the theory of sound absorption and dispersion in gases, liquids and solids,* Oxford Clarendon Press

Bhattachayra, D. K.; Jayakumar, T.; Palanichamy, P. & Raj, Baldev. (1994). Ultrasonic Measurements for Microstructural Characterisation in 17-4 PH Steel Journal of Non-Destructive Evaluation, 13., 15-21

Brugger, K. (1964). Thermodynamic definition of higher order elastic constants. Physical Review, 133., 6., A1611-A1612.

Gomez Alvarez T.E.; Segura L.E.; Franco de Sarabia E. R. (2002). Characterization of suspension of particles in water by an ultrasonic resonant cell. *Ultrasonics*, 39., 715-727

Green, R.E. (1973). *Ultrasonic Investigation of Mechanical Properties*, Academic Press, New York,

Ghate, P. B. (1965). Third order elastic constants of alkali halide. *Physical Review*, 130., 5A., A1666-A1674.

Jackle, J.; Piche, L.; Arnold, W. & Hunklinger, S. (1976). Elastic effects of structural relaxation in glasses at low temperatures. *Journal of Non- Crystalline Solid, 20*., 365-391

Jayakumar, T.; Palanichamy, P. & Raj, B. (1998). Detection of hard inter-metallics in beta-quenched and termally aged Zircaloy-2 using ultrasonic measurements, *Journal of Nuclear Materials*, 255., 243-249.

Kumar, Anish.; Laha K.; Jayakumar, T.; Rao K. B. S. & Raj B. (2001). Imaging of modified 9Cr-1Mo ferritic steel weldment via ultrasonic velocity measurements, *Science & Technology of Welding & Joining*, 6., 6., 383-386

Krautkramer, J. & Krautkramer, H. (1993). *Ultrasonic testing of materials*, Narosa Publishing House, New Delhi, India

Krautkramer, J. & Krautkramer, H. (1977). *Ultrasonic testing of materials*, Springer-Verlag, New York

Lowrance, L. C. (1975). Industrial applications of ultrasound: A review-measurements, test and process control using low intensity ultrasound. *IEEE Transactions on sonics and Ultrasonics,* SU-22., 2., 71-101

Lynnworth, L. C.; Papadakis, E. & Flower K. A., (1977). In *International Advancesin Nondestructive Testing, McGonnagle W.J., Gordon and Breach (Ed.),* Academic Press, New York.

Manghnami, M. H. (1974). Pressure and temperature studies of glass properties related to vibrational spectra report HIG -74-11. *University of Hawaii, Honolubu*

Mason, W. P. (1950). *Piezoelectric crystals and their applications to Ultrasound*. D. Van Nostrand and Co., Princeton

Mason, W. P. (1958). *Physical Acoustics and properties of solids.* D. Van Nostrand Co., Inc, Princeton, N. J., vol. 13 p. 368-373.

Mason, W. P. & Bateman ,T. B. (1964), Ultrasonic wave propagation in doped n-Germanium and p-Silicon. *Physical Review*, 134., 5A, A1387-A1396

Mason W. P. (1965). Effect of Impurities and phonon processes on the ultrasonic attenuation of Germanium, Crystal Quartz and Silicon. In *Physical Acoustics IIIB.* W. P. Mason (Ed.). Academic Press Inc., New York, 235-285.

Mason, W. P. & Rosenberg, A. (1966). Phonon and electron drag coefficient in single crystal Aluminum. *Physical Review*, 151., 2., 434-441.

Mason, W. P. & Rosenberg, A. (1969). Thermal and electronic attenuation and dislocation drag in hexagonal crystal cadmium. *Journal of Acoustical Society America,* 45., 2., 470–475.

Mori, S. & Hiki, Y. (1978). Calculation of third-forth order elastic constants of alkali halide crystal. *Journal of Physical Society Japan*, 45., 1449-1456

Oligschleger, C.; Jones, R. O.; Reimann, S. M. & Schober, H. R. (1996). Model interatomic potential for simulation in selenium. *Physical Review B,* 53., 10., 6165–6173.

Palanichamy, P.; Joseph, A.; Jayakumar, T. & Raj, B. (1995). Ultrasonic velocity measurements for estimation of grain size in austenitic stainless steel. *NDE&E Int,* 28., 3., 179-185

Pandey, D. K.; Yadawa, P.K. & Yadav, R.R. (2007). Ultrasonic properties of hexagonal Zns at nanoscale. *Materials Letters, 61*., 5194-5198

Pandey, D. K.; Singh, D. & Yadav, R.R. (2007). Ultrasonic wave propagation in IIIrd group nitrides. *Applied Acoustics, 68*., 766-777

Pandey, D. K.; Yadawa, P.K. & Yadav, R.R. (2007). Acoustic wave propagation in Laves-phase compounds. *Materials Letters, 61*., 4747-4751

Pandey, D. K. & Yadav, R.R. (2009). Temperature dependent ultrasonic properties of aluminium nitride. *Applied Acoustics, 70.*, 412-415

Pandey, D. K.; Singh, D. & Yadawa, P. K. (2009). Ultrasonic study of Osmium and Ruthenium. *Platinum Metals review*, *53*., 2., 91-97

Papadakis, E. P. (1976). Ultrasonic Velocity and Attenuation: measurements methods with scientific and industrial applications. In *Physical Acoustics XII*, W.P. Mason and N. Thurston (Ed.), Academic Press, New York. 277-374

Roth, D.J.; Cosgriff, L.M.; Martin, R.E.; Verrilli, M.J. & R.T. Bhatt. (2003). Microstructural and defect characterization in ceramic composites using an ultrasonic guided wave scan system. *NASA/TM – 2003-212518,* 1-9

Rajendran V.; Begum, A. Nishara; Azooz M. A. & Batal F. H. El. (2002) Microstructural dependence on relevant physical–mechanical properties on SiO_2–Na_2O–CaO–P_2O_5 biological glasses, *Biomaterials, 23.,21., 4263-4275*

Rosen, M. & Klimker, H. (1970). Low temperature elasticity and magneto-elasticity of dysprosium single crystal. *Physical Review B* , 1., 9., 3748–3756.

Rao, R. R. & Menon, C. S. (1974). Lattice dynamics, third order elastic constants and thermal expansion of Gadolinium. *Phys. Chem.,* 35., 425-432.

Raj, B.; Moorthy V.; Jayakumar, T. & Rao K. B. S. (2003). Assessment of microstructures and mechanical behaviour of metallic materials through non-destructive characterization, *International Materials Reviews* , 48., 5., 273-325.

Raj, B.; Rajendran, V. & Palanichamy. (2004). *Science and Technology of Ultrasonics,* Narosa Publishing House, New Delhi, India

Renolds, W.N. & Wilkinson, S. J. (1978). The analysis of Fibre-reinforced composite materials by measurement of ultrasonic velocity. *Ultrasonics*, 16., 4., 159-163

Singhal, R. L. (2003). *Solid State Physics*, Kedar Nath Ram Nath & co. Publishers, Meerut, India, p. 73-88.

Singh, D.; Yadav, R.R. & Tiwari, A. K. (2002). Ultrasonic attenuation in semiconductors. *Indian Journal of Pure and Applied Physics, 40.*, 845-849

Singh, D.; Pandey, D. K.; Yadawa, P.K. & Yadav, A. K. (2009). Attenuation of ultrasonic waves in V, Nb and Ta at low temperatures. *Cryogenics, 49*., 12-16

Singh, D.; Pandey, D. K. & Yadawa, P.K. (2009). Ultrasonic wave propagation in rare-earth monochalcogenides. *Central European Journal of Physics, 7*., 1., 198-205

Singh, D. & Pandey, D. K. (2009). Ultrasonic investigations in intermetallics. *Pramana-journal of physics, 72*., 2., 389-398

Singh, D. (2009). Behaviour of acoustic attenuation in rare-earth chalcogenides .*Materials Chemistry and physics, 115*., 65-68

Skumiel, A.; Hornowski, T. & Jozefczak, A. (2000). Investigation of magnetic fluids by ultrasonic and magnetic methods. *Ultrasonics*, 38., *864-867*

Skumiel, A.; Jozefczak, A.; Hornowski, T. & Labowski, M. (2003).The influence of the concentration of ferroparticles in a ferrofluid on its magnetic and acoustic properties. *Journal of Physics D:Applied Physics*, 36., *3120-3124*

Skumiel, A. (2004).The effect of temperature on the anisotropy of ultrasound attenuation in ferrofluid. *Journal of Physics D:Applied Physics*, 37., *3073-3079*

Smith, R.L (1987). Ultrasonic materials characterization. *NDT International,* 20.,1., 43-48

Teagle, P. R. (1983). The quality control and non-destructive evaluation of composite aerospace components. *Composites,* 14., 2., 115-128

Temkin S. (1998). Sound propagation in dilute suspension of rigid particles. *J. Acoust. Soc. Am.*, 103(2)., 838-849

Thompson, R. B. (1996). Ultrasonic measurements of mechanical properties. *IEEE Ultrasonic Symposium,* 735-744

Taketomi, S. (1986). The anisotropy of the sound attenuation in magnetic fluid under an external magnetic field. *Journal of Physical Society Japan*, 55., 3., 838-844

Varry, A. (1987). *Material analysis by ultrasonic: metals, ceramics, composites*, Noyes Data Corporation, NJ

Yadav, R. R. & Singh D. (2001). Behaviour of ultrasonic attenuation in intermetallics. *Intermetallics, 49*., 5., 189-194

Yadav, R. R. & Singh D. (2001). Ultrasonic attenuation in Lanthanum Monochalcogenides. *Journal of the Physical Society of Japan, 7*., 6., 1825-1832

Yadav, R. R. & Singh D. (2003). Effect of thermal conductivity on ultrasonic attenuation in Preseodymium Monochalcogenides. *Acoustical Physics, 49*., 5., 595-604

Yadav, R. R. & Pandey, D. K. (2005). Ultrasonic properties at the nanoscale in some metals. *Materials Letters, 59*., 564-569

Yadav, R.R. & Pandey, D. K. (2005). Size dependent acoustical properties of bcc metal. *Acta Physica Polonica A, 107*., 6., 933-946

Yadav, R. R. & Pandey, D. K. (2006). Ultrasonic characterization of gallium nitride. *Materials Research Innovations, 10*., 4., 113-115

Yadav, R. R. & Pandey, D. K. (2006). Ultrasonic properties of β– phase nickel aluminide . *Journal of Pure and Applied Ultrasonics, 28*., 4-11

Yadav, A. K.; Yadav, R. R.; Pandey, D. K. & Singh, D. (2008). Ultrasonic study of fission products precipitated in the nuclear fuel. *Materials Letters,* 62., 3258-3261

Yadawa, P. K.; Singh, D.; Pandey, D. K. & Yadav, R.R. (2009). Elastic and acoustic properties of heavy rare-earth metals. *The Open Acoustics Journal*, 2., 80-86

19

Dissipation of Acoustic Waves in Barium Monochalcogenides

Rajendra Kumar Singh
Department of Physics, Banaras Hindu University, Varanasi-221005, India

1. Introduction

The term *acoustic* refers to a periodic pressure wave. The term includes waves in the audio frequency range as well as those above audio frequency range (ultrasonic and hypersonic) and below the audio frequency range. Acoustic waves are characterized by their speed and absorption. Acoustic absorption is a measure of the energy removed from the acoustic waves by conversion to heat as the wave propagates through a given thickness of material; it has unit dB/cm (or Np/cm). Absorption is a material property, in contrast to attenuation, which includes energy loss due to scattering and reflection as well as and depends upon sample size and experimental configuration.

The elastic and inelastic properties of solids are suitable for the study of acoustic dissipation which account for the direct conversion of acoustic energy into thermal energy. In measurement of the attenuation of acoustic waves in solids using pulse echo method, the attenuation is usually found to be greater than the absorption due to intrinsic dissipation. Acoustic energy is removed from the propagating acoustic wave, but is not immediately converted into heat.

The most important cause of the attenuation is the scattering of acoustics wave from imperfections. In terms of phonon description of acoustic waves, this is a two-phonon process, in which incoming and outgoing phonons have different wave vectors. The perturbation at the scattering centre may be due to a mass difference of an impurity atom from the normal mass or to a change in interatomic forces.

In polycrystalline solids, sound is scattered from the boundaries between the microcrystal grains. The grain boundaries act as scattering centers due to the discontinuity of the elastic constants, and the amount of loss depends on the grain size and on the wavelength of the acoustic wave. When the wavelength of the acoustic wave is small compared to the grain size, the loss is independent of frequency and inversely proportional to the mean grain diameter. When the wavelength is large compared to the grain size, the loss is proportional to the fourth power of the frequency (Rayleigh scattering) and to the third power of grain diameter.

Another source of nondissipative loss is diffraction of acoustic field from the transducer. This is an important loss mechanism in megahertz frequency range, but at higher frequencies it is negligible. At higher frequencies, a loss mechanism occurs due to lack of flatness and parallelism of the end faces of the specimen. At higher frequencies (10 GHz), the wavelength of the sound in a solid is of the order of optical wavelength in visible range and hence, the surface of the specimen should be polished with optical quality.

Various causes can be attributed to the dissipation of acoustic waves propagating in different types of solids. These causes depend primarily on the physical conditions of the material under investigation. Having control over the physical conditions of the material, one cause can be studied eliminating others. Most of the energy from the propagating acoustic wave through the medium is absorbed and converted into heat. Following causes may be attributed to the attenuation of the acoustic wave propagating through a solid;
(a) Electron-phonon interaction, (b) Phonon-phonon interaction, (c) Lattice imperfection, (d) Thermoelastic loss (e) Ferromagnetic and Ferroelectric losses and (f) NMR and Thermal relaxation etc.. In non conducting non-ferromagnetic solids at 50 K and above, phonon-phonon interaction is the principal cause of acoustical dissipation.
Among the wide band gap II-IV semiconductors, the barium chalcogenides [BaX, X=S, Se, Te] are interesting in connection with optoelectronic applications in blue light wavelength regime. The Barium Chalcogenides form very important closed shell ionic systems crystallized in the NaCl (B1) type and CsCl (B2) type structures at ambient conditions. Alkaline earth chalcogenides are currently under intense investigations driven by their applications in light emitting diodes (LEDs) and laser diodes (LDs). It is expected that these compounds may provide new II-IV candidates for the fabrication of various electrical and optical devices [Charifi et al. (2005) and Bouhemadou et al. (2006)].
Experimental as well as theoretical work on different aspects of these compounds has been reported in the recent past [Charifi et al. (2005), Bouhemadou et al. (2006), Hassan and Akbarzadeh (2006) and Cervantes et al (1998)]. However, results on temperature dependent acoustical behaviour of these chalcogenides viz. acoustical dissipation due to phonon-phonon interaction, thermoelastic loss, dislocation damping, Gruneisen parameter, non-linearity parameters and thermal relaxation time etc, which are very important parameters necessary to explain the microstructure and other related physical properties of these chalcogenides have not been studied. Recently, we studied in detail [Singh and Singh 2010] acoustical behaviour of these compounds starting from second and third order elastic constants (obtained at different temperatures), which were used to evaluate Gruneisen parameters and non-linearity parameters along different crystallographic directions viz. <100>, <110> and <111> for longitudinal and shear modes in the temperature range 50K-500 K. Taking electrostatic and Born repulsive potentials and utilizing some parameters viz. nearest neighbour distance, hardness parameter and specific heat as a function of Debye temperature; acoustical dissipation coefficients were obtained at different temperatures.

2. Absorption of acoustic waves by thermal phonons

The anharmonic interactions among phonons in a solid are responsible for attenuation of ultrasonic waves, and are particularly important in insulators where absorption due to free electrons is absent. Also, when a longitudinal wave propagates in a crystalline solid, compression and rarefaction is produced and heat is transmitted from compressed part to rarefied parts and dissipation of acoustic waves occurs. Dislocation damping due to screw and edge dislocations also produces appreciable loss in solids.

2.1 Phonon-phonon interaction

In perfect, insulating, non-ferromagnetic and non-ferroelectric substances, dissipation of acoustical energy occurs mainly due to phonon-phonon (p-p) interaction and thermoelastic

loss. Akhiezer (1939) was first to propose the phonon-viscosity mechanism for acoustical dissipation, but he did not include the finite value of relaxation time for thermal equilibrium process. Bommel and Dransfeld (1960) later took this work considering the relaxation time to be finite. They obtained results comparable to the experimental results. Further, Woodruf and Ehrenreich (1960) used the Boltzmann equation method to evaluate the steady state distribution of thermal phonons and acoustical attenuation. They considered the N (normal) and U (Umklapp) processes. But due to insufficient information regarding parameter ' γ ', used, Mason (1965) used Gruneisen constant (γ_i^j), which is related to second and third order elastic constants and this approach is found to be very useful for the estimation of ultrasonic attenuation in various crystals.

At room temperature and in a wide temperature region also, thermal phonon relaxation time, τ_{th}, varies from 10^{-10} sec to 10^{-12} sec from metallic to dielectric crystals. As temperature increases, τ_{th} decreases. Hence, condition $\omega\tau_{th} << 1$ holds good and at the same time the individual phonon looses its significance and idea of the phonon gas having macroscopic parameter is described. In the Akhiezer regime (ωτ « 1), a sound wave passing through a solid can be attenuated by two processes. First, if the wave is longitudinal, periodic contractions and dilations in the solid induce a temperature wave via thermal expansion. Energy is dissipated by heat conduction between regions of different temperatures. This is called thermoelastic loss. Second, dissipation occurs as the gas of thermal phonons tries to reach an equilibrium characterized by a local (sound wave induced) strain. This is internal friction mechanism.

The physical basis for obtaining attenuation coefficient is that the elastic constants contributed by thermal phonons relax [Bommel and Dransfield (1960), Pippard (1955) and Mason (1955)]. The phonon contribution to the unrelaxed elastic constants is evaluated by taking into consideration the change in energy of the thermal phonons due to applied instantaneous strain. The frequency of each mode ν_i is changed by $\frac{\partial \nu_i}{\nu_i} = -\gamma_i^j S_j$, where γ_i^j is generalised Gruneisen parameter & S_j is instantaneous strain. It is assumed that all the phonons of a given direction of propagation and polarization have equal change in frequency. Then phonons of ith branch and jth mode suffer a change in temperature $\frac{\Delta T_i}{T_0} = -\gamma_i^j S_j$ (T is the temperature). A relaxed elastic constant is obtained after there is phonon-phonon coupling among various branches and ΔT_i relax to a common temperature change, ΔT given by $\frac{\Delta T}{T} = -\langle\ \gamma_i^j\ \rangle S_j$; where $\langle\ \gamma_i^j\ \rangle$ is the average value of γ_i^j.

The sudden application of acoustical pressure to a body at temperature T causes different temperature increments for different phonon modes, which relax back to new equilibrium at a temperature $T + \Delta T$ through the phonon-phonon collision. This temperature difference lags behind the periodic stress and causes a relaxational absorption.

The relation between the attenuation and ΔC_e (change in the elastic constant due to non-equilibrium temperature separation of the phonon modes by the applied strain) is given as:

$$\alpha = \Delta C_e \omega^2 \tau_{th} / 2dV^3(1 + \omega^2 {\tau_{th}}^2) \quad (1)$$

Where α the is attenuation in dB/cm, d is density, ω is angular frequency of the ultrasonic wave and V is the velocity of the wave.

When the strain S_j is applied to the crystal, there is change in mode frequency given by:

$$\omega_i = \omega_{i0}(1 - \sum_{j=1} \gamma_i^j S_j) \tag{2}$$

ω_{i0} is frequency of the mode in the standard state. By measurements of SOEC and TOEC, it could be predicted that γ_i^j do not vary much. When the above expression is differentiated, one obtains:

$$\gamma_i^j = (\partial\omega_i / \partial S_j) / \omega_{i0} \tag{3}$$

γ_i^j is known as Gruneisen number. A general formula for γ_i^j has been given by Brugger in terms of tensor notation:

$$-\gamma_i^j = -\gamma_i^{jk} = U_i U_j + N_p N_q (C_{jkpq} + U_r U_s C_{jkpqrs}) / 2C_e \tag{4}$$

Where jk are the two index symbols for strain S_j. N_p and N_q are the direction cosines for the propagation direction and C_e is the required elastic constant determined by the type of the wave and the direction of propagation. U_j and U_k are the direction cosines for the particle displacements. C_{jkpq} and C_{jkpqrs} are the second and third order elastic constants in tensor notations. Now a suddenly applied strain neither changes the number of modes nor their entropy. Mason considered thermal energy of the modes under Debye's approximation,

$$U_{th} = 3\hbar \sum_i (N_i / {\omega_{gi}}^2) \int_0^{\omega_{gi}} (\omega^2 / (\exp(\hbar\omega / kT) - 1)) d\omega \tag{5}$$

on differentiation of the sum of the elastic energy plus the total thermal energy of all modes, one obtains:

$$T_j = \partial U_{th} / \partial S_j \tag{6}$$

$$= {C_{ij}}^S S_j + 3\hbar(\partial / \partial S_j)(\sum ({N_i} / {\omega_{gi}}^3) \int_0^{\omega_{gi}} (\omega^3 / (\exp(\hbar\omega / kT) - 1)) d\omega$$

and finally one gets:

$$T_j = {C_{ij}}^S + 3\sum_i E_i (\gamma_i^j)^2 S_j + 3\sum_i E_i \gamma_i^j \tag{7}$$

where T_j is the stress associated with the strain S_j, C_{ij} is the corresponding elastic constants resulting from no entropy exchange between any of the modes and γ_i^j is the Gruneisen number. E_i is the thermal energy associated with each direction and each mode. The above expression, shows that elastic constant changes by

$$\Delta C_e = 3\sum_i E_i (\gamma_i^j)^2 \tag{8}$$

This development is valid for shear modes for which the average rise in temperature is zero. For longitudinal modes, the increase in modulus resulting from the difference between the adiabatic and isothermal conditions is to be supported and it is given by:

$$\Delta C_e = (3\sum_i E_i(\gamma_i^j) - \gamma^2 CT) \tag{9}$$

Replacing the value of ΔC_e , one gets:

$$\alpha = ED\omega^2\tau_{th} / 6dV^3(1+\omega^2{\tau_{th}}^2) \tag{10}$$

for

$$\omega\tau_{th} << 1 \text{ , } \alpha = ED\omega^2\tau_{th} / 6dV^3 \tag{11}$$

Equation (1) reduces to

$$(\alpha / f^2)_l = 2\pi^2 E_0(D_l / 3)\tau_l / dV_l^3 \tag{12}$$

and

$$(\alpha / f^2)_S = 2\pi^2 E_0(D_S / 3)\tau_S / dV_S^3 \tag{13}$$

for longitudinal and shear waves, respectively.

$$\text{where} \quad D = 9 < (\gamma_i^j)^2 > -(3CT < (\gamma_i^j) >^2 / E) \tag{14}$$

Here D is the non-linearity constant. Mason and co-workers (1964) have obtained a number of tables in terms of second and third order elastic constants to calculate $< (\gamma_i^j) >^2$ and $< (\gamma_i^j)^2 >$ for different directions of propagation and polarization. Gruneisen numbers along different directions of propagation viz. <100>, <110> and <111> can be obtained using Mason (1965) approach.

Thermal relaxation time, τ (subscripts l and s for longitudinal and shear waves) is given as,

$$\tau = \tau_s = \frac{\tau_l}{2} = \frac{3K}{C_v < V >^2} \tag{15}$$

Where K is thermal conductivity, C_v is specific heat per unit volume and $< V >$ is Debye average velocity given by

$$\frac{3}{\left(< V >\right)^3} = \frac{1}{V_L^3} + \frac{2}{V_S^3} \tag{16}$$

The Debye temperature is given by [Jasiukiewicz & Karpus (2003)],

$$\Theta_D = \hbar <V> q_d / K_B \tag{17}$$

K_B is Boltzmann constant and

$q_d = (6\Pi^2 N_a)^{1/3}$ where N_a is atom concentration

According to Mason and Batemann [1964], SOEC and TOEC are related by Gruneisen parameter γ_i^j and hence by non-linearity parameter, D. $\left\langle \left(\gamma_i^j\right)^2 \right\rangle$ and $\left\langle \gamma_i^j \right\rangle^2$ are square average & average square Gruneisen parameters, V is sound wave velocity (V_l) for longitudinal wave and (V_s) for shear wave and d is density.

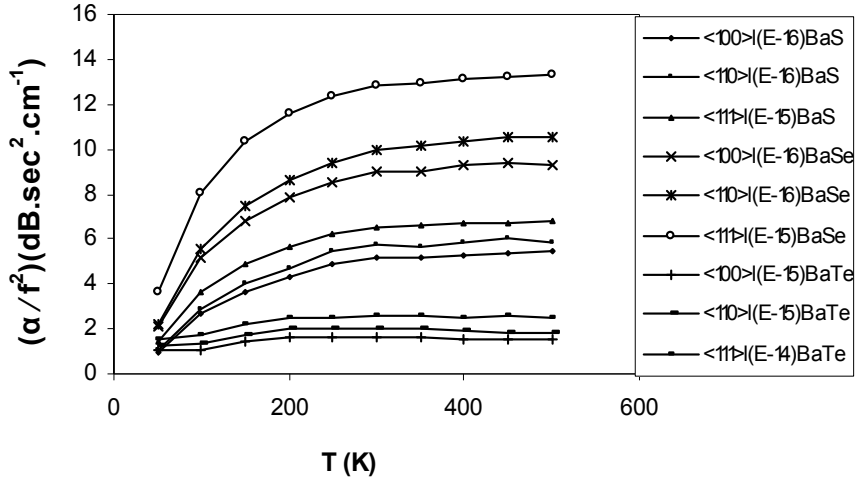


Fig. 1.Temperature variation of $(\alpha/f^2)_l$ along different directions.

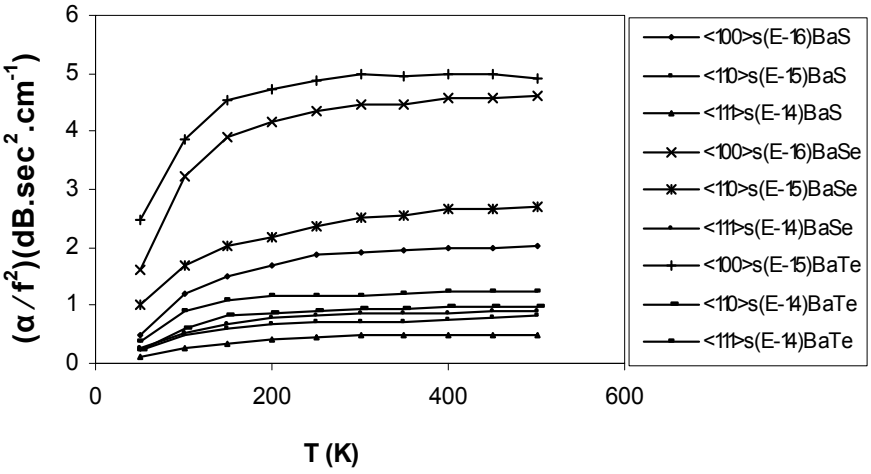


Fig. 2. Temperature variation of $(\alpha/f^2)_s$ along different direction directions.

The ultrasonic attenuation due to phonon-phonon interaction for longitudinal, $(\alpha/f^2)_l$ and shear waves, $(\alpha/f^2)_s$ are evaluated using equations (12) and (13), respectively. Typical attenuation versus temperature curves ($(\alpha/f^2)_l$ and $(\alpha/f^2)_s$ vs Temperature) along [100], [110] and [111] directions of propagation are shown in Figs. (1-2) , and it can be seen that the temperature dependence divides into two regions. Region 1, (upto Debye temperature of respective solids, which has been shown in Table 1) attenuation coefficient varies rapidly and in Region II, attenuation coefficient becomes temperature independent. To understand the physical processes involved, it is helpful to consider region I and II separately. When $\omega\tau_{th}$ <1 (Region 2), where ω is the acoustic frequency and τ_{th} is the mean lifetime of thermal phonons, the phonon mean free path is short compared to the acoustic wavelength and phonons see a very gradual spatial gradient of the acoustic starin. In the opposite extreme ($\omega\tau_{th}$ >1), the phonon mean free path is long compared to the acoustic

wavelength, and the acoustic wave is best thought of as a beam of coherent phonons which are on an actual footing with the thermal phonons. The mode of interaction is then by phonon-phonon interaction.

Compound	Θ_D (K)	<V> (10^5cm/sec)	M (Mol. Weight)
BaS	200	4.18	169.39
BaSe	170	2.61	216.28
BaTe	143	2.39	264.92

Table 1. Debye temperature (Θ_D) and average Debye velocity (<V>) at 300 K

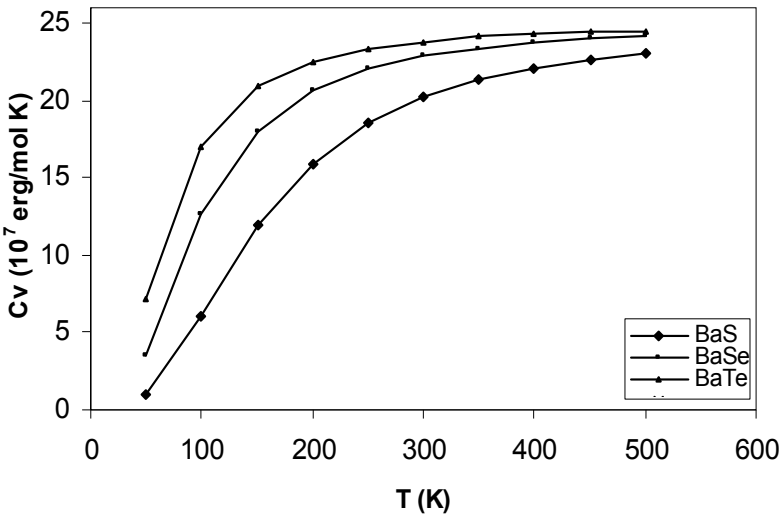


Fig. 3.Temperature variation of specific heat (C_v)

The attenuation due to phonon-phonon interaction for longitudinal and shear waves $(\alpha/f^2)_l$, $(\alpha/f^2)_s$ increases up to Θ_D and then becomes constant. When $(\Theta_D/T) \geq 1$, $(\alpha/f^2)_l$ and $(\alpha/f^2)_s$ increase and for values of temperatures satisfying $(\Theta_D/T) < 1$, attenuation becomes nearly constant, because $(\alpha/f^2)_l$ or $(\alpha/f^2)_s$ due to p-p interaction is mainly affected by the specific heat, C_v (since (α/f^2) due to phonon-phonon interaction is related to C_v, (through the relaxation time). For $(\Theta_D/T) \geq 1$, C_v increases and becomes nearly constant for the values satisfying $(\Theta_D/T) < 1$, (Fig. 3).

The value (α/f^2) at a given temperasture is minimum for BaS and maximum for BaTe. The value of (α/f^2) depends upon Debye temperature (Θ_D). The Debye temperature is maximum for BaS and minimum for BaTe, (Table 1). Thus greater the Θ_D value, smaller is the attenuation. The value of Θ_D depends on the Debye average velocity <V> and inverse of cube root of molecular weight i.e. $M^{-1/3}$ through $(N/V)^{1/3}$ where N is the Avogadro number and V (V = M/d, M= mol. wt. and d= density) is volume. <V> is maximum for BaS and minimum for BaTe, therefore larger is the <V>, smaller will be attenuation. The attenuation increases in these chalcogenide series with increasing the Molecular weight. The Θ_D and <V> are SOEM dependent. Thus the increase in the value of $(\alpha/f^2)_l$, $(\alpha/f^2)_s$ and $(\alpha/f^2)_{th}$ from BaS to BaTe is mainly influenced by SOEM values and Molecular weight.

2.2 Thermoelastic loss

In an isotropic polycrystalline solid, strain varies according to applied stress from one grain to another (Lucke, 1956). The substance is isotropic due to random orientation of grains. The individual grain may be anisotropic. The propagation of longitudinal wave creates compression and rarefactions throughout the crystal. The rarified regions are cooler than compressed regions and hence there is a flow of heat between the two regions and the direction of flow of this energy will be reversed after every half cycle. Since there is a relaxational phenomenon, there is a loss of energy. Attenuation due to this effect is given by (Mason, 1965)

$$\alpha = 2\pi^2 f^2 K(C_{nn}{}^{\sigma} - C_{nn}{}^{\theta}) / dV^3 C_V C_{nn} \tag{17}$$

where $C_{nn}{}^{\sigma}$ and $C_{nn}{}^{\theta}$ are adiabatic and isothermal elastic constants, K is thermal conductivity and C_V is specific heat per unit mass. The difference between $C_{nn}{}^{\sigma}$ and $C_{nn}{}^{\theta}$ may be obtained with the help of SOEC and TOEC. This loss does not make any appreciable contribution to the total ultrasonic attenuation in case of dielectric and semiconducting crystals due to low value of thermal conductivity. In case of metals thermal conduction arises due to electronic and lattice contribution so it is large enough to cause appreciable contribution to the total ultrasonic attenuation. For shear wave propagation no compression or rarefaction occurs hence no thermoelastic loss. Propagation of sound wave through crystal produces compression and rarefactions as a result heat are transmitted from compressed region (at higher temperature) to rarefied region (at lower temperature) and hence thermoelastic loss occurs, which is given by.

$$\alpha_{th} = \frac{4\pi^2 f^2 \left\langle \gamma_i{}^j \right\rangle^2 KT}{2dV_L{}^5} \tag{18}$$

Ultrasonic attenuation due to this effect has also been evaluated in case of barium monochalcogenides and is given in Fig. 4..

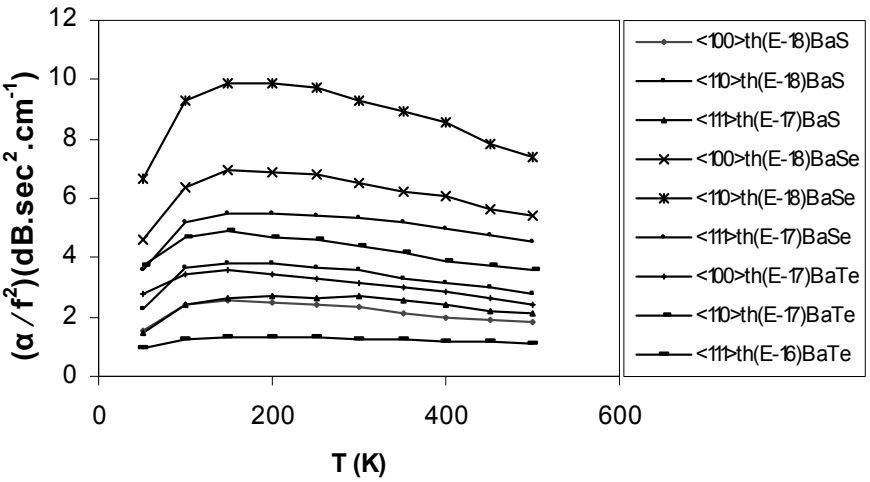


Fig. 4. Temperature variation of $(\alpha/f^2)_{th}$ along diffrent direction

$(\alpha/f^2)_{th}$ is directly proportional to rate of heat transfer from compressed regions to rarefied regions. In the low temperature range, 50-200 K, heat is transferred at faster rate from compressional regions to the rarefied regions resulting larger rate of thermoelastic loss. The rate of increase of thermoelastic loss is small beyond 200 K.

2.3 Phonon processes and drag on dislocations

A dislocation is a linear imperfection in a crystal. In edge dislocation, near the dislocation line, the crystal is severely strained. In a screw dislocation, Burger vectors are parallel to the dislocation line. In general, a dislocation is composed of mixtures of screw and edge dislocations. Another process for which thermal losses due to p-p interaction can produce an appreciable effect is the drag on disclocations as they are moved through a lattice. Leibfried et al. (1954) discussed the mechanism of scattering of phonons by moving disclocations and the results show that the resulting differential produces a drag force which is proportional to the velocity of the disclocation. Mason (1965) proposed a theory to explain the mechanism involved in the drag produced on a dislocation by phonon-viscosity. This was evaluated on the basis of the effect caused by the change in dimensions of phonon modes and their subsequent equilibrium through a thermal relaxation process.

Dislocation damping due to screw and edge dislocations also produces appreciable loss due to phonon-phonon interaction. The loss due to this mechanism can be obtained by multiplying dislocation viscosities by square of dislocation velocity. Dislocation damping due to screw and edge dislocations is given by equations (21) and (22).

The Phonon-viscosity, which is analogous to shear-viscosity in liquids damps the motion of both type (screw and edge) disclocations and has the value

$$\eta = EDk / C < V >^2 = ED\tau_{th} / 3 \tag{19}$$

These phonon-viscosities are presented in the form of drag coefficients for the motion of screw and edge type of disclocations. Here the Cortell's (Cortell, 1963) condition $a_0 = 3b/4$ is valid, where a_0 the disclocation core radius and 'b' is is the Brugger's vector. B_{screw} and B_{edge} are given by

$$B = b^2 / 8\pi a^2 \tag{20}$$

substituting $a_0 = 3b/4$ the above equation reduces to,

$$B_{screw} = 0.071\eta \tag{21}$$

and

$$B_{edge} = (0.0532\eta + 0.0079(\mu / K)^2 \chi / (1-\sigma)^2 \tag{22}$$

where σ, μ, K and χ are Poisson's ratio, shear modulus, bulk modulus and compressional viscosity respectively. These values can be calculated using the relations

$$\mu = (C_{11} - C_{12} + C_{44}) / 3 \,,\; K = (C_{11} + 2C_{12}) / 3 \,,$$

$$\text{and } \chi = (4/3\eta_l - \eta_S) \tag{23}$$

Compound	B_{screw}		B_{edge}	
	Long.	Shear	Long.	Shear
BaS	0.23	0.10	0.45	0.55
BaSe	0.29	0.17	0.60	0.79
BaTe	0.47	1.30	1.07	3.22

Table 2. Phonon viscosity due to screw and edge dislocation at 300K longitudinal (in cp) and shear (in mp.) waves.

Debye average velocity and Debye temperature have been calculated using equations (16) and (17) and are presented in Table 2.

Square average Gruneisen numbers $< \gamma_i^{j2} >_l$ and $< \gamma_i^{j\ 2} >_{s^*}$ and average square Gruneisen parameter $< \gamma_i^{j} >^2_l$ and $< \gamma_i^{j} >^2_s$ and $< \gamma_i^{j} >^2_{s^*}$ for longitudinal and shear waves, nonlinearity coupling constants D_l, D_s, D_s^* and their ratios D_l/D_s, and D_l/D_s^* along different directions of propagation are given in Table 3. Results are as expected [Mason (1967), Kor and Singh (1993)].

Compound	Direction	$< \gamma_i^{j2} >_l$	$< \gamma_i^{j} >^2_l$	$< \gamma_i^{j} >^2_s$	$< \gamma_i^{j} >^2_{s^*}$	D_l	D_s	D_{s^*}	D_l/D_s	D_l/D_{s^*}
BaS	100	0.94	0.17	0.04	--	7.82	0.37	--	20.81	--
	110	1.06	0.26	0.15	1.93	8.63	1.43	17.37	6.03	0.49
BaSe	100	0.90	0.24	0.04	--	7.27	0.43	--	16.79	--
	110	1.04	0.36	0.22	1.80	8.04	1.98	16.28	4.06	0.49
BaTe	100	1.68	1.28	0.30	--	10.63	2.73	--	3.88	--
	110	2.14	1.75	4.49	1.33	12.93	40.46	12.00	0.31	1.00

Table 3. Square Average and average square Gruneisen number for longitudinal $< \gamma_i^{j2} >_l$, $< \gamma_i^{j} >^2_l$ and shear $< \gamma_i^{j} >^2_s$, $< \gamma_i^{j} >^2_{s^*}$ Waves, nonlinearity coupling constants D_l, D_s and nonlinearity coupling constants ratios D_l / D_s, D_l / D_s^* at 300K

$_l$ for longitudinal wave

$_s$ for shear wave, polarized along [001]

$_{s^*}$ for shear wave, polarized along $[1\bar{1}0]$

Viscous drag due to screw (B_{screw}) and edge dislocations have been obtained (B_{edge}) using equation (21) and (22), as given in Table 2.

The phonon mean free path due to phonon-phonon collision is a rapidly changing function of temperature at low temperatures. Fig. 4 shows the τ_{th} vs T plot for barium monochalcogenides.. Thermal relaxation time is evaluated using equation (6). Temperature variation of thermal relaxation time is shown in Fig. 4 which shows exponential decay according to relation $\tau = \tau_o \exp(-t/T)$, where τ_o and t are constants.

From the values of thermal relaxation time, it can be seen that the condition $\omega\tau_{th} << 1$ is satisfied even at GHz range acoustic wave frequency.

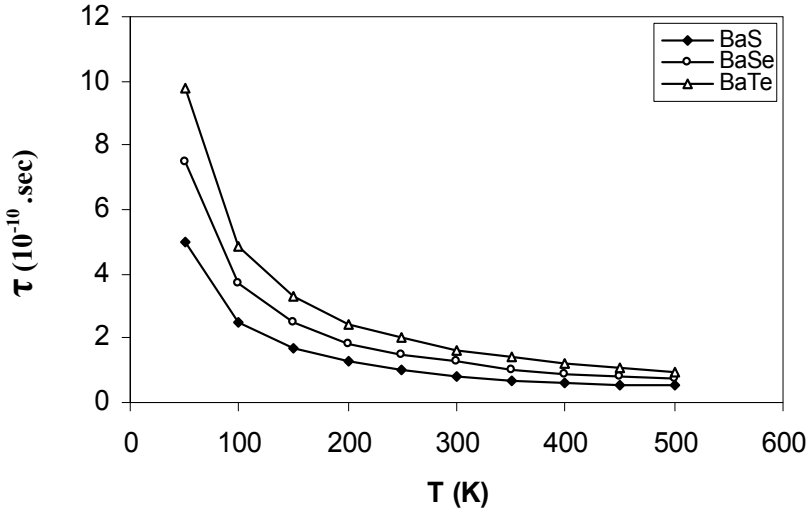


Fig. 4. Temperature variation of thermal relaxation time (τ).

3. Conclusions

Acoustical dissipation and related parameters have been evaluated over a wide temperature range using simple approach and starting from second and third order elastic constants. These values of second and third order elastic constants have been used to obtain acoustical Gruneisen parameters and non-linearity coupling constants. Utilizing values of non-linearity coupling constants, ultrasonic arttenuation due to phonon-phonon interaction, thermoelastic loss and dislocation dampming due to screw and edge dislocations have been obtained over a wide temperature range. In the present approach, Grunesen parameters have been evaluated for longitudinal and shear modes by considering only finite number of modes (39 modes for longitudinal wave while 18 modes for shear waves). However, a more rigorous approach is needed, in which all possible phonon modes can be incorporated.

4. Acknowledgements

I am thankful to the University Grants Commission, New Delhi (Government of India) for financial assistance.

5. References

Akhiezer, A., Absorption of sound in metals, *J. Phys. (USSR*), 1 (1939) 289-298.

Bouhemadou, A, Khenata, R., Zegrar, F., Sahnoun, M, Baltache, H., Resh, A.H, *Computational Material Science* 38, 263 (2006)

Bommel H. E and Dransfield , K. Excitation and attenuation of hypersonic waves in quartz , *Phys. Rev.* 177, 145 (1960).

Bommel, H.E. and Dransfeld, K., *Phys. Rev*., 117 (1960) 245.

Breazeale, M.A. and Philip, J., *J. Phys. (Colloq)*, 42 (1981) 134.

Brugger, K, Phys. Rev. A. 133, 1611 (1964)

Charifi Z, Baoziz H., Hassan, F El Haj and Bouarissa, N, *J. Phys. Condens. Matter 17*, 4083 (2005)

Cervantes, P., Williams, Q, Cote, M, Rohlfing, M, Cohen, M. L. and .Louie, S G *Phys. Rev*. B, 58 (15) 9793 (1998)

Elmore, P.A. and Breazeale, M.A., Dispersion and frequency dependent nonlinearity parameters in a graphite–epoxy composite, *Ultrasonics*, 41 (2004) 709-718.

Fabian, J. and Allen, P.B., Theory of sound attenuation in Glasses: The role of thermal vibrations, *Phys. Rev. Let*., 82 (7) (1999) 1478-1481.

Ghate, P.B., Third order elastic constants of Alkali halide crystals, *Phys. Rev*., 139 (1965) A1666-A1674.

Ghate , P.B., Phys. Rev. B 139 (5A) A1666. (1965)

Hassan, F. El. Haj and Akbarzadeh, H.Computational and Material *Science*. 38, 362 (2006)

Leibfried, G. and Hahn, H., Temperature dependent elastic constants of alkali halides, *Z. Physik*, 150 (1958) 497-525.

Ludwig, W. and Leibfried, G., Theory of anharmonic effects in crystals, Solid State Physics, Academic Press New York, 12 (1967).

Mason, W.P., Ultrasonic attenuation due to lattice-electron interaction in normal conducting metals, *Phys. Rev*., 97 (1955) 557-558.

Mason, W.P. and Bateman, T.B., Ultrasonic wave propagation in pure Si and Ge, *J. Acoust. Soc. Am*., 36 (1964) 645.

Mason, W.P., Effect of impurities and phonon processes in the ultrasonic attenuation of germanium crystal, quartz and silicon, Physical Acoustics, Academic Press New York, IIIB (1965) 237.

Mason, W.P., Relation between thermal ultrasonic attenuation and third order elastic moduli for waves along <110> axis of a crystal, *J. Acoust. Soc. America*, 42 (1967) 253.

Mason, W.P. and Rosenberg, A., Thermal and electronic attenuations and dislocation drag in the hexagonal crystal Cadmium, *J. Acoust. Soc. America*, 45 (2) (1969) 470-480.

Pippard, A.B., Ultrasonic attenuation in metals, *Philos. Mag*., 46 (1955) 1104.

Singh, R.K., Singh R. P. Singh and Singh M.P.., *Proc. 19th International Congress on Acoustics (ICA-2007), Spain (Madrid*) 2007.

Singh R. K. *, Singh R. P., Singh M. P., and Chaurasia, S. K., *Acoustic Wave Propagation in Barium Monochalcogenides in the B1 Phase Acoustical Physics, 2009*, Vol. 55, No. 2, pp. 186–191.

Woodruff, R.O. and Ehrenreich, H., *Phys. Rev*., 123 (1962) 1553.

20

Statistical Errors in Remote Passive Wireless SAW Sensing Employing Phase Differences

Y.S. Shmaliy, O.Y. Shmaliy, O. Ibarra-Manzano,
J. Andrade-Lucio, and G. Cerda-Villafana
Electronics Department, Guanajuato University
Mexico

1. Introduction

Passive remote wireless sensing employing properties of the surface acoustic wave (SAW) has gained currency during a couple of decades to measure different physical quantities such as temperature, force (pressure, torque, and stress), velocity, direction of motion, etc. with a resolution of about 1% [1]. The basic principle utilized in such a technique combines advantages of the precise piezoelectric sensors [2, 3, 4], high SAW sensitivity to the environment, passive (without a power supply) operation, and wireless communication between the sensor element and the reader (interrogator). Several passive wireless SAW devices have been manufactured to measure temperature [1], identify the railway vehicle at high speed [5], and pressure and torque [6].

The information bearer in such sensors is primarily the time delay of the SAW or the central frequency of the SAW device. Most passive SAW sensors are designed as reflective delay lines with M reflectors[1] and operate as sketched in Fig. 1. At some time instant $t_0 = 0$, the reader transmits the electromagnetic wave as an interrogating radio frequency (RF) pulse ($K = 1$), pulse burst ($K > 1$), pulse train, or periodic pulse burst train. The interdigital transducer (IDT) converts the electric signal to SAW, and about half of its energy distributes to the reflector. The SAW propagates on the piezoelectric crystal surface with a velocity v through double distances ($2L_1$ and $2L_2$), attenuates (6 dB per μs delay time [5]), reflects partly from the reflectors (R_1 and R_2), and returns back to the IDT. Inherently, the SAW undergoes phase delays on the piezoelectric surface. The returned SAW is reconverted by the IDT to the electric signal, and retransmitted to the interrogator. While propagating, the RF pulse decays that can be accompanied with effects of fading. At last, K pairs of RF pulses (Fig. 1b) appear at the coherent receiver, where they are contaminated by noise. In these pulses, each inter distance time delay $\Delta\tau_{(2k)(2k-1)} = 2(L_2 - L_1)/v$, $k \in [1,K]$, bears information about the measured quantity, i.e., temperature [1], pressure and torque [5], vehicle at high speed [6], etc.

To measure $\Delta\tau_{(2k)(2k-1)}$, a coherent receiver is commonly used [7], implementing the maximum likelihood function approach. Here, the estimate of the RF pulse phase relative to the reference is formed to range either from $-\pi/2$ to $\pi/2$ or from $-\pi$ to π by, respectively,

[1] Below, we consider the case of $M = 2$.

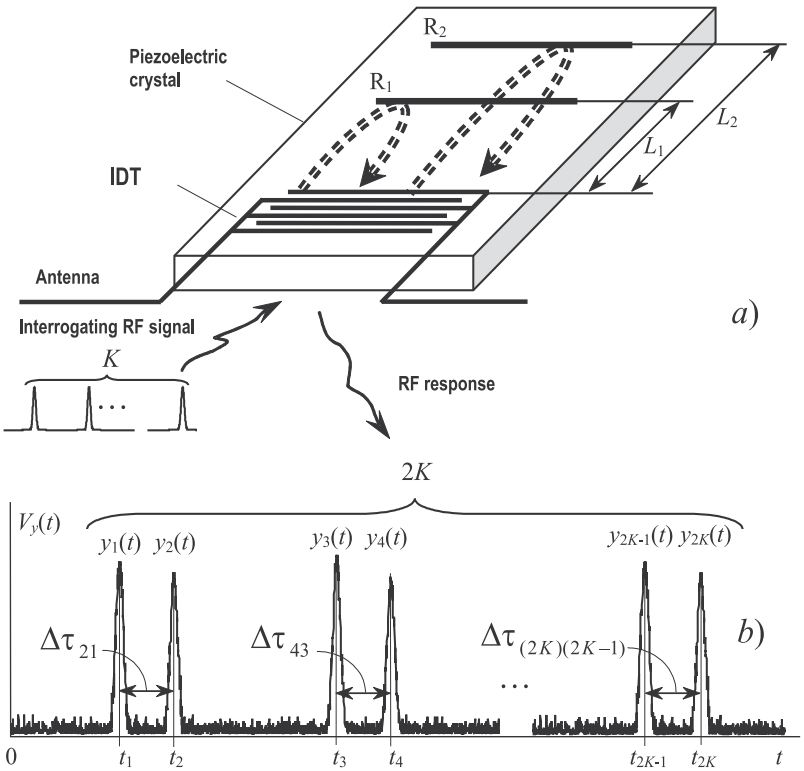


Fig. 1. Operational principle of remote SAW sensing with phase measurement: a) basic design of passive SAW sensors and b) reflected pulses at the coherent receiver detector [25].

$$\hat{\theta} = \arctan\frac{Q}{I}, \tag{1}$$

$$\hat{\theta} = \begin{cases} \arctan(Q/I), & I \geq 0 \\ \arctan(Q/I) \pm \pi, & I < 0, \begin{cases} Q \geq 0 \\ Q < 0 \end{cases} \end{cases}, \tag{2}$$

where I and Q are the in-phase and quadrature phase components obtained for the received pulse. With differential phase measurement (DPM), the phase difference in every pair of pulses is calculated by

$$\hat{\Theta}_k = \hat{\theta}_{2k} - \hat{\theta}_{2k-1} \tag{3}$$

and used as a current DPM. Here several estimates may be averaged to increase the signal-to-noise ratio (SNR) [7]. Averaging works out efficiently if the mean values are equal. Otherwise, the differential phase diversity is of interest to estimate either the vehicle's velocity (Doppler shift) or the random error via

$$\hat{\Psi}_k = \hat{\Theta}_k - \hat{\Theta}_{k-1}. \tag{4}$$

An accurate estimate $\hat{\Theta}_k$ is a principal goal of the receiver. To obtain it with a permitted inaccuracy in the presence of noise, the interrogating signal must be transmitted with a sufficient peak power that, however, should not be redundant. The peak power is coupled with the SNR. Therefore, statistical properties of $\hat{\Theta}_k$ and $\hat{\Psi}_k$ are of prime interest. Knowing these properties and the peak power of the interrogating pulse, one can predict the measurement error and optimize the system. In this Chapter, we discuss limiting and approximate statistical errors in the estimates (3) and (4).

2. Signal model

For SAW sensors with identification marks, the readers are often designed to interrogate the sensors with a linear frequency modulated (LFM) RF impulse request signal [8, 9]

$$x(t) = \sqrt{2S}a(t)\cos\left(2\pi f_0 t + \frac{\alpha t^2}{2} + \theta_0\right), \tag{5}$$

where $2S$ and θ_0 are the peak-power and initial phase, respectively, f_0 is the initial carrier frequency, and t is the current time. The LFM pulse can have a near rectangular normalized waveform $a(t)$ of duration T such that $\alpha = \Delta\omega/T$, where $\Delta\omega$ is a required angular frequency deviation, overlapping all the sensor responses.

It turns out that noise does not perturb $x(t)$ substantially in the sensor. Therefore, assuming Gaussian envelope in the reflected pulses, the induced SAW reflected from the reflectors R_1 and R_2 and then reconverted and retransmitted can be modeled with, respectively,

$$\begin{aligned} s_1(t) &= \sum_{k=1}^{K} u_{2k-1}(t) \\ &= \sqrt{2S}\sum_{k=1}^{K} \beta_{2k-1}(t) e^{-b^2(t-t_{2k-1})^2} \cos[2\pi f_{2k-1}t + \vartheta_{2k-1}(t)], \end{aligned} \tag{6}$$

$$\begin{aligned} s_2(t) &= \sum_{k=1}^{K} u_{2k}(t) \\ &= \sqrt{2S}\sum_{k=1}^{K} \beta_{2k}(t) e^{-b^2(t-t_{2k})^2} \cos[2\pi f_{2k}t + \vartheta_{2k}(t)], \end{aligned} \tag{7}$$

where $\beta_i^2(t)$, $i \in [1,2K]$, is a normalized instantaneous power caused by attenuation and fading. The full phase shifts relative to the carrier and its constituent induced during the SAW propagation are given by, respectively,

$$\vartheta_{2k-1} = \phi_{2k-1} - \psi_{2k-1} + \theta_0, \tag{8}$$

$$\vartheta_{2k} = \phi_{2k} - \psi_{2k} + \theta_0, \tag{9}$$

where $k \in [1,K]$, ϕ_{2k-1} and ϕ_{2k} are phase shifts caused by various reasons, e.g., RF wave propagation, Doppler effect, frequency shift between the signals, etc. Here, the relevant information bearing phase shifts can be evaluated with, respectively,

$$\psi_{2k-1} = 4\pi f_{2k-1}\frac{L_1}{v} \quad \text{and} \quad \psi_{2k} = 4\pi f_{2k}\frac{L_2}{v}. \tag{10}$$

At the receiver, each of the RF pulses $u_i(t)$, $i \in [1, 2K]$, is contaminated by zero mean additive stationary narrowband Gaussian noise $n(t)$ with a known variance σ^2, so that, at $t = t_i$, we have a mixture

$$y_i(t) = u_i(t) + n(t) = V_i(t)\cos[2\pi f_i t + \theta_i(t)], \tag{11}$$

where $V_i \geq 0$ is a positive valued envelope with the Rice distribution and $|\theta_i| \leq \pi$ is the modulo 2π random phase[2]. Although the frequency f_i in the reflected pulses can be different, below we often let the frequencies be equal, by setting $f_k = f_0$. The instantaneous peak SNR in $y_i(t)$ (Fig. 1b) is calculated by

$$\gamma_i = \frac{S\beta^2(t_i)}{\sigma^2}. \tag{12}$$

Because of noise, the actual phase difference[3]

$$\bar{\Theta}_k = \vartheta_{2k} - \vartheta_{2k-1} = \psi_{2k-1} - \psi_{2k} \tag{13}$$

$$= \frac{4\pi}{v}[f_{2k-1}L_1 - f_{2k}L_2] \tag{14}$$

$$\cong -2\pi f_0 \Delta\tau_{(2k)(2k-1)}, \tag{15}$$

where $\vartheta_{2k-1} = \vartheta_1(t_{2k-1})$, $\vartheta_{2k} = \vartheta_2(t_{2k})$, $\psi_{2k-1} = \psi_1(t_{2k-1})$, and $\psi_{2k} = \psi_2(t_{2k})$ cannot be measured precisely and are estimated at the coherent receiver via the noisy phase difference $\theta_{2k} - \theta_{2k-1}$ as (3), using (1) or (2). Similarly, the time drift in $\bar{\Theta}_k$ is evaluated by

$$\bar{\Psi}_k = \bar{\Theta}_k - \bar{\Theta}_{k-1} \tag{16}$$

$$= -\psi_{2k} + \psi_{2k-1} + \psi_{2k-2} - \psi_{2k-3}. \tag{17}$$

So, instead of the actual angle $\bar{\Theta}_k$, the coherent receiver produces its random estimate $\hat{\Theta}_k$ and instead of $\bar{\Psi}_k$ we have $\hat{\Psi}_k$. Note that, in the ideal receiver, Θ_k and $\hat{\Theta}_k$ as well as Ψ_k and $\hat{\Psi}_k$ have the same distributions [11].

3. Probability density of the phase difference

Because both the received signal and noise induced by the receiver are essentially narrowband processes, the instantaneous phase θ_i in (11) has Bennett's conditional distribution

[2] Throughout the paper, we consider the modulo 2π phase and phase difference.

[3] For the sake of simplicity, we assume equal phases $\phi(t_{2k})$ and $\phi(t_{2k-1})$. It is important that a linearly drifting phase difference $\phi(t_{2k}) - \phi(t_{2k-1})$ does not affect distribution of Θ_k [8] and may be accounted as a regular error.

$$p(\theta_i \mid \gamma_i, \vartheta_i) = \frac{e^{-\gamma_i}}{2\pi} + \sqrt{\frac{\gamma_i}{\pi}} e^{-\gamma_i \sin^2 \tilde{\theta}_i} \Phi\left(\sqrt{2\gamma_i} \cos\tilde{\theta}_i\right) \cos\tilde{\theta}_i , \tag{18}$$

where $\tilde{\theta}_i = \theta_i - \vartheta_i$ and $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} \mathrm{d}t$ is the probability integral. It has been shown in [17, 13] that (18) is fundamental for the interrogating RF pulses of arbitrary waveforms and modulation laws.

Employing the maximum likelihood function approach, the coherent receiver produces an estimate $\hat{\theta}_i$ of θ_i [11]. Assuming in this paper an ideal receiver, we let $\hat{\theta}_i = \theta_i$. Provided (18), the pdf of the information bearing phase difference Θ_k can be found for equal and different SNRs in the pulses and we notice that the problem is akin to that in two channel phase systems.

3.1 Different SNRs in the RF pulses

Most generally, one can suppose that the SNRs in the reflected pulses are different, $\gamma_{2k-1} \neq \gamma_{2k}$, owing to design problems and the SAW attenuation with distance. The relevant conditional pdf was originally published by Tsvetnov in 1969 [16]. Independently, in 1981, Pawula presented an alternative formula [21] that soon after appeared in [18] in a simpler form of

$$p(\Theta_k \mid \gamma_{2k-1}, \gamma_{2k}, \bar{\Theta}_k) = \frac{e^{-\bar{\gamma}}}{2\pi} \left[\cosh\breve{\gamma} + \frac{1}{2} \int_0^{\pi} \left(\bar{\gamma} \sin y + \lambda\right) \cosh(\breve{\gamma} \cos y) \times e^{\lambda \sin y} \mathrm{d}y \right], \tag{19}$$

where $\bar{\gamma} = (\gamma_{2k-1} + \gamma_{2k})/2$, $\breve{\gamma} = (\gamma_{2k} - \gamma_{2k-1})/2$, $\xi = \arctan\frac{\breve{\gamma}}{\lambda}$, $\lambda = \sqrt{\gamma_{2k-1}\gamma_{2k}} \cos\tilde{\Theta}_k$, and $\tilde{\Theta}_k = \Theta_k - \bar{\Theta}_k$. An equivalence of the Tsvetnov and Pawula pdfs was shown in [22].

To avoid computational problems, Tsvetnov expended his pdf in [20] to the Fourier series

$$p(\Theta_k \mid \gamma_{2k-1}, \gamma_{2k}, \bar{\Theta}_k) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{n=1}^{N} c_n(\gamma_{2k-1}, \gamma_{2k}) \cos n(\Theta_k - \bar{\Theta}_k), \tag{20}$$

where N is proportional to the maximum SNR in the pulses, $c_n(\gamma_{2k-1}, \gamma_{2k}) = c_n(\gamma_{2k-1}) c_n(\gamma_{2k})$, and

$$c_n(\gamma_i) = \frac{\sqrt{\pi\gamma_i}}{2} e^{-\gamma_i/2} \left[I_{(n+1)/2}\left(\frac{\gamma_i}{2}\right) + I_{(n-1)/2}\left(\frac{\gamma_i}{2}\right) \right], \tag{21}$$

where $I_v(x)$ is the modified Bessel function of the first kind and fractional order v. The mean and mean square values associated with (20) have been found in [24] to be, respectively,

$$\left\langle \Theta_k \right\rangle = 2 \sum_{n=1}^{N} \frac{(-1)^{n+1}}{n} c_n(\gamma_{2k-1}, \gamma_{2k}) \sin n\bar{\Theta}_k , \tag{22}$$

$$\left\langle \Theta_k^2 \right\rangle = \frac{\pi^2}{3} + 4 \sum_{n=1}^{N} \frac{(-1)^n}{n^2} c_n(\gamma_{2k-1}, \gamma_{2k}) \cos n\bar{\Theta}_k . \tag{23}$$

3.2 Equal SNRs in the RF pulses

In a special case when the SNRs in the pulses ara supposed to be equal, $\bar{\gamma}_k = \gamma_{2k-1} = \gamma_{2k}$, the phase difference has the conditional Tsvetnov pdf [20]

$$p(\Theta_k \mid \bar{\gamma}_k, \bar{\Theta}_k) = \frac{e^{-\bar{\gamma}_k}}{2\pi}\left[1 + \int_0^{\pi/2} \left(\bar{\gamma}_k \cos z + \lambda_k\right) e^{\lambda_k \cos z} dz\right], \tag{24}$$

where $\lambda_k = \bar{\gamma}_k \cos \tilde{\Theta}_k$. Note that Tsvetnov published his pdf in the functional form. The integral equivalent (24) shown in [11] does not appear in Tsvetnov's works. It can be observed that, by equal SNRs, (19) becomes (24), although indirectly.

3.3 Probability density of the differential phase difference

It has been shown in [22] that the pdf of the differential phase difference (DPD) has two equivalent forms.

The first form of this pdf appears to be

$$\begin{aligned} p_\Psi &\triangleq p(\Psi_k \mid \gamma_{2k}, \gamma_{2k-1}, \gamma_{2k-2}, \gamma_{2k-3}, \bar{\Psi}_k) \\ &= \frac{e^{-\bar{\gamma}_1 - \bar{\gamma}_2}}{2\pi}\left[A + \frac{1}{4}\int_0^{\pi}\int_0^{\pi} \cosh(\bar{\gamma}_1 \cos x)\cosh(\bar{\gamma}_2 \cos y) F(x, y, \tilde{\Psi}_k) \mathrm{d}x\mathrm{d}y\right], \end{aligned} \tag{25}$$

where $\bar{\gamma}_1 = (\gamma_{2k-1} + \gamma_{2k})/2$, $\bar{\gamma}_2 = (\gamma_{2k-3} + \gamma_{2k-2})/2$,

$$\begin{aligned} F = &\left[\bar{\gamma}_1\bar{\gamma}_2 \sin x \sin y + \frac{1}{2}\sqrt{\gamma_{2k}\gamma_{2k-1}\gamma_{2k-2}\gamma_{2k-3}} \cos\tilde{\Psi}\right] I_0(a) \\ &+ \left[\bar{\gamma}_1\sqrt{\gamma_{2k-3}\gamma_{2k-2}} \cos(\tilde{\Psi} + \zeta)\sin x + \bar{\gamma}_2\sqrt{\gamma_{2k-1}\gamma_{2k}} \cos\zeta \sin y\right] I_1(a) \\ &+ \frac{1}{2} I_2(a)\sqrt{\gamma_{2k}\gamma_{2k-1}\gamma_{2k-2}\gamma_{2k-3}} \cos(\tilde{\Psi} + 2\zeta), \end{aligned} \tag{26}$$

$$\begin{aligned} a(x, y, \tilde{\Psi}) = &\left[2\sqrt{\gamma_{2k}\gamma_{2k-1}\gamma_{2k-2}\gamma_{2k-3}} \sin x \sin y \cos\tilde{\Psi}\right. \\ &\left. + \gamma_{2k-1}\gamma_{2k}\sin^2 x + \gamma_{2k-3}\gamma_{2k-2}\sin^2 y\right]^{1/2}, \end{aligned} \tag{27}$$

$$\zeta(x, y, \tilde{\Psi}) = \begin{cases} \arctan(Q/I), & I \geq 0 \\ \arctan(Q/I) \pm \pi, & I < 0, \begin{cases} Q \geq 0, \\ Q < 0 \end{cases} \end{cases} \tag{28}$$

$$Q = -\sqrt{\gamma_{2k-3}\gamma_{2k-2}} \sin y\tilde{\Psi}, \tag{29}$$

$$I = \sqrt{\gamma_{2k-1}\gamma_{2k}} \sin x + \sqrt{\gamma_{2k-3}\gamma_{2k-2}} \sin y \cos\tilde{\Psi}. \tag{30}$$

By changing the variables, namely by substituting sin x with x and sin y with y, the pdf transforms to its second equivalent form of

$$p_\Psi = \frac{e^{-\bar{\gamma}_1-\bar{\gamma}_2}}{2\pi}\left[A+\int_0^1\int_0^1 \frac{\cosh(\breve{\gamma}_1\sqrt{1-x^2})\cosh(\breve{\gamma}_2\sqrt{1-y^2})}{\sqrt{(1-x^2)(1-y^2)}}\times G(x,y,\tilde{\Psi}_k)\mathrm{d}x\mathrm{d}y\right], \tag{31}$$

where

$$\begin{aligned} G = &\left[\bar{\gamma}_1\bar{\gamma}_2 xy + \frac{1}{2}\sqrt{\gamma_{2k}\gamma_{2k-1}\gamma_{2k-2}\gamma_{2k-3}}\cos\tilde{\Psi}_k\right]I_0(b) \\ &+\left[\bar{\gamma}_2\sqrt{\gamma_{2k-1}\gamma_{2k}}\,y\cos\zeta + \bar{\gamma}_1\sqrt{\gamma_{2k-3}\gamma_{2k-2}}\,x\cos(\tilde{\Psi}_k+\zeta)\right] \\ &\times\ I_1(b)+\frac{1}{2}I_2(b)\sqrt{\gamma_{2k}\gamma_{2k-1}\gamma_{2k-2}\gamma_{2k-3}}\cos(\tilde{\Psi}_k+2\zeta), \end{aligned} \tag{32}$$

$$\begin{aligned} b(x,y,\tilde{\Psi}_k) = &\Big[2xy\sqrt{\gamma_{2k}\gamma_{2k-1}\gamma_{2k-2}\gamma_{2k-3}}\cos\tilde{\Psi}_k \\ &+\gamma_{2k-1}\gamma_{2k}x^2+\gamma_{2k-3}\gamma_{2k-2}y^2\Big]^{1/2}, \end{aligned} \tag{33}$$

and $\zeta(x,y,\tilde{\Psi})$ is given by (28) with $Q=-y\sqrt{\gamma_{2k-3}\gamma_{2k-2}}\sin\tilde{\Psi}_k$ and $I=x\sqrt{\gamma_{2k-1}\gamma_{2k}}+y\sqrt{\gamma_{2k-3}\gamma_{2k-2}}\cos\tilde{\Psi}_k$.

One may arrive at the conclusion that neither (25) nor (31) allow for further substantial simplifications and closed forms even in the special case of equal SNRs in the first and second pulses.

3.3.1 Equal SNRs in the first and second pulses

By letting $\gamma_1 = \gamma_{2k-1}$ and $\gamma_2 = \gamma_{2k}$, the pdf p_ψ attains the form shown in [22]

$$p_\Psi = \frac{e^{-2\bar{\gamma}}}{2\pi}\left[2e^{\breve{\gamma}}\cosh\breve{\gamma}-\cosh^2\breve{\gamma}+\int_0^1\int_0^1\frac{\cosh(\breve{\gamma}\sqrt{1-x^2})\cosh(\breve{\gamma}\sqrt{1-y^2})}{\sqrt{(1-x^2)(1-y^2)}}\times G(x,y,\tilde{\Psi}_k)\mathrm{d}x\mathrm{d}y\right], \tag{34}$$

where $\bar{\gamma}=(\gamma_2+\gamma_1)/2$, $\breve{\gamma}=(\gamma_2-\gamma_1)/2$,

$$\begin{aligned} G = &\left(\bar{\gamma}^2xy+\frac{\gamma_1\gamma_2}{2}\cos\tilde{\Psi}_k\right)I_0(b)+\bar{\gamma}\sqrt{\gamma_1\gamma_2}\left[y\cos\zeta+x\cos(\tilde{\Psi}_k+\zeta)\right] \\ &\times I_1(b)+\frac{\gamma_1\gamma_2}{2}I_2(b)\cos(\tilde{\Psi}_k+2\zeta), \end{aligned} \tag{35}$$

$$b(x,y,\tilde{\Psi}_k)=\left[\gamma_1\gamma_2(x^2+2xy\cos\tilde{\Psi}_k+y^2)\right]^{1/2}, \tag{36}$$

and $\zeta(x,y,\tilde{\Psi}_k)$ is given by (28) with $Q=-y\sqrt{\gamma_1\gamma_2}\sin\tilde{\Psi}_k$ and $I=\sqrt{\gamma_1\gamma_2}(x+y\cos\tilde{\Psi}_k)$.

3.3.2 Equal SNRs in the pulses

For SAW sensors with closely placed reflectors, one may suppose that all of the received RF pulses have equal SNRs, $\gamma = \gamma_{2k-1} = \gamma_{2k}$. By setting $\bar{\gamma} = \gamma$ and $\breve{\gamma} = 0$, substituting the

hyperbolic functions with the exponential ones, and providing the routine transformations, we arrive at the pdf originally derived in [25],

$$\begin{aligned} p_\Psi &= p(\Psi_k \mid \gamma, \overline{\Psi}_k) \\ &= \frac{2e^{-\gamma} - e^{-2\gamma}}{2\pi} + \frac{\gamma^2 e^{-2\gamma}}{2\pi} + \int_0^1 \int_0^1 \frac{\mathrm{d}x\mathrm{d}y}{\sqrt{(1-x^2)(1-y^2)}} \\ &\quad \times [E_1 I_0(\gamma E_2) + E_3 I_1(\gamma E_2)], \end{aligned} \tag{37}$$

where

$$E_1 = xy(1 + E_2^{-2}\sin^2\tilde{\Psi}) + \cos\tilde{\Psi}, \tag{38}$$

$$E_2 = \sqrt{x^2 + 2xy\cos\tilde{\Psi} + y^2}, \tag{39}$$

$$E_3 = (E_2\cos\tilde{\Psi} + 2xyE_2^{-2}\sin^2\tilde{\Psi})(1 - \gamma^{-1}E_2^{-2}). \tag{40}$$

Certainly, (37) can be used when there is no substantial difference in the RF pulses of the received burst, although (34) gives us a more realistic picture. Notwith-standing this fact, neither of the above discussed pdfs has engineering features. Below, we shall show that this disadvantage is efficiently circumvented with quite simple and reasonably accurate approximations.

4. Von-Mises/Tikhonov-based approximations

Observing the above-described probability densities of the phase difference and DPD, one can deduce they all these relations are not suitable for the engineering use and approximations having simpler forms would be more appropriate. It has been shown in [14] that efficient approximations can be found employing the von Mises/Tikhonov distribution known as circular normal distribution. The von Mises/Tikhonov pdf [15] is

$$p(\varphi) = \frac{1}{2\pi I_0(\alpha)} e^{\alpha\cos(\varphi - \varphi_0)}, \tag{41}$$

where $\alpha(\gamma)$ is the SNR-sensitive parameter, φ is the mod 2π variable phase, and φ_0 is some constant value. Commonly, (41) is used by the authors to approximate Bennett's pdf (18) for the instantaneous phase θ with the error of about 5% and $\alpha(\gamma)$ specified in the least mean squares (LMS) sense. Shmaliy showed in [14] that (41) fits better the phase difference pdf with equal SNRs allowing for the approximation error lesser 0.6%. Referring to this fact, below we give simple and reasonably accurate von Mises/Tikhonov-based distributions for the phase difference and DPD.

4.1 Phase difference

To fit the phase difference Θ_k with arbitrary SNRs, Tsvetnov proposed in [16] an approximation

$$p(\Theta_k \mid \overline{\gamma}_k, \overline{\Theta}_k) = \frac{1}{2\pi I_0(\overline{\gamma}_k)} e^{\overline{\gamma}_k\cos(\Theta_k - \overline{\Theta}_k)}, \tag{42}$$

where

$$\bar{\gamma}_k = \frac{2\gamma_{2k-1}\gamma_{2k}}{\gamma_{2k-1}+\gamma_{2k}}, \tag{43}$$

allowing for a maximum error of about 20%. Referring to [16], Shmaliy showed in [14] that (41) works out with a maximum error at $\gamma_{2k-1} = \gamma_{2k} \simeq 0.6$ of about 0.6% if to write

$$p(\Theta_k \mid \bar{\gamma}_k, \bar{\Theta}_k) = \frac{1}{2\pi I_0(\alpha_k)} e^{\alpha_k \cos(\Theta_k - \bar{\Theta}_k)} \tag{44}$$

and set

$$\alpha_k = \bar{\gamma}_k (1 + a e^{-b\bar{\gamma}_k}), \tag{45}$$

where a = 0.525 and b = 1.1503 are determined in the least mean squares (LMS) sense. With $\gamma_{2k-1} \neq \gamma_{2k}$, the error increases up to about 5 % with the SNRs difference tending toward infinity. The mean and mean square values associated with (44), for a fixed α_k, are, respectively,

$$\left\langle \Theta_k \right\rangle = 2\sum_{n=1}^{N} \frac{(-1)^{n+1}}{n} \varrho_{n0}(\alpha_k) \sin n\bar{\Theta}_k, \tag{46}$$

$$\left\langle \Theta_k^2 \right\rangle = \frac{\pi^2}{3} + 4\sum_{n=1}^{N} \frac{(-1)^n}{n^2} \varrho_{n0}(\alpha_k) \cos n\bar{\Theta}_k, \tag{47}$$

where N is proportional to the SNR and

$$\varrho_{ij}(x) = \frac{I_i(x)}{I_j(x)} \tag{48}$$

is a ratio of the modified Bessel functions of the first kind and integer order. It can be shown that, by zero and large SNRs, (44) becomes uniform and normal, respectively,

$$p(\Theta_k) = \frac{1}{2\pi}, \tag{49}$$

$$p(\Theta_k \mid \bar{\gamma}_k, \bar{\Theta}_k) = \sqrt{\frac{\bar{\gamma}_k}{2\pi}} e^{-\frac{\bar{\gamma}_k}{2}(\Theta_k - \bar{\Theta}_k)^2}, \tag{50}$$

having in the latter case (50) the variance $\sigma_\Theta^2 = 1/\bar{\gamma}_k$.

4.2 Differential phase difference

It has also been shown in [14] that, to fit Ψ_k, the following von Mises/Tikhonov- based approximations may be used with a maximum error of about 0.41 % at equal unit SNRs. For different and equal SNRs these pdfs are, respectively[4],

[4] (52) was originally derived in [17].

$$p(\Psi_k \mid \bar{\gamma}_{k-1}, \bar{\gamma}_k, \bar{\Psi}_k) = \frac{1}{2\pi} \frac{I_0(r_k)}{I_0(\alpha_{k-1}) I_0(\alpha_k)}, \tag{51}$$

$$p(\Psi_k \mid \bar{\gamma}_k, \bar{\Psi}_k) = \frac{1}{2\pi} \frac{I_0(r_{1k})}{I_0^2(\alpha_k)}, \tag{52}$$

where

$$r_k = \sqrt{\alpha_{k-1}^2 + 2\alpha_{k-1}\alpha_k \cos(\Psi_k - \bar{\Psi}_k) + \alpha_k^2},$$
$$r_{1k} = \alpha_k \sqrt{2[1 + \cos(\Psi_k - \bar{\Psi}_k)]}.$$

The mean and mean square values associated with (51) and (52) are, respectively,

$$\langle \Psi_k \rangle = 2 \sum_{n=1}^{N} \frac{(-1)^{n+1}}{n} \varrho_{n0}(\alpha_{k-1}, \alpha_k) \sin n\bar{\Psi}_k, \tag{53}$$

$$\langle \Psi_k^2 \rangle = \frac{\pi^2}{3} + 4 \sum_{n=1}^{N} \frac{(-1)^n}{n^2} \varrho_{n0}(\alpha_{k-1}, \alpha_k) \cos n\bar{\Psi}_k, \tag{54}$$

where $\varrho_{n0}(\alpha_{k-1}, \alpha_k) = \varrho_{n0}(\alpha_{k-1}) \varrho_{n0}(\alpha_k)$.
Several important limiting cases can now be distinguished.

4.2.1 Case 1: Large SNR in one of the signals

With $\bar{\gamma}_k \gg 1$ and $\bar{\gamma}_{k-1} \ll \bar{\gamma}_k$, the pdf (51) degenerates to the von Mises/Tikhonov density (41).

4.2.2 Case 2: One of the signals is a pure noise

Let $\bar{\gamma}_k = 0$ and $\bar{\gamma}_{k-1} \neq 0$. With $\bar{\gamma}_k = 0$, (51) transforms to the uniform density (49) disregarding the other SNR value. Therefore, (52) also becomes uniform.

4.2.3 Case 3: Large and equal SNRs in the vectors

With $1 \ll \bar{\gamma}_{k-1} = \bar{\gamma}_k$, the pdf (52) degenerates to the normal density

$$p(\Psi_k \mid \bar{\gamma}_k, \bar{\Psi}_k) \cong \sqrt{\frac{\bar{\gamma}_k}{4\pi}} e^{-\frac{\bar{\gamma}_k}{4}(\Psi_k - \bar{\Psi}_k)^2}, \tag{55}$$

in which the variance is $\sigma_\Psi^2 = 2 / \bar{\gamma}_k$.
As can be observed, all the von Mises/Tikhonov-based approximations have simple engineering forms allowing for small errors with typically near equal SNRs in the received SAW sensor pulses.

5. Errors in the phase difference estimates

To evaluate errors in the estimates of phase angles, let us assume that the actual phase difference between the received pulses of the SAW sensor is $\bar{\Theta}_k$. At the receiver, this

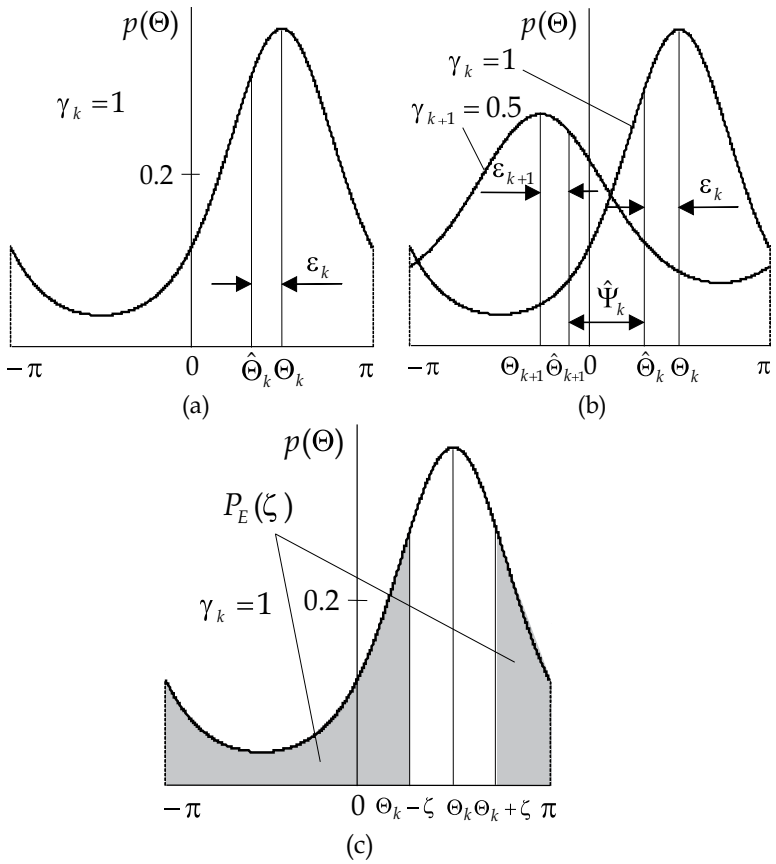


Fig. 2. Errors in passive SAW sensing with DPM and $\gamma_k = 1$: (a) instantaneous ϵ_k, (b) differential phase diversity $\hat{\Psi}_k$, and (c) error probability $P_E(\zeta)$ [25]. Note that, with $\gamma_k \gg 1$, the pdf tends to be normal and, by $\gamma_k \to 0$, it becomes uniform.

difference becomes noisy and is estimated as $\hat{\Theta}_k$ with the probability density (19) or (44). We thus have an estimate with the instantaneous error $\epsilon_k = \bar{\Theta}_k - \hat{\Theta}_k$. Figure 2a illustrates the estimate errors for the case of $\gamma_k = 1$, in which we recognize the mean error (bias) and the mean square error (MSE), respectively,

$$\langle \epsilon_k \rangle = \bar{\Theta}_k - \langle \hat{\Theta}_k \rangle, \tag{56}$$

$$\langle \epsilon_k^2 \rangle = \bar{\Theta}_k^2 - 2\bar{\Theta}_k \langle \hat{\Theta}_k \rangle + \langle \hat{\Theta}_k^2 \rangle. \tag{57}$$

When two neighboring values, $\hat{\Theta}_k$ and $\hat{\Theta}_{k+1}$, are unequal then the differential phase diversity $\hat{\Psi}_k$ occurs (Fig. 2b). If $\langle \hat{\Psi}_k \rangle = 0$, then the estimates $\hat{\Theta}_k$ are mutually unbiased. Otherwise, $\langle \hat{\Psi}_k \rangle$ causes a bias in the multiple DPM that may be associated with the sensor movement (Doppler effect).

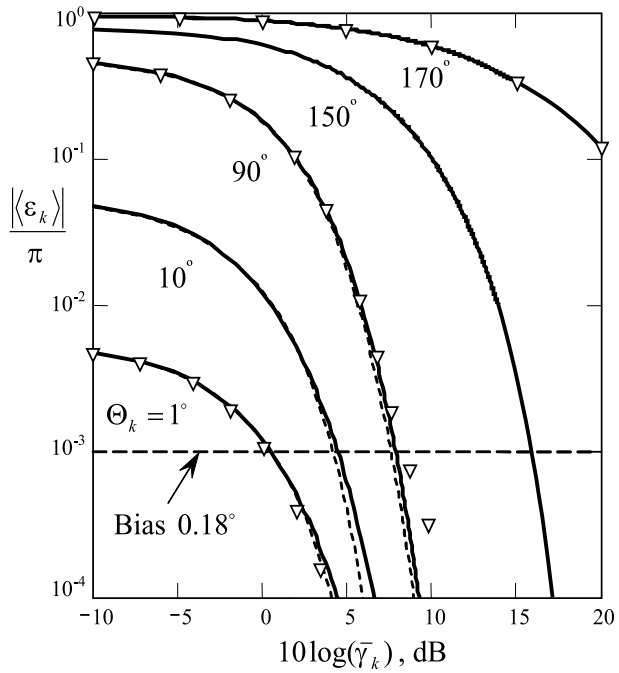


Fig. 3. Mean errors calculated for equal SNRs by (19) rigorously (bold) with c_n (21) and approximately (dashed) by (44) with ϱ_{n0} (48)[25].

5.1 Mean error

The mean error (bias) in the estimate can be evaluated if we test the rigorous and approximate distributions by (56). That gives us

$$\langle \epsilon_k \rangle = \bar{\Theta}_k - 2\sum_{n=1}^{N} \frac{(-1)^{n+1}}{n} A_n \sin n\bar{\Theta}_k , \tag{58}$$

where A_n is the amplitude of the Fourier series component. One must let $A_n = c_n(\gamma_{2k-1}, \gamma_{2k})$ as specified with (21) if the bias is calculated via the Tsvetnov/Pawula pdf (19). If the von Mises/Tikhonov-based approximating pdf (44) is used, then substitute A_n with $\varrho_{n0}(\alpha_k)$ specified with (48). Note that in each of these cases, the length N of the series is practically limited by the doubled maximum SNR in the pulses. Figure 3 exhibits $\langle \epsilon_k \rangle$ calculated rigorously, by c_n, and approximately, by ϱ_{n0}, for equal SNRs $\bar{\gamma}_k$. Here, triangle points represent simulation. It is seen that the approximation is accurate in a whole range of angles and SNRs.

5.2 Mean square error

Employing (57) and reasoning along similar lines, one can find the MSE in the estimate in the form of

$$\langle \epsilon_k^2 \rangle = \frac{\pi^2}{3} + \bar{\Theta}_k^2 + 4\sum_{n=1}^{N} \frac{(-1)^n}{n} A_n \left(\frac{1}{n} \cos n\bar{\Theta}_k + \bar{\Theta}_k \sin n\bar{\Theta}_k \right). \tag{59}$$

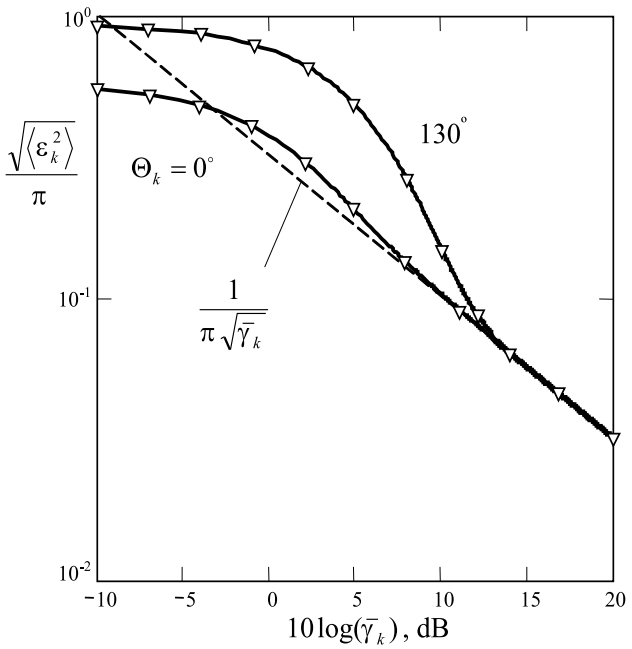


Fig. 4. Root MSEs calculated with equal SNRs. No visible error is observed between the rigorous estimate, by c_n, and approximate estimate, by ϱ_{n0}. The asymptotic behavior (61) is dashed [25].

It may also be calculated approximately in two important special cases:

- $\bar{\gamma}_k < 0$ dB, then

$$\begin{aligned} \left\langle \epsilon_k^2 \right\rangle \cong\ & \frac{\pi^2}{3} + \bar{\Theta}_k^2 - \pi\bar{\gamma}_k e^{-\bar{\gamma}_k}\left(1 + \frac{\bar{\gamma}_k}{4} + \frac{\bar{\gamma}_k^2}{16}\right)^2 \left(\cos\bar{\Theta}_k + \bar{\Theta}_k \sin\bar{\Theta}_k\right) \\ & + \frac{\bar{\gamma}_k^2 e^{-\bar{\gamma}_k}}{16}\left(\frac{\bar{\gamma}_k}{3} + 2\right)^2 \left(\cos 2\bar{\Theta}_k + 2\bar{\Theta}_k \sin 2\bar{\Theta}_k\right), \end{aligned} \tag{60}$$

- If $\bar{\gamma}_k > 13$ dB, then

$$\left\langle \epsilon_k^2 \right\rangle \cong \frac{1}{\bar{\gamma}_k}. \tag{61}$$

Figure 4 sketches the root MSE (RMSE) calculated rigorously, by c_n, and approximately, by ϱn_0, for equal SNRs, $\bar{\gamma}_k$. One can observe that there is no visible difference between two curves and thus the von Mises/Tikhonov approximation is highly accurate. Furthermore, by large SNR $\bar{\gamma}_k > 13$ dB, both curves converge to the asymptotic line given by (61).

5.3 Error variance and Cramér-Rao lower bound

A measure of noise in the estimate is the variance calculated for a single DPM by

$$\sigma_{\Theta k}^2 = \langle \hat{\Theta}_k^2 \rangle - \langle \hat{\Theta}_k \rangle^2. \tag{62}$$

With multiple DPM, the variance is often substituted with the Cramér-Rao lower bound (CRLB) having approximate, although typically simple representations. Supposing that the measurement vector **x** is formed with N readings as $\mathbf{x} = [x_0\, x_1 \ldots x_{N-1}]^T$ with the uncorrelated phase difference components x_k having equal SNRs, $\alpha_k = \alpha$, and actual phase difference $\bar{\Theta}$, the likelihood function can be written as

$$p(\mathbf{x};\bar{\Theta}) = \frac{1}{[2\pi I_0(\alpha)]^N} e^{\alpha \sum_{k=0}^{N-1} \cos(x_k - \bar{\Theta})}. \tag{63}$$

Accordingly, the CRLB calculates

$$\sigma_\Theta^2 \geq \frac{1}{-\langle \frac{\partial^2}{\partial \bar{\Theta}^2} ln p(\mathbf{x};\bar{\Theta}) \rangle}$$

$$= \frac{1}{\alpha N r \cos(\bar{\Theta} - \phi)}, \tag{64}$$

where $r = \sqrt{\bar{x}_c^2 + \bar{x}_s^2}$, $\phi = arctan \frac{\bar{x}_s}{\bar{x}_c}$, $\bar{x}_c = \frac{1}{N} \sum_{k=0}^{N-1} \cos x_k$, and $\bar{x}_s = \frac{1}{N} \sum_{k=0}^{N-1} \sin x_k$. It is seen that large N causes a substantial decrease in the estimate variance. It can also be shown that, for single DPM, the CRLB reduces to

$$\sigma_\Theta^2 \geq \frac{1}{\alpha \cos(\bar{\Theta} - x_0)}, \tag{65}$$

where x_0 may be assumed to be either $\langle \Theta \rangle$ or Θ_0.

Figure 5 illustrates this analysis. We first notice that the von Mises/Tikhonov approximation does not produce a visible error, like in Fig. 4. By (65) and $x_0 = \Theta_0$, The CRLB holds true for all angles and traces below the asymptotic line (61) owing to the function (45) (see the case of $\Theta_k = 0^o$). With $x_0 = \langle \Theta \rangle$, (65) produces more realistic values if SNRs > 0 dB and too large values with SNRs < 0 dB that may not be appropriate in applications (see the case of $\Theta_k = 130^o$).

6. Error probability for the estimate to exceed a threshold

The error probability P_E of passive remote SAW sensing is the conditional probability $P(\zeta \leq | \bar{\Theta}_k - \hat{\Theta}_k \| \gamma_{2k-1}, \gamma_{2k})$ for the error in the estimate of the phase difference between two pulses $| \bar{\Theta}_k - \hat{\Theta}_k |$ to exceed a threshold ζ. The P_E is represented by the shadowed area in Fig. 2c. Its quantity does not depend on $\bar{\Theta}_k$ and can be approximately estimated by

$$P_E(\zeta \,|\, \gamma_{2k-1}, \gamma_{2k}) = 1 - \frac{\zeta}{\pi} - \frac{2}{\pi} \sum_{n=1}^{N} \frac{1}{n} A_n \sin n\zeta, \tag{66}$$

$$\cong 1 - \frac{\zeta}{\pi} \left(1 + 2 \sum_{n=1}^{N} A_n \right), \quad \zeta \ll \pi, \tag{67}$$

$$\cong 1 - \frac{\zeta}{\pi} \left(1 + 2 \sum_{n=1}^{N} e^{-n^2 \frac{\gamma_{k+}}{2\gamma_{kx}^2}} \right), \quad \zeta \ll \pi, \quad 1 \ll \gamma_i. \tag{68}$$

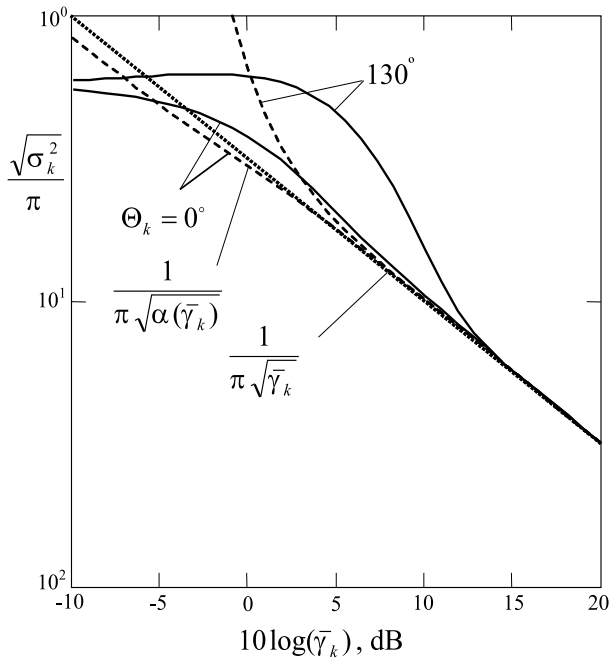


Fig. 5. Variance estimate with single DPM for two angles Θ = 0° and Θ = 130°: rigorous (bold), asymptotic (dotted), and CRLBs (dashed) [24].

where $A_n = c_n(\gamma_{2k-1}, \gamma_{2k})$ is specified with (21), if the bias is calculated via the Tsvetnov/Pawula pdf (19), and A_n must be substituted with $\varrho_{n0}(\alpha_k)$, if the von Mises/Tikhonov-based approximating pdf (44) is used.

The estimate of P_E is akin to that in the calculus of the conditional symbol error (SER) for the differentially coherent detection in digital communication channels with M-ary phase difference shift keying (MPDSK).We exploit it below to find approximate solutions for remote SAW sensing.

6.1 Approximate estimates for equal SNRs

Since $P_E(\zeta)$ is $\bar{\Theta}_k$ invariant, one can set $\bar{\Theta}_k = 0$ and go to the symmetric pdf $p(\Theta \mid \bar{\gamma}_k)$. The error probability can thus be calculated for a threshold ζ to range from 0 to π as [11]

$$P_E(\zeta \mid \bar{\gamma}_k) = 2\int_{\zeta}^{\pi} p(\Theta_k \mid \bar{\gamma}_k) d\Theta_k \tag{69}$$

$$= \frac{e^{-\bar{\gamma}_k}}{\pi} \sin\zeta \int_0^{\pi/2} \frac{e^{\bar{\gamma}_k \cos\zeta \cos t}}{1 - \cos\zeta \cos t} dt. \tag{70}$$

A similar formula employed in [18] is known as the conditional SER,

$$P_E(M \mid \bar{\gamma}_k) = 2\int_{\pi/M}^{\pi} p(\Theta_k \mid \bar{\gamma}_k) \mathrm{d}\Theta_k, \tag{71}$$

where M is an integer but may be arbitrary in a common case. By $M = \pi/\zeta$, (71) becomes (70) and we notice that (71) was performed in [18] in the integral form that, by symmetry of the integrand, simplifies to (70).
Alternatively, one can substitute (44) into (69) and arrive at the familiar von Mises/Tikhonov approximation

$$P_{ET}(\zeta \mid \overline{\gamma}_k) = \frac{1}{\pi I_0(\alpha_k)} \int_{\zeta}^{\pi} e^{\alpha_k \cos x} dx. \tag{72}$$

By employing the Fourier series analysis, one can also transform (72) to the computationally more preferable form of

$$P_{ET}(\zeta \mid \overline{\gamma}_k) = 1 - \frac{\zeta}{\pi} - \frac{2}{\pi} \sum_{n=1}^{N} \frac{1}{n} \varrho_{n0}(\alpha_k) \sin n\zeta. \tag{73}$$

Observing the above-given results concerning the error probability, several special cases can be distinguished.

6.1.1 Large and equal SNRs

It has been shown in [11] that large SNRs allow for the approximation of (69) by

$$P_E(\zeta \mid \overline{\gamma}_k \gg 1) \cong 2Q\left(\zeta \sqrt{\overline{\gamma}_k}\right), \tag{74}$$

where $Q(x)$ = $\frac{1}{2}$ erfc $(\frac{x}{\sqrt{2}})$ is the Gauss Q-function and erfc(x) is the complimentary error function. Several other approximate solutions found by Fleck and Trabka, Arthurs and Dym, Bussgang and Leiter, and Salz and Stein for $\pi/\zeta > 2$ where reported in [18]. These solutions are, respectively,

$$P_E(\zeta \mid \overline{\gamma}_k \gg 1) \cong \operatorname{erfc} X + \frac{X e^{-X^2}}{4\sqrt{\pi}(\overline{\gamma}_k + 0.125)}, \tag{75}$$

$$\cong \operatorname{erfc}\left(\sqrt{\overline{\gamma}_k} \sin \frac{\zeta}{\sqrt{2}}\right), \tag{76}$$

$$\cong \operatorname{erfc}\left(\frac{\overline{\gamma}_k}{\sqrt{1 + 2\overline{\gamma}_k}} \sin \zeta\right), \tag{77}$$

$$\cong \sqrt{\frac{1 + \cos\zeta}{2\cos\zeta}} \operatorname{erfc} X, \tag{78}$$

where $X(\zeta, \overline{\gamma}_k) = \sqrt{\overline{\gamma}_k (1 - \cos\zeta)}$. In turn, (44) becomes normal (50) with large SNRs, and one more approximation can be proposed for this case, namely

$$P_E(\zeta \mid \overline{\gamma}_k \gg 1) \cong 2Q\left(\zeta \sqrt{\overline{\gamma}_k}\right) - 2Q\left(\pi \sqrt{\overline{\gamma}_k}\right). \tag{79}$$

It has been shown in [14] that, among all other known approximations, the von Mises/Tikhonov-based one (72) is most accurate.

6.1.2 Low and equal SNRs

The case of SNR $\cong$ 0 dB allows for an asymptotic form of (70), by letting $e^x \cong 1-x$ and then integrating; that is,

$$P_E(\zeta \mid \overline{\gamma}_k \le 1) \cong \frac{2}{\pi}\arctan\frac{\sin\zeta}{1-\cos\zeta} - \frac{\overline{\gamma}_k}{2}\sin\zeta. \tag{80}$$

It turns out, however, that the inaccuracy of (80) is larger than that produced by (73), if we set $N = 2$. The latter approximation is given by

$$P_{ET}(\zeta \mid \overline{\gamma}_k \cong 1) \cong 1 - \frac{\zeta}{\pi} - \frac{2}{\pi}\varrho_{10}(\alpha_k)\left[1+\varrho_{21}(\alpha_k)\cos\zeta\right]\sin\zeta \tag{81}$$

$$= 1 - \frac{\zeta}{\pi} - \frac{2}{\pi}\varrho_{10}(\alpha_k)\left(1-\frac{2}{\alpha_k}\cos\zeta\right)\sin\zeta - \frac{1}{\pi}\sin 2\zeta. \tag{82}$$

6.1.3 Very low and equal SNRs

Let us finally consider the case of $\overline{\gamma}_k \ll 1$. With such values of $\overline{\gamma}_k$, (80) simplifies to

$$P_E(\zeta \mid \overline{\gamma}_k \ll 1) \cong \frac{2}{\pi}\arctan\frac{\sin\zeta}{1-\cos\zeta} \tag{83}$$

and, because $\varrho_{10}(x \ll 1) \cong 4x/(x^2+8)$, the approximation (82) becomes

$$P_{ET}(\zeta \mid \overline{\gamma}_k < 1) \cong 1 - \frac{\zeta}{\pi} - \frac{2}{\pi}\varrho_{10}(\alpha_k)\sin\zeta \tag{84}$$

$$\cong 1 - \frac{\zeta}{\pi} - \frac{8\alpha_k}{\pi(\alpha_k^2+8)}\sin\zeta. \tag{85}$$

Again, we notice that (80) is not a rival here, being still lesser accurate, and, as a matter of fact, we note that only with SNR < –20 dB the approximate functions (83)-(85) trace along the same trajectory.

Figure 6 illustrates the rigorous, by (70), and approximate, by (72), calculation of the error probability for equal SNRs and different values of the threshold. A splendid property of the von Misis/Tikhonov-based approximations is indicated instantly: the error is negligible in the whole range of angles with arbitrary values of SNRs.

7. Phase difference drift rate

A measure D of the drift rate of the phase difference Θ_k in the received RF pulse- burst has three critical applications: 1) It represents the drift rate error when the burst is used to increase the SNR in the received signal [7]; 2) When the SAW sensor is intended to measure a physical quantity, then D characterizes speed of change of this quantity; and 3) If the SAW reader system measures velocity of a moving object, then D gives a measure of acceleration. In applications, of interest are the mean value $\langle D \rangle$ and variance σ_D^2 of the drift rate.

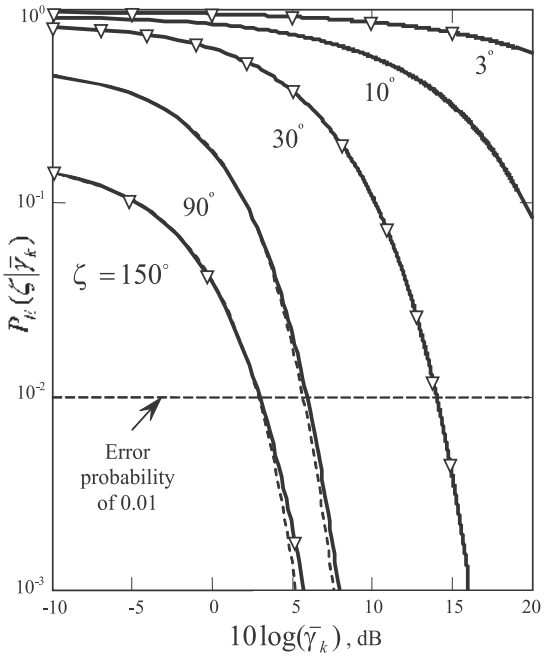


Fig. 6. Error probability for equal SNRs: rigorous (bold), by (70), and approximated by von Mises/Tikhonov's distribution (dashed), by (72) [25].

7.1 Mean drift rate

The mean drift rate $\langle D\rangle$ can be evaluated employing both (51) and the mean value $\langle \Psi_k\rangle$ given in [11] as in the following,

$$\langle D\rangle = \frac{\langle \Psi_k\rangle}{T} \tag{86}$$

$$= \frac{1}{T}\int_{-\pi}^{\pi} \Psi_k \bar{p}(\Psi_k \,|\, \gamma_{2k}, \gamma_{2k-1}, \gamma_{2k-2}, \gamma_{2k-3}, \tilde{\Psi}_k)\mathrm{d}\Psi_k \tag{87}$$

$$\cong \frac{2}{T}\sum_{n=1}^{N} \frac{(-1)^{n+1}}{n} \frac{I_n(\alpha_1)}{I_0(\alpha_1)} \frac{I_n(\alpha_2)}{I_0(\alpha_2)} \sin n\bar{\Psi}_k\,, \tag{88}$$

where a reasonable series length is limited with $N \geq 2 \max\ \bar{\gamma}_{1,2}$ [3]. The estimate of $\langle D\rangle$ can be found by averaging $\hat{\Psi}_k$ at the coherent receiver as

$$\langle \hat{D}\rangle = \frac{1}{TK}\sum_{k=1}^{K} \hat{\Psi}_k\,. \tag{89}$$

Example 1: Burst length vs. the drift rate [25]. Consider a passive SAW sensor of temperature [10] operating at the frequency f_0 = 2.45 GHz with the temperature sensitivities of the delay time difference S_t = 0.017 ns/K and phase difference $S_p = 2\pi f_0 S_t$ = 0.262 rad/K. Suppose that the temperature rate at the sensor substrate is 1 K per 10 sec; that is

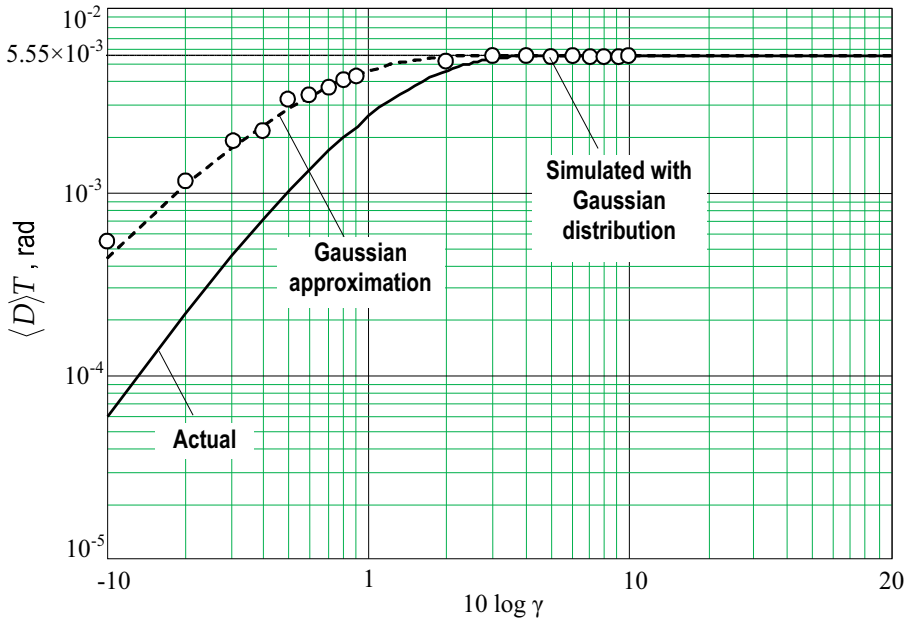


Fig. 7. The mean drift rate $\langle D\rangle$ for $\bar{\Psi} = 5.55 \times 10^{-3}$ rad: actual, by (88); Gaussian approximation; and simulation with Gaussian distribution, using (89) [25].

γ	$\lvert\bar{\Psi}\rvert$, rad					
	0.1π	0.3π	0.5π	0.7π	0.8π	0.9π
dB	> 4	> 5	> 7	> 11	> 13	> 21

Table 1. Allowed γ for Accurate Estimation of $\langle D\rangle$ with Different $\bar{\Psi}$

$R_t = 10^{-10}$ K/ns. The phase difference mean drift rate is thus $\langle D\rangle = S_p R_t = 2.62\times10^{-11}$ rad/ns. The sensor is interrogated with the pulse-burst of K pulses and period T. During the burst length $L = KT$, temperature is changed at $\Delta T = R_t L$ K and the phase difference at $\Delta\Theta = \langle D\rangle L$ rad. For the allowed error of $\epsilon = 0.1^o$ in the temperature range of $T_r = 300^o$, the mean phase error is $\epsilon_p = \pi\epsilon / T_r = 1.047 \times 10^{-3}$ rad. By $\Delta\Theta = \epsilon_p$, the pulse-burst length is thus limited with $L \le \frac{\pi\epsilon}{\langle D\rangle T_r}\times 10^{-9} = 0.04\,\text{sec}$.

Example 2: Effect of the SNR on the mean drift rate [25]. Figure 7 shows effect of the SNR, by its equal values in each of the pulses, on $\langle D\rangle$ for $\bar{\Psi} = 5.55 \times 10^{-3}$ rad. Actual values are calculated by (88) for $\gamma \le 20$ and by (55) when $\gamma > 20$. Supposing that Ψ the Gaussian pdf (55) over all values of γ, we arrive at an approximation (dashed). For the latter case, the process was simulated and $\langle D\rangle$ calculated numerically (circles), by (89). As can be seen, the approximation errors practically vanish when γ exceeds 4 dB. Otherwise, bias occurs in the estimate. Table 1 gives the relevant values for $\bar{\Psi}$ ranging from 0.1π to 0.9π. A simple measure of accuracy used here is when the exact and approximate values become visually indistinguishable.

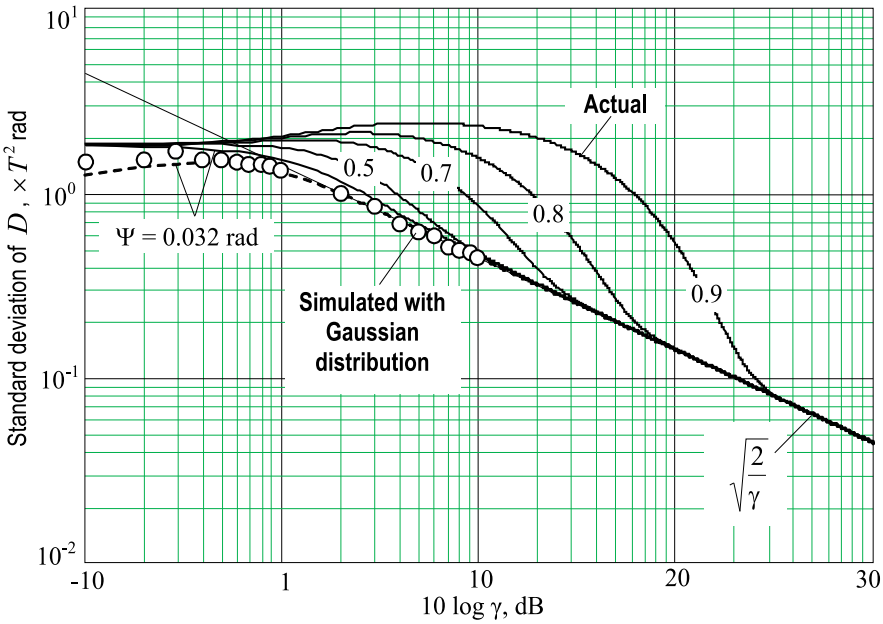


Fig. 8. The standard deviation $\sqrt{\sigma_D^2}$ for different $\bar{\Psi}$: actual (bold), by (90); Gaussian approximation (dashed) for $\bar{\Psi}$ = 0.032 rad; and simulation with Gaussian distribution, using (91).

γ	$\lvert\bar{\Psi}\rvert$, rad					
	0.1π	0.3π	0.5π	0.7π	0.8π	0.9π
dB	> 8	> 9	> 10	> 13	> 18	> 23

Table 2. Allowed γ for Accurate Evaluation of σ_D^2 with Different $\bar{\Psi}$

7.2 Drift rate variance

The phase difference drift rate variance can be represented, using (88), via the mean square value found in [11] to be [26]

$$\begin{aligned}\sigma_D^2 &= \frac{\left\langle \Psi_k^2 \right\rangle}{T^2} - \langle D \rangle^2 \\ &= \frac{1}{T^2}\left[\frac{\pi^2}{3} + 4\sum_{n=1}^{N}\frac{(-1)^n}{n^2}\frac{I_n(\alpha_1)}{I_0(\alpha_1)}\frac{I_n(\alpha_2)}{I_0(\alpha_2)}\cos n\bar{\Psi}_k\right] - \langle D \rangle^2\end{aligned} \tag{90}$$

and its estimate obtained by averaging as

$$\hat{\sigma}_D^2 = \frac{1}{TK}\sum_{k=1}^{K}(\hat{\Psi}_k - \langle D \rangle)^2. \tag{91}$$

Figure 8 sketches the standard deviation $\sqrt{\sigma_D^2}$ evaluated by (90) and (91) for different values of $\bar{\Psi}$. Along, we show the estimates (dashed) and simulated values calculated by

(91) for $\bar{\Psi}$ = 0.032 rad assuming Gaussian approximation (55). For the comparison with the mean drift rate (Table 1), Table 2 gives minimum values of γ, for several $\bar{\Psi}$, allowing for accurate evaluation of σ_D^2. An important inference follows instantly. For the sake of minimum errors, the SNR in the pulses must be obtained larger than 23 dB for $\bar{\Psi}$ ranging from -0.9π to 0.9π. We notice that similar values were found in [3] and [14] for the phase difference Θ. An analysis shows that the CRLB cannot be found for multiple DPM in simple functions and the best candidate for the estimate of σ_D^2 still remains (91).

8. Error probability for the drift rate to exceed a threshold

In applications, the phase difference drift rate is often required to range below some allowed value. The relevant error probability P_E can be characterized by the probability for the DPD to exceed a threshold ζ. Because the pdf of the modulo 2π angular measure is 2π-periodical, the P_E is commonly ascertained by setting $\bar{\Psi}_k = 0$. Using (51), we thus have

$$P_E \triangleq P_E(\zeta \mid \bar{\gamma}_1, \bar{\gamma}_2)$$

$$= 2\int_{\zeta}^{\pi} \bar{p}(\Psi_k \mid \bar{\gamma}_1, \bar{\gamma}_2)\mathrm{d}\Psi_k \tag{92}$$

$$= \frac{1}{\pi I_0(\alpha_1) I_0(\alpha_2)} \int_{\zeta}^{\pi} I_0\left(\sqrt{\alpha_1^2 + 2\alpha_1\alpha_2 \cos z + \alpha_2^2}\right)\mathrm{d}z. \tag{93}$$

Expanding the integrand in (93) to the Fourier series

$$I_0\left(\sqrt{\alpha_1^2 + 2\alpha_1\alpha_2 \cos z + \alpha_2^2}\right) = \sum_{n=0}^{\infty} \epsilon_n I_n(\alpha_1) I_n(\alpha_2) \cos nz,$$

where $\epsilon_0 = 1$ and $\epsilon_{n>0} = 2$, brings (93) to several useful estimates

$$P_E = \sum_{n=0}^{\infty} \frac{\epsilon_n}{\pi n} \frac{I_n(\alpha_1)}{I_0(\alpha_1)} \frac{I_n(\alpha_2)}{I_0(\alpha_2)} \left(\sin n\pi - \sin n\zeta\right) \tag{94}$$

$$\cong 1 - \frac{\zeta}{\pi}\left[1 + 2\sum_{n=1}^{N} \frac{I_n(\alpha_1)}{I_0(\alpha_1)} \frac{I_n(\alpha_2)}{I_0(\alpha_2)} \frac{\sin n\zeta}{n\zeta}\right] \tag{95}$$

$$\cong 1 - \frac{\zeta}{\pi}\left[1 + 2\sum_{n=1}^{N} \frac{I_n(\alpha_1)}{I_0(\alpha_1)} \frac{I_n(\alpha_2)}{I_0(\alpha_2)}\right], \zeta \ll \pi, \tag{96}$$

$$\cong 1 - \frac{\zeta}{\pi}\left(1 + 2N\right), \quad \zeta \ll \pi, \quad 1 \ll \bar{\gamma}_{1,2}, \tag{97}$$

$$\le 1 - \frac{\zeta}{\pi}\left(1 + 4\max \bar{\gamma}_{1,2}\right), \zeta \ll \pi, 1 \ll \bar{\gamma}_{1,2}. \tag{98}$$

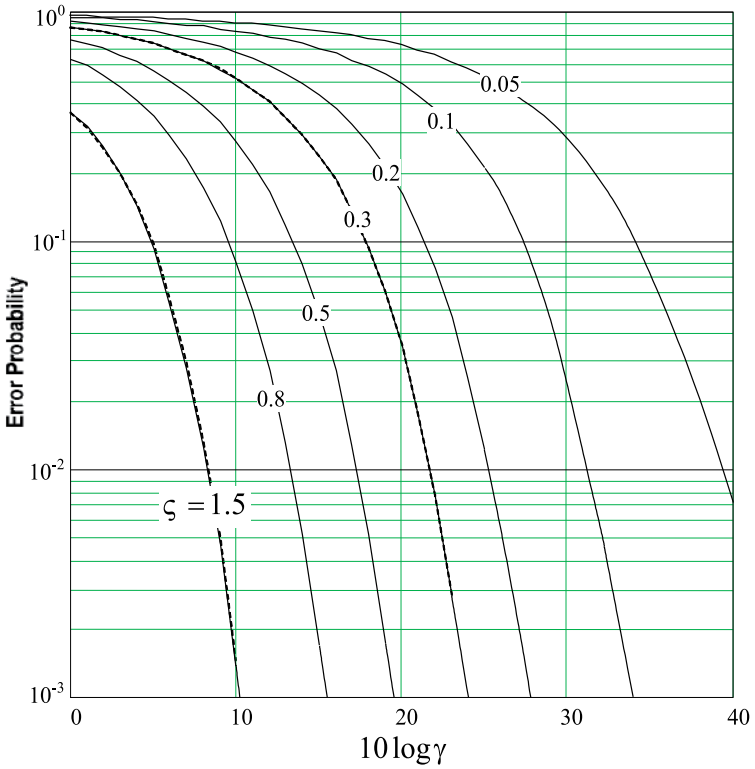


Fig. 9. Error probability of the drift rate to exceed a limit ζ with equal SNRs in the pulses: rigorous (dashed), by (37), and approximate (bold), by (95).

Figure 9 illustrates the error probability calculated rigorously (dashed), by (37) and (92), and approximately (bold), by (95) for equal SNRs in the pulses. One infers that the approximation error is negligibly small in the whole range of angular measures.

9. Conclusions

This Chapter gives a statistical analysis of errors in passive remote wireless surface acoustic wave sensing. By using the relations discussed, one can design the sensor reader system in an optimal way from the standpoint of maximum accuracy in measurement and minimum energy in the interrogating pulse. We were concerned with both rigorous and approximate estimates of the phase difference errors. It was shown that the rigorous pdfs cannot be represented in closed forms. In turn, the von Mises/Tikhonov-based densities are simple and reasonably accurate that makes them very attractive for engineering applications.

10. References

[1] Bulst, W.-E.; Fischerauer, G. & Reindl, L. (2001). State of the art in wireless sensing with surface acoustic waves. *IEEE Trans. on Industrial Electronics*, Vol. 48, No. 2, (April 2001)(265-271), ISSN 0278-0046.

[2] EerNisse, E.P. & Wiggins, R. B. (2001) Review of thickness-shear mode quartz resonator sensors for temperature and pressure, *IEEE Sensors J.*, Vol. 1, No. 1, (June 2001)(79-87), ISSN 1530-437X.

[3] Shmaliy, Y.S. (2000) Toward resonator anharmonic sensors for precision crystal oscillators: A Gaussian approach, *IEEE Trans. on Ultrason., Ferroel., and Freq. Contr.*, Vol. 47, No. 2, (Mar. 2000)(379-389), ISSN 0885-3010.

[4] Vig, J. R. (2001) Temperature-insensitive dual-mode resonant sensors – a review, *IEEE Sensors J.*, Vol. 1, No. 1, (June 2001)(62-68), ISSN 1530-437X.

[5] Schmidt, F.; Sczesny, O.; Ruppel, C. & Mágoni, V. (1996) Wireless interrogator system for SAW-Identification-Marks and SAW-Sensor components, *Proc. Int. Freq. Contr. Symp.*, pp. 208-215, ISBN 0-7803-3309-8, Hilton Hawaiian Village, June 1996, IEEE, Honolulu.

[6] Scholl, G.; Schmidt, F.; Ostertag, T.; Reindl, L.; Scherr, H. &Wolff, U. (1998) Wireless passive SAWsensor systems for industrial and domestic applications, *Proc. Int. Freq. Contr. Symp.*, pp. 595-601, ISBN 0-7803-4373-5, Ritz-Carlton Hotel, May 1998, IEEE, Pasadena.

[7] Pohl, A. A review of wireless SAW sensors, *IEEE Trans. on Ultrason, Ferroelect. and Freq. Contr.*, Vol. 47, No. 2, (Mar. 2000)(317-332), ISSN 0885-3010.

[8] Scheiblhofer, S.; Schuster, S. & Stelzer, A. 'Modeling and performance analysis of SAW reader systems for delay-line sensors, *IEEE Trans. Ultrason. Ferroel. Freq. Contr.*, Vol. 56, No. 10, (Oct. 2009)(2292-2303), ISSN 0885-3010.

[9] Shmaliy, Y. S. (2006) *Continuous-Time Signals*, Springer, ISBN-10 1-4020- 4817-3, Dordrecht.

[10] Reindl, L. M. & Shrena, I. M. Wireless Measurement of Temperature Using Surface Acoustic Waves Sensors, *IEEE Trans. on Ultrason, Ferroelect. And Freq. Contr.*, Vol. 51, No. 11, (Nov. 2004)(1457-1463), ISSN 0885-3010.

[11] Shmaliy, Y. S. Limiting phase errors of passive wireless SAW sensing with differential measurement, *IEEE Sensors J.*, Vol. 4, No. 6, (Dec. 2004)(819- 827), ISSN 1530-437X.

[12] Shmaliy, Y. S. On the multivariate conditional probability density of a vector perturbed by Gaussian noise, *IEEE Trans. Inform. Theory*, Vol. 53, No. 12, (Dec. 2007)(4792-4797), ISSN 0018-9448.

[13] Shmaliy, Y. S. Probability density of the phase of a random RF pulse in the presence of Gaussian noise, *Int. J. Electron. Commun.*, Vol. 63, No. 1, (Jan. 2009)(15-23), ISSN 1434-8411.

[14] Shmaliy, Y. S. Von Mises/Thkhonov-based distributions for systems with differential phase measurement, *Signal Process.*, Vol. 85, No. 4, (Apr. 2005)(693-703), ISSN 0165-1684.

[15] Tikhonov, V. I. The effect of noise on phase-lock oscillation operation, *Automatika i Telemekhanika*, Vol. 20, No. 9, (Sep. 1959)(1188-1196).

[16] Tsvetnov, V. V. Unconditional statistical characteristics of signals and uncorrelated Gaussian noises in two-channel phase systems, *Radiotekh. I Elektronika*, Vol. XIV, No. 12, (Dec. 1969)(2147-2159).

[17] Shmaliy, Y. S.; Shkvarko, Y. V.; Torres-Cisnerros, M.; Rojas-Laguna, R. & Ibarra-Manzano, O. A stochastic analysis of an anharmonic sensor phase response, *IEEE Sensors J.*, Vol. 3, No. 2, (Apr. 2003)(158-163), ISSN 1530- 437X.

[18] Pawula, R. F.; Rice, S. O. & Roberts, J. H. Distribution of the phase angle between two vectors perturbed by Gaussian noise, *IEEE Trans. Comm.*, Vol. COM-30, No. 8 (Aug. 1982)(1828-1841), ISSN 0090-6778.

[19] Bennett, W. R. Methods of solving noise problems, *Proc. IRE*, Vol. 44, No. 5, (May 1956)(609-638).

[20] Tsvetnov, V. V. Statistical properties of signals and noises in two-channel phase systems, *Radiotekh.*, Vol. 12, No. 5, (May 1957)(12-29).

[21] Pawula, R. F. On the theory of error rates for narrow-band digital FM, *IEEE Trans. Comm.*, Vol. COM-29, No. 11, (Nov. 1981)(1634-1643), ISSN 0090- 6778.

[22] Shmaliy Y. S. & Shmaliy, O.Y. Probability density of the differential phase difference in applications to passive wireless surface acoustic wave sensing, *Int. J. Electron Commun.*, Vol. 63, No. 1, (Jan. 2009)(623{631), ISSN 1434-8411.

[23] Karpov, A. F. Parameter estimation of the distribution function of a radio signal phase, *Radioelectron. and Commun. Systems*, Vol. 26, No. 7, (Jul. 1983)(23-29).

[24] Shmaliy, Y. S.; Ibarra-Manzano, O.; Andrade-Lucio, J. & Rojas-Laguna, R. Approximate estimates of limiting errors of passive wireless SAW sensing with DPM, *IEEE Trans. on Ultrason., Ferroel., Freq. Control*, Vol. 52, No. 10, (Oct. 2005)(1797{1805), ISSN 0885-3010.

[25] Shmaliy, Y. S.; Shmaliy, O.Y. & Ibarra-Manzano, O. Drift errors in passive remote wireless SAW sensing with multiple DPM, *IEEE Sensors J.*, Vol. 9, No. 7, (Jul. 2009)(774-781), ISSN 1530-437X.

Edited by Don Dissanayake

SAW devices are widely used in multitude of device concepts mainly in MEMS and communication electronics. As such, SAW based micro sensors, actuators and communication electronic devices are well known applications of SAW technology. For example, SAW based passive micro sensors are capable of measuring physical properties such as temperature, pressure, variation in chemical properties, and SAW based communication devices perform a range of signal processing functions, such as delay lines, filters, resonators, pulse compressors, and convolvers. In recent decades, SAW based low-powered actuators and microfluidic devices have significantly added a new dimension to SAW technology. This book consists of 20 exciting chapters composed by researchers and engineers active in the field of SAW technology, biomedical and other related engineering disciplines. The topics range from basic SAW theory, materials and phenomena to advanced applications such as sensors actuators, and communication systems. As such, in addition to theoretical analysis and numerical modelling such as Finite Element Modelling (FEM) and Finite Difference Methods (FDM) of SAW devices, SAW based actuators and micro motors, and SAW based micro sensors are some of the exciting applications presented in this book. This collection of up-to-date information and research outcomes on SAW technology will be of great interest, not only to all those working in SAW based technology, but also to many more who stand to benefit from an insight into the rich opportunities that this technology has to offer, especially to develop advanced, low-powered biomedical implants and passive communication devices.

ISBN 978-953-51-4543-1

Photo by jimmyan / iStock

IntechOpen