SPECTRUM AND PROPERTIES OF MESOSCOPIC SURFACE-COUPLED PHONONS IN RECTANGULAR WIRES

By

STEVEN EUGENE PATAMIA

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2001

Copyright 2001

by

Steven E. Patamia

ACKNOWLEDGMENTS

I would like to thank my advisor, Prof. Pradeep Kumar, for indulging my interest in pursuing this vestigial problem. I want to acknowledge publically that I understand the risk to both our careers that accompanied the continued pursuit of a fundamental and yet elusive problem which, despite its significance, has been declared without solution.

I want to thank Dr. Albert Migliori and the Los Alamos National Laboratories for supporting the completion of this dissertation and for confirming in a material way that there were physicists and scientific enterprises for whom the results truly mattered.

I want to thank my children, Mandy and Sarah, for their tolerance of my determination to pursue a PhD late in life.

On the chance that she will someday read this page, I want to express appreciation to Michelle for reasons only she will know.

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

SPECTRUM AND PROPERTIES OF MESOSCOPIC
SURFACE-COUPLED PHONONS
IN RECTANGULAR WIRES

By

Steven Eugene Patamia

December 2001

Chairman: Pradeep Kumar
Major Department: Physics

This dissertation presents original analytical derivations of the propagating modes of coupled mesoscopic phonons in an isotropic rectangular wire with stress-free surfaces. Incidental to the derivations, novel consequences of the derived cutoff modes are presented as they affect the low-energy heat conductance of such wires, or indeed any property that depends upon the dimensionality of the phase space within which the modes reside. Owing to nonseparability of the free-surface boundary conditions, an analytic description of coupled mesoscopic modes has heretofore been presumed to be underivable. Results presented herein show that the mode spectrum of coupled mesoscopic phonons is both subtle and rich, but considerable success in their analytic derivation is achieved. Using numerical methods developed for resonance problems, at least one contemporary researcher has purported to exhibit the lowest dispersion branches of propagating mesoscopic phonon modes in GaAs–which is not isotropic. The accuracy of these branches has not been measured, but they bear a qualitative consistency with isotropic modes derived herein. Since before the beginning of the 20th century, analytical solutions have been known for the infinite thin plate and even the

case of waveguides with circular cross sections. Solutions for these special cases take the form of transcendental relations among the wavenumber and boundary parameters, but the underlying wavefunctions are separable in the coordinates. The analytical results presented herein for the general rectangular case involve nonseparable solutions whose separable components do not individually satisfy the boundary conditions. These solutions also take the form of transcendental relations, but there are sets of transcendental relations for each family of the cases that partition the problem. Consequently, the eigenspectrum, while defined by exact forms, must be enumerated by identifying plotted intersections of the root families of these transcendental relations. The resulting spectrums are more complex and have less apparent order than the spectrum produced using either periodic boundary conditions or rigid boundary conditions for uncoupled phonons.

CHAPTER 1
INTRODUCTION

1.1 Nature of the Problem

Despite some progress in exhibiting the low-lying band structure of mesoscopic phonons by numerical methods, the study of phenomena involving phonon interactions has been hampered by an inability to analytically derive phonon modes involving surface coupling within rectangular geometries [1]. In unbounded bulk media, long-wavelength phonons are adequately represented by independent families of elastic plane waves distinguished by polarization. However, at nonrigid surfaces, elastic boundary conditions cannot in general be satisfied except by a coupled superposition of shear and longitudinal vibrations. In some geometries–particularly those involving edges and corners–this surface constraint cannot even be satisfied by separable wave functions let alone a derivable superposition of simple plane waves. Reflecting this difficulty, no general analytical solution has heretofore been published for the eigenvalues or wave functions of phonons in a mesoscopic wire of rectangular cross-section [2–4].

As contemplation shifts from bulk material to bounded samples with elongated geometries, the manifestation of long-wavelength phonons will include torsion and flexing of the sample. Bulk phonon models based on longitudinal scalar potentials alone are intrinsically incapable of exhibiting any flexural or torsional behavior since these require both some mechanism for shear and specific kinds of distortion of a specified surface geometry. Torsional and flexural behavior can be modeled to some extent using bending moduli and minimizing the potential energy associated

with the distortion of finite objects (See, e.g., [5] §4.12). Yet, correct torsional and flexural modes ought to be automatically included within a comprehensive phonon model which correctly identifies phonons as mixed longitudinal and shear vibrations which satisfy nonrigid boundary conditions within a specific geometry.

In modern terms, the nature of the problem can be succinctly described in the following useful way. Within the limits of linear elasticity, phonons in bulk are distinguished by polarization direction and propagation velocity and they do not interact. At any non-rigid surface, the reduction of external stress to zero is a boundary condition that can be satisfied in general only if the two species of modes become coupled at the surface. In effect, each species scatters into the other until an energy transfer balance is achieved. A detailed examination of the boundary conditions clarifies an important feature of this coupling. Namely, the nondegenerate longitudinal phonons polarized along their propagation direction (I am ignoring quasi-shear and quasi-longitudinal situations which can occur in non isotropic materials) only couple to shear phonons which have a polarization component normal to the surface. The surface interaction that arises as the surfaces are permitted to distort acts as a perturbation that splits the natural degeneracy of the shear phonons. Shear phonons polarized parallel to the surface are not affected and reflect specularly subject to their displacements vanishing at the surface. With respect to the phonon band structure, this degeneracy breaking is manifest as level repulsions among phonon dispersion subbands.

It should immediately be noted that degenerate perturbation techniques are not fruitful in deriving solutions for the coupled phonon problem. The reason for this is that such an approach depends upon an ability to find eigenfunctions of the perturbation itself. Pursuing this inevitably encounters the fundamental source of mathematical difficulty which plagues the problem generally. Namely, solutions

that incorporate the full boundary conditions at adjacent surfaces in rectangular geometries are easily shown to be non-separable. Defining a perturbation that reflects coupling of adjacent sides and then identifying a basis for its eigenfunctions become intractable.

1.2 History

Early in the history of elasticity theory, vibrational modes which accompany full relaxation of applied surface forces (formally, as the projection of the stress tensor onto a vector normal to the surface vanishes) were successfully derived for certain geometries. A solution was first obtained for circular, infinite-length bars by Pochammer [6] in 1875 and independently by Chree [7] in 1886. In 1889 Lord Rayleigh [8] published the analogous solution for an infinite plate with stress free surfaces.

In the foregoing cases, the "solutions" obtained took the form of so-called "frequency equations" which define a transcendental relationship among shear and longitudinal wave vector components. The roots of these equations, for each propagation wave number, constitute the eigenspectrum. Reflecting the difficulties inherent in analytically exposing the roots of such transcendental relations, only some asymptotic roots of the Pochammer-Chree frequency equations accompanied them immediately, and asymptotic solutions to Rayleigh's plate solution did not appear until it was revisited by Lamb [9] in 1917. As Lamb explores asymptotic roots and the displacement functions, he makes perhaps the first observation that, at increasing propagation wave numbers, two of the fundamental modes converge to form Rayleigh Surface Waves. Over succeeding decades, frequency equations for plates and rods were extensively studied and numerically derived roots of the

underlying transcendental relations are extensively characterized and well known in engineering applications [10–14].

It might seem surprising that despite difficulties identified above, that the frequency equation for circular waveguides was found at all, let alone before, the published analogous solution for the infinite plate. The reason is that a cylindrical coordinate system readily accommodates the fact that surface coupling only involves shear waves whose polarizations result in displacements normal to the surface. For cylindrical environments, the shear displacement field is naturally decomposable into a radial component coupled to the longitudinal field and a component parallel to the surface which is not. By contrast, for surfaces with sharp corners, only an unobvious superposition of waves forms (not necessarily plane) will accommodate coupling at adjacent faces. Making the geometry finite–whether with flat or rounded ends–further complicates this difficulty, as was noted even by Chree.

Once the wave guide takes on a rectangular cross section, the resultant boundary value problem develops pathologies linked to the fact that solutions are generally nonseparable in the coordinates. The lack of an analytical solution to the rectangular cross section problem has consequently been widely conceded in the literature (see e.g., [2–4]) In view of this perceived limitation, approximate solutions for various subsets of the modes for a bar have been developed–Timoshenko [15] "beam theory" being an important example in the realm of engineering applications.

In the absence of direct analytical solutions for the rectangular cross section, it remained possible to apply a variational approach to generate solutions that coincidentally satisfy the boundary conditions. The possibility of this tactic emerges in the literature at least as early as 1966 when Medick [16] attempted to develop a

1-dimensional wave propagation theory for rectangular bars. Medick reviewed the important result [17,18], already then a decade old, that solutions which render the Lagrangian stationary are precisely those which simultaneously satisfy the bulk wave and natural boundary conditions. Medick, however, did not apply the variational principle directly to numerical computation.

The basic observation that the eigenfrequencies of a freely vibrating elastic body are stationary in the space of functions was deduced early on by Lord Rayleigh who incorporated it into his treatise *The Theory of Sound* [19] in 1877. This result is usually referred to as "Rayleigh's Principle" [20]. Note that using Rayleigh's Principle to support a variational solution can be distinguished from approaches discussed within that rely explicitly upon stationarity of the time-independent Lagrangian as a manifestation of Hamilton's principle.

Early direct application of a variational principle to compute the low-order modes of parallelepipeds was presented by Holland [21] in 1968. Beginning with stationarity of the time-independent Lagrangian, Holland articulates an essentially Rayleigh-Ritz approach wherein finite combinations of trial functions are minimized. Holland's particular interest is in piezoelectric materials, but his computations are not in any way restricted by this. Holland examines the resulting eigenfunctions in an attempt to describe the displacement pattern of the modes themselves along with the spectrum. Holland's trial functions for components of the displacement field are products of sines and cosines of the coordinates. Later, Demarest (below) believed that such choices could precipitate inaccuracy in the results of a variational calculation. In recognition of the importance of this convergence issue, Holland and EerNisse examine this factor critically in a later monograph [22].

Building on Holland's 1968 result, Demarest [23] in 1971 improved the accuracy of Holland's mode computations for a free-surface cube–in part by changing the nature of the trial functions. The trial functions of Demarest's own Rayleigh-Ritz expansion are products of Legendre Polynomials of the coordinates. Along with identifying an inadequacy in Holland's choice of trial function, Demarest makes the observation that while Legendre Polynomials exhibit a computationally useful orthogonality in the relevant space, in principle they could be replaced by a power series in the coordinates of which Legendre Polynomials are just a linear combination.

Neither Holland's 1968 paper nor Demarest's 1971 paper derive the general relationship between satisfaction of the boundary-value condition and the efficacy of the variational principle utilized. Since they are able to demonstrate a satisfactory correspondence between actual measurement results and the numerical predictions, they appear to consider the theoretical, albeit important, question of boundary conditions to be moot. Demarest instead focuses on the fact that derivatives of his trial functions constitute a more "relatively complete" set of basis functions than Holland's. By "more relatively complete," he means that linear combinations of selections from this finite basis are able to converge to a broader range of possible functions. Holland, as Demarest points out, manipulates his trial function choices to favor convergence of their derivatives to zero at the boundary. Demarest's choice of Legendre Polynomials enriches the basis so that convergence both to zero displacements and zero derivatives of displacement are possible. While this is sufficient to improve the accuracy of numerical results as compared against measurements, it cannot be decisive. The actual boundary conditions which will be explored herein cannot be simply partitioned into cases involving zero displacements versus zero derivatives of displacements. Moreover, in most cases

unobvious superpositions are needed to assemble boundary solutions which are provably nonseparable and the nature of their boundary behavior is impossible to discern in advance of actually exhibiting them. Whatever its incremental improvement, Demarest's choice of basis also thwarts the expectation that the final solution, and presumably its components, must be constituted of no more than two parts which separately satisfy a basic wave equation–as is provable from Elasticity Theory.

Both Holland and Demarest utilize a Rayleigh-Ritz kind of procedure which approximates the solution to an eigenvalue problem by finding stationary points among a combination of trial functions. The conventional Rayleigh-Ritz procedure is implemented using trial functions which individually satisfy applicable boundary conditions (see any text on mathematical physics, e.g., [24] §17.8). This, however, is not possible under circumstances where there are no available basic solutions for the geometry that are known to satisfy those boundary conditions in the first place. It thus was openly expressed by Visscher, Migliori and Bell [25] in 1991 to be "fortuitous" that displacement functions which could be found to minimize the time-independent Lagrangian would automatically satisfy the bulk-wave and free-boundary conditions simultaneously. As I will point out more particularly within, this fortuity is less dramatically useful than its discovery infers. The reason is that it is only as helpful as the set of trial functions is versatile and is literally true only if the set of trial functions happens to include a combination that is a solution. Again, the central impediment is that basis functions which at least satisfy the boundary conditions do not even exist. Visscher et al., therefore, concentrated upon assembling a basis chosen for its versatility and the convenience with which numerical computations could incorporate its elements. In fact, the basis elements they choose, while linearly independent, are not orthogonal, and do not even satisfy

a wave equation let alone boundary conditions. Remarkably, at least in terms of resonances of closed objects, the basis is nevertheless sufficient to produce useful results to a sometimes high degree of accuracy [26].

Possibly because engineering and physics communities do not consistently overlap, Visscher et al. were probably not aware that Medick had previously noticed the same fortuity a quarter-century before and had been reviewing results available a decade before that. Nevertheless, Visscher, Migliori and Bell offer an extremely elegant derivation of the principle and clarify greatly the theoretical underpinnings of the algorithm developed by Holland and Demarest. Moreover, equipped with access to some of the earliest tools to bear the designation "supercomputer," Visscher et al. tested directly the efficacy of a particularly useful and elegant choice of basis functions. Specifically, they implement Demarest's suggestion that products of powers of the coordinates should constitute a set of basis functions that would be sufficiently "complete" to converge to accurate solutions–at least for the frequency eigenvalues in resonance problems.

Like Holland and Demarest, Visscher et al. were motivated by the need to find accurate low-order eigenmodes which could be verified by comparison to actual measured resonances of samples. They showed that an algorithm employing a basis built from simple products of powers of the coordinates would accomplish this even for geometric shapes far more challenging then the parallelepipeds of Demarest and Holland. Subsequently, Migliori has perfected this technique into an accurate standard means of finding the microscopic elastic constants of materials having arbitrary crystal or composite structure [26].

The directness of computability and broad applicability of the computational algorithm refined by Visscher, Migliori, and Bell has recently been adapted and applied to exhibit the modes of an infinite rectangular wire. Through the simple

expedient of replacing one coordinate factor of the basis functions with a common periodic function of the long coordinate alone,

$$\text{i.e.,} \quad x^p y^r z^q \quad \rightarrow \quad x^p y^r e^{iqz} \quad ,$$

Nishiguchi, Ando and Wybourne [1] resolved the transverse cross section components using the same variational methods as Visscher et al. while retaining translational invariance along the length of the infinite wire. Instead of obtaining a finite set of resonant eigensolutions for a resonator, he obtains a finite family of solutions for each value of q in a waveguide. they then generate the dispersion relations along the z-axis by plotting the resulting solutions which remain continuous in the parameter q. Since the wires are infinite length, there are no considerations of mixing from end-effects. There is also no need for the variational algorithm to resolve the modal pattern along the long dimension. As a consequence, convergence is limited only by the algorithm's ability to resolve the transverse displacements whose modest periodicity over a limited range is more readily approximated by a finite power expansion. Their specific results are computed for GaAs, but as Visscher et al. had emphasized, there is nothing in the nature of the numerical algorithm which inherently limits it to any material or crystal structure.

The mode dispersions developed by Nishiguchi et al. lack comparison with measurement. Directly measuring the eigenfrequencies of a resonator is straightforward, but directly measuring the long-wavelength subband dispersions of phonons subject to a specific geometric boundary may not be. The technique of finding bulk phonon dispersions using neutron scattering cannot be easily applied to a geometrically confined low-wavelength environment–particularly if the sample is microscopic. Also, full resonances are precise by definition and represent a severely constrained computational problem. It is not obvious whether a set of linearly

independent, but not inherently orthogonal, trial functions remain similarly able to converge strongly to a solution where the basis is constrained conceptually, but not actually, along one direction. These concerns will be further addressed in the results presented.

In general contemporary research, mesoscopic acoustic phonons have been investigated theoretically, and to some extent experimentally, in primarily two kinds of situations. The first is in respect of their interactions with phonons in quantum wires and the second is their role as carriers of thermal energy. The two topics have some overlap. In the next section, I will summarize the findings with respect to electronic interactions in quantum wires and in Ch. 7 I will take up the more emergent topic of kinetic heat transport by acoustic phonons in the low-temperature high-confinement regime.

1.3 Acoustic Phonons in Quantum Wires

1.3.1 Confinement

From the end of the 1980's and continuing into the present, mesoscopic physics has enjoyed an obvious prominence driven largely by the relentless downscaling of electronic devices. Commercial applications aside, the emergence of an ability to fabricate devices within which quantum behavior becomes observable as a consequence of dimensional confinement itself stimulates persistent and widespread interests in the effects of small size and low temperatures. Certainly, today, there is no more ubiquitous device in the theoretical and experimental condensed matter inventory as the "quantum wire."

In this context, so-called "phonon confinement" has also become a phenomenon of specific interest. Phonons, being a classical rather than quantum phenomenon

per se, are not subject to the sharp dimensional confinements possible for quantum particles, like electrons. When characterizing phonons as "confined," authors are only distinguishing phonon modes subject to quantization by physical boundaries from bulk phonons quantized using periodic boundary conditions. There is no inferred loss of dimensionality of the phonons, though, of course, the lowest subbands typically have imaginary transverse wave numbers and higher subbands may sometimes become energetically inaccessible.

While "confinement" at drastically small scales can lead to actual dimensional reduction of quantum particles, its generic effect is to render the physical boundaries relevant. The signature feature of what is called "mesoscopic" remains the commensurateness of the characteristic wavelength with the distance between boundaries. For phonons, however, actual small scales and low temperatures tend to magnify the departures from bulk behavior provoked by mesoscopic considerations. Certainly, at any scale, mesoscopic phonons exist since some subset of the phonon spectrum consists of phonons at wavelengths comparable to the size of the physical environment. At small scales and low temperatures, however, mesoscopic phonons are the dominant ones and their actual frequencies, which scale inversely with overall size, become large enough to augment the probability of electronic interactions.

Mesoscopic phonons are, by definition, of a wavelength large enough that surface effects and geometry control their dispersion and wave structure. Periodic boundary conditions and plane wave modeling give way to explicit boundary solutions for specific geometries. Surface coupling endemic to the non-rigid boundary conditions determines the eigenspectrum and mixes longitudinal and shear contributions.

1.3.2 Mesoscopic Electron-Phonon Interactions

Absent an analytical solution for surface-coupled phonons, research into confined phonon interactions progressed by using approximations or rough models which were believed to mimic the most important features of actual phonons. A brief survey of the results produced in this manner thus serves to identify conclusions and calculations which should be revisited in the light of analytical results from this investigation.

By the 1990's, papers explicitly calculating the interaction of electrons in a quasi-one-dimensional gas (i.e., in a quantum wire) with "confined phonons" were appearing regularly [27,28]. Initial investigations focused primarily on hot-electron relaxation via emission of "confined" optical phonons. The phonon models utilized during this time were typically uncoupled symmetric and antisymmetric transverse resonances where the transverse components were taken either to be zero (denominated "guided" modes) or a maxima at the boundaries (denominated "slab" modes). The latter condition is actually similar to stress-free conditions for elastic phonons. Results from the two flavors of boundary conditions are compared. Among the interesting conclusions is that boundary conditions resembling stress-free surfaces lead to an order of magnitude faster relaxation rate than obtained by assuming zero displacement boundaries [27].

Longitudinal optical phonons in polar semiconductors will decay to acoustic phonons (see Ref. [29] §6.2) within some average LO lifetime which can actually augment dissipation by virtue of precluding LO phonons in the process of decay from being absorbed back by the carriers. However, at low temperatures in quantum wires (e.g., below 30 K for a 100 Å wide GaAs wire) hot-electron scattering to LO phonons becomes exponentially weak and direct "acoustic phonon

emission is the only important dissipative process" [30] though for quantum wire wells, LO scattering may continue to dominate into lower temperature regimes [31].

Early in this period, theoretical interest in hot electron relaxation to acoustic phonons in wires was encouraged by experimental developments. In 1992 Seyler and Wybourne published a PRL [32] reporting on the detection of resonances with presumably acoustic modes in small (approximately 20 nm thick and between 30 and 90 nm wide) Au-Pd wires over a broad range of low temperatures (1-20 K). For the balance of the decade through to the present, this particular experimental observation is among the most consistently cited as justification for further theoretical exploration of the "importance of acoustic-phonon confinement in reduced dimensional electronic structures" [4].

The essential role of acoustic phonons in energy dissipation within quantum wires has been scrutinized in various ways. Senna and Das Sarma have investigated a "Giant Many-Body Enhancement" of electron-acoustic phonon coupling at low temperatures by renormalizing the phonons in the presence of the electron gas [33]. They find that at low temperature (viz. 1 K), below which direct plasma resonance cannot be significant, the quantum-mechanical (which they distinguish from thermal) uncertainty in the phonon modes creates an interaction enhancement orders of magnitude greater than for bare phonons. Mickevičius, Mitin, and Kochelap used Monte Carlo calculations to investigate phonon radiation from a Quasi-1D electron gas in rectangular GaAs quantum wires of 80x80 Å in the neighborhood of 4 K [34]. They report that virtually the entire power dissipation is due to the transverse radiation of ballistic acoustic phonons.

All of the above cited theoretical explorations involve GaAs and/or AlAs as the specific material of choice. While GaAs remained an almost predominant material chosen for theoretical studies during and since this time, Si has also been

explored. In a comprehensive set of calculations, Sanders, Stanton, and Chang calculated [35] a range of transport properties for a Si quantum wire–including deformation potential scattering by "quantum confined phonons." Although the authors explicitly reserve the term "acoustic" for the lowest subband, the phonon model used corresponds exclusively to longitudinal elastic waves coupled to electrons via a deformation potential. Possibly to avoid intimating that higher subbands of these confined elastic modes should be called "optical" the authors describe them as being "excited quantum-confined phonons." In any case, the focus in this paper is on a second-quantization representation of the phonons without attempting to mimic specific surface boundary conditions.

In a 1994 study, Yu, Kim, Stroscio, Iafrate, and Ballato [4] (liberally cited elsewhere herein) build upon the 1949 Ph.D. thesis work of Morse [36,3] to implement a more realistic confined acoustic phonon model that incorporates some surface coupling in a rectangular quantum wire environment. Upon reciting that "As is well known, there are no exact solutions for the complete set of phonon modes for a rectangular wire" Yu et al. adopt Morse's compromise strategy which was expected to be adequate when the cross sectional aspect ratio was greater than two. Morse had made progress solving the rectangular problem (which I investigate more successfully herein) by treating the closest parallel surfaces as an independent plate scenario (see section 4.4 herein) and finding what amounts to Rayleigh-Lamb modes relative to those surfaces. He then observed that two of the three surface-normal stress components become small at the adjacent surfaces as the aspect ratio grows and tunes a free parameter in his plate solution to force one of the two diminished stresses to be zero. Morse's transcendental equation derived from considering the closer of the surfaces (equation (14) in Ref. [4] under discussion) is very similar to, though less general than, my own intermediate equation (5.27)

which appears within my derivation of propagating coupled modes. His starting assumptions are far less general, but at that corresponding point of the derivations the situation is artificially similar. Restricting themselves to dilatational modes (see comments below), Yu et al. proceed to compute a normalization of Morse's separable solution and then compute electron-phonon scattering from deformation potential interaction for a range of cross sections (principally 28x57 and 50x200 Å) at 77 K. Their essential finding is that scattering rates are notably higher for such confined coupled acoustic modes than for bulk acoustic modes and increase dramatically as the overall scale is reduced. In fact, they compute that the scattering rate for a 28x57 Å wire are an order of magnitude larger than for a 100x200 Å cross section.

1.3.3 Confined Acoustic Phonons and the Deformation Potential

Hereafter, my own use of the label "acoustic" will, absent qualification, encompass all phonons that are a manifestation of elastic deformation. Consistent with much of the literature being reviewed (see, e.g., Ref [4] discussed above), this means that the subbands that arise from real boundary confinement of elastic waves will continue to be referred to as acoustic despite the fact that in bulk environments this term refers only to a dispersion branch that originates with zero frequency at zero k. My convention is appropriate to this study where only confined phonons are relevant, is consistent with texts and monographs devoted to acoustics of solids [5,10,37,38]–and it appears consistent with emergent text books devoted to phonons in nanostructures [29]. Unfortunately, past publications are not always as inclusive, but in the worst case it will simply be important to notice whether acoustic modes under discussion have been restricted to a specified dispersion branch or explicitly stated to be bulk-like. As is consistent in the

historical literature of acoustics per se, I will distinguish modes whose dispersion goes to zero as $k \to 0$ by calling them "fundamental."

Within the most inclusive use of the term, "acoustic" phonons are coupled to carriers solely by the deformation potential which tracks the ion density fluctuations. In each of the above-cited articles, only longitudinal (i.e., compressional) phonons (see, especially, [39]) participate in this coupling. Since bulk shear mode displacements can be generated by a vector potential, the divergence of their displacements is zero and they cannot alter the ion potential field via fluctuations in ion density.

There are, nevertheless, circumstances where a deformation potential can be associated with non-compressional phonons. For materials with degeneracies in band structure that can be broken by shear distortions, it is long been known that deformation potential should be generalized to a tensor such that the potential change due to deformation incorporates shear as well as divergence effects [40,41]. Even if such special circumstances do not apply, there remains a more fundamental reason to revisit the exclusive focus on compressional modes once the phonons are subject to boundary confinement.

In the continuum limit, acoustic phonons are merely elastic modes irrespective of whether the environment is confined or bulk-like. This study is concerned, however, exclusively with elastic modes confined to rectangular waveguides. As will be developed in detail in subsequent chapters, elastic waveguide modes can be partitioned with respect to the parity patterns of their displacements relative to a coordinate systems aligned with the long axis and normal to the sides. For the fundamental modes (i.e., those confined branches which do go to zero as $k \to 0$) these patterns correspond to macroscopic motion of the waveguide– dilatational, torsional, and flexural. Of these, the dilatational is so named because

the displacement pattern of the fundamental modes resembles propagating density fluctuations on the scale of the width of the waveguide. It is no doubt for this reason that as phonon models in quantum wires take boundary confinement into account, there is an explicit assumption made that only the part of the model corresponding to dilatational modes need be considered. Torsional and flexural modes correspond to twisting and bending and appear, on first impression, to be dominated by shear and are assumed to have no significant deformation potential (as assumed in Ref. [4]).

Transferring allegiance from longitudinal bulk modes to dilatational confined modes promotes a significant oversight. In bulk media, torsion and flexural displacement patterns do not even exist, but once the phonons are confined by actual surfaces, these categories arise precisely because surface coupling mixes longitudinal and shear contributions into the displacement pattern. Ignoring the torsion and flexural modes–perhaps for lack of an analytical model for the surface coupling–comes at the cost of ignoring definite sources of density variation. Aggravating the potential consequences of this oversight, it will later be seen that the branch shape of some fundamental torsional, along with the fundamental flexural modes, shows a high density of states at low frequencies.

Anyone who has ever wrung out a washrag realizes that fundamental torsional modes involve patterns of change in local density. Bending a foam rubber object creates a clear opportunity to observe the compressions and stretches which accompany that movement. Consistent with these real-life macroscopic observables, the derivations presented in subsequent chapters reveal that the longitudinal potential makes a necessary contribution to virtually every coupled mode. While it may turn out that some subset of modes within a subband are dominated either by shear or longitudinal patterns, there is no a priori justification for assuming that

any category of coupled modes can be safely ignored even if only the deformation potential from density fluctuation remains relevant. Even Demarest took care to point out that the dilatational, torsional, and flexural naming attached to parity patterns should not be taken too seriously as each contained elements of shear and dilatation [23].

CHAPTER 2
ASSUMPTIONS AND CONVENTIONS

2.1 Physical Model and Coordinate System

I assume that the waveguide is composed of an isotropic material formed into an infinite bar whose rectangular cross section is invariant along its length. The stiffness tensor for an isotropic material has only two independent elastic constants and these, together with the material density, determine the bulk shear and longitudinal velocities, denoted c_s and c_ℓ respectively [5]. To keep the results completely general, I will reduce all derived observables to a dimensionless form rescaled relative to the shear velocity, the smaller of the half-widths of the cross section, h, and the aspect ratio a between the widths. Derived results will thus be in terms of a single material characteristic–the ratio of velocities $R = c_\ell/c_s$–and a single geometric parameter–the cross sectional aspect ratio, a.

When computing definite eigenvalues to display representative quantitative results, R will be arbitrarily set to a value of $\sqrt{3}$. Some results will turn out to be independent of R. For perspective, it is a well known result of linear elasticity theory [42] that R cannot be less than $\sqrt{2}$ for isotropic materials. Commercial aluminum, a modestly soft and nearly isotropic metal, has an R of roughly 2 whereas GaAs, a non-isotropic semiconductor, has an R averaging close to the $\sqrt{2}$ limit.

All phonons are contemplated in the "mesoscopic" regime, by which I mean that their wavelengths are not less than an order of magnitude smaller than the smallest cross sectional width. I also assume that treating the material as an elastic

continuum is justified by first assuming the waveguide material has a typical inter lattice spacing much smaller than the smallest phonon wavelength considered. As a practical matter, this would still permit important results to apply to quantum wires with widths on the order of a few hundred atoms. The displacements will be assumed sufficiently small that applied elasticity theory is well within the linear regime. For phonons that are thermally excited or scattered from interactions with itinerate electrons, magnons, and similar particles, this is a reasonable physical assumption consistent with remaining within the phonons' own mesoscopic regime. Calculations and representations will be rendered within a right-handed Cartesian

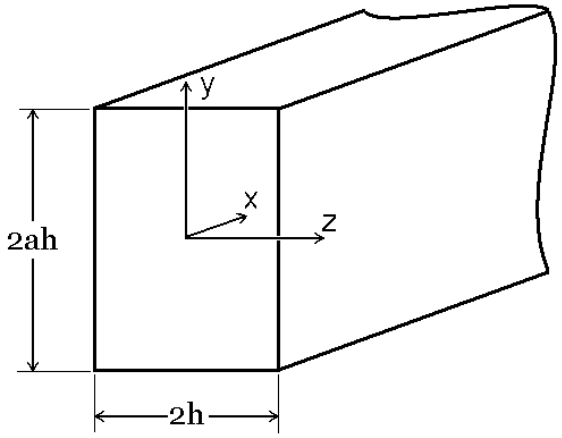

Figure 2.1: Nomenclature of Elastic Waveguide. Frequency and wave number will be rescaled relative to h (and c_s). The smaller halfwidth will be denoted $h_z = h$. The only geometric factor in final results will be a, the cross-section aspect ratio.

coordinate system. The long axis of the bar will be considered the x axis to facilitate easy comparison with publications of historical significance in which this is the more common convention. The long coordinate will be embedded along the geometrical center of the bar. This placement of the long axis symmetrically divides the bar. Accordingly, the bar will be transversely bounded by $-h_y \leq y \leq h_y$ and $-h_z \leq z \leq h_z$. For convenience, h_z will be consistently taken as the smaller half-

width should $h_y \neq h_z$ and h without subscript will refer to this smaller quantity. As already indicated, a will denote the cross sectional aspect ratio so that $h_y = ah_z$.

2.2 Symbolic Consistencies and Adopted Tensor Notation

In addition to foregoing nomenclature, the base symbol for all quantities related to elastic displacements will be the letter u. Vectors, as opposed to their components, will be bolded. When necessary to distinguish shear versus longitudinal displacement contributions, a parenthesized superscript will be used, as in $\mathbf{u}^{(\ell)}$ and $\mathbf{u}^{(\mathbf{s})}$. The displacement field will be decomposed into longitudinal and shear parts generated by a scalar potential φ and a vector potential $\mathbf{H}$ respectively. The Greek letter η will be consistently chosen as a base symbol for longitudinal wave numbers and σ for shear wave numbers so that these associations can be perceived at first glance.

Einstein notation will be used in tensor equations. Repeated indices will, unless otherwise noted, imply a summation over $\{x, y, z\}$ coordinates. Commas preceding one or more indices will abbreviate derivatives taken with respect to them. However, the problem will be investigated within a Cartesian metric and so there will be no distinction between covariant and contravariant vectors and correspondingly there will be no methodical use of raised versus lowered indices. Absent the ability to restrict summations to matching upper and lower indices, summation may be assumed for all matching pairs of indices on any one side of an equation—subject to contrary commentary.

The symmetric and antisymmetric parts of tensors with respect to any subset of indices will sometimes be explicitly indicated using the standard notational devices of enclosing the list of p indices that are symmetrized in parenthesis and any that

are to be antisymmetrized in square brackets.

$$A_{\ldots(ij\ldots)} = \tfrac{1}{p!}[A_{\ldots ij\ldots} + A_{\ldots ji\ldots} + \ldots]$$

$$A_{\ldots[ij\ldots]} = \tfrac{1}{p!}[A_{\ldots ij\ldots} - A_{\ldots ji\ldots} + \ldots]$$

$$H_{z,y} - H_{y,z} = 2H_{[z,y]} = -2H_{[y,z]}$$

The "totally antisymmetric" tensor of any appropriate rank will be indicated using an ϵ (as opposed to ε), as in

$$[\nabla \times \mathbf{H}]_i = \epsilon_{ijk} H_{j,k}$$

2.3 Linear Elasticity Theory

The phonons in this study are modeled as elastic vibrations. It is helpful and important to identify and outline the origin of specific elements of elasticity theory essential to the derivations developed herein. Besides clarifying notation and connecting it to standard literature, this will serve the important purpose of exposing implied and explicit assumptions which underlie my own derivations and which could limit applicability of the results.

The basic results of the linear theory are straightforward, but elasticity theory has many subtleties and complexities which will not be needed in what follows. For full development of the topic, monographs and texts range from comprehensive classics [43], to relatively recent standard texts which are cited by most contemporary researchers who utilize an elasticity model of long-wavelength phonons [37,38]. As a recommended supplement, a succinct development of elementary linear elasticity theory is provided within the theoretical physics series of

Landau and Lifshitz [42], but some lesser known texts and monographs are more relevant to the study of elastic waveguides [5,10].

Linear elasticity can be viewed as a generalization of Hooke's Law applied to a continuum whose elastic response need not be isotropic. Hence, the basic stress-strain relation law takes the following general form:

$$\sigma_{ij} = c_{ijk\ell}\, u_{k\ell} \tag{2.1}$$

The tensor $c_{ijk\ell}$ is sometimes called the "stiffness" tensor–in contradistinction to the "compliance" tensor whose elements are those of the inverse of the stiffness tensor's matrix representation. It is not uncommon to refer to the stiffness tensor as the elastic tensor. The tensor $u_{k\ell}$ is a dimensionless strain defined to encapsulate deformation in a form invariant under pure rotation while omitting non-linear terms. This is accomplished by defining it to be the symmetric part of the gradient of the displacement vector (sometimes denoted $\nabla_s \mathbf{u}$ [38]).

$$[\nabla \mathbf{u}]_{(k\ell)} = u_{k\ell} = u_{(i,j)} = \tfrac{1}{2}(u_{k,\ell} + u_{\ell,k}) \tag{2.2}$$

The potential energy density then has the following form:

$$V = \tfrac{1}{2} u_{k\ell}\, c_{ijk\ell}\, u_{k\ell} \tag{2.3}$$

Since u_{ij} is symmetric, directional invariance of V leads ultimately to essential symmetries of the elastic tensor as follows:

$$c_{ijk\ell} = c_{ijk\ell} = c_{jik\ell} = c_{ij\ell k} = c_{ji\ell k} \quad \text{and} \quad c_{ijk\ell} = c_{k\ell ij} \tag{2.4}$$

These intrinsic symmetries substantially reduce the number of unique components (81, before applying the symmetries) and give rise to an abbreviated notation, ubiquitous in materials science, for denoting the, at most 21, unique elastic constants which remain. Abbreviated indexing reflects the pattern of symmetries by denoting pairs with single digits. The strain, stress, and elastic tensor components are then expressed in terms of the correspondingly reduced number of indices. The abbreviation scheme is simply

$$\begin{array}{ccc} 11 \leftrightarrow 1 & 22 \leftrightarrow 2 & 33 \leftrightarrow 3 \\ 23, 32 \leftrightarrow 4 & 13, 31 \leftrightarrow 5 & 12, 21 \leftrightarrow 6 \end{array}$$

This allows equation (2.1) to be expressed as the following two-dimensional matrix-vector product relation:

$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} \\ & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} \\ & & c_{33} & c_{34} & c_{35} & c_{36} \\ & & & c_{44} & c_{45} & c_{46} \\ & & & & c_{55} & c_{56} \\ & & & & & c_{66} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix} \tag{2.5}$$

where omitted matrix elements are symmetric reflections. Various crystal symmetries will reduce further the number of unique elements of c_{pq}. Although I will sometimes rely upon and refer to general properties of c_{pq}, the important novel results depend upon the material being isotropic. Isotropy can be represented by imposing rotational invariance on an existing cubic symmetry. The full elastic

matrix for a cubic material (in abbreviated notation) is

$$\begin{bmatrix} c_{11} & c_{12} & c_{12} & 0 & 0 & 0 \\ c_{12} & c_{11} & c_{12} & 0 & 0 & 0 \\ c_{12} & c_{12} & c_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{44} \end{bmatrix} \tag{2.6}$$

which can be rendered rotationally invariant by requiring that

$$c_{44} = \tfrac{1}{2}(c_{11} - c_{12}) \tag{2.7}$$

The result is that the three components can be parameterized using only two constants (or moduli), and these can be chosen to be equivalent to the so-called Lamé constants which emerge naturally when the theory is derived starting from an assumption of isotropy (see, e.g., [42]).

$$\lambda = c_{12} \quad \text{and} \quad \mu = c_{44} \quad \text{with} \quad c_{11} = \lambda + 2\mu \tag{2.8}$$

For an isotropic material, the stress-strain relationship (2.1) can then be summarized as

$$\sigma_{ij} = \lambda \delta_{ij} \sum_r u_{rr} + 2\mu u_{ij} \tag{2.9}$$

Notice that I have not been content to allow u_{rr} to imply summation because one of the penalties of avoiding the use of upper indices is that, under the definition (2.2), u_{ii} is proper notation for an individual diagonal component of $\mathbf{u}$. Having called attention to this problem, it can be mitigated by adopting a convention

that, absent clarifying comments or explicit summation, repeated indices in a rank two tensor will imply summation (i.e., contraction) only when the usual letters set aside as coordinate variables are not used. Under this rule, the use of r as a subscript above in lieu of i, j, or k would have signaled that the ambiguity should be resolved in favor of implying summation.

The free-body equation of motion within the media follows immediately from Newton's law and substituting equation (2.1) into (2.3).

$$\rho\, u_{i,tt} = V_{,i} = \sigma_{ij,j} \tag{2.10}$$

If the right side of equation (2.10) is expanded using equations (2.1), (2.2), and the symmetries identified in (2.4), then the equation of motion can be expressed in terms of displacements as

$$\rho\, u_{i,tt} = c_{ijk\ell}\, u_{k,\ell j} \tag{2.11}$$

which, in the case of a normal mode (i.e., $e^{i\omega t}$ time dependency), immediately yields a form of vector wave equation with possibly non-factorable "velocities".

$$\rho\, \omega^2 u_i + c_{ijk\ell}\, u_{k,\ell j} = 0 \tag{2.12}$$

To exhibit the wave picture in the case of isotropy, I can use, instead, equation (2.9) in expanding equation (2.10). This leads ultimately to a useful coordinate-free equation of motion

$$\mathbf{u}_{,tt} = \left(\frac{\lambda+\mu}{\rho}\right) \nabla(\nabla\cdot\mathbf{u}) + \frac{\mu}{\rho}\nabla^2\mathbf{u} \tag{2.13}$$

This expression is "useful" in the sense that an assumption that the displacement field vanishes at infinity permits, via a well-known theorem of vector algebra, $\mathbf{u}$

to be separated into divergenceless and solenoidal parts which, in turn, produces separated wave equations with distinct velocities as follows:

$$\begin{aligned} \mathbf{u}^{(\ell)}_{,tt} &= \left(\frac{\lambda+2\mu}{\rho}\right)\nabla^2(\nabla\varphi) \\ \mathbf{u}^{(s)}_{,tt} &= \left(\frac{\mu}{\rho}\right)\nabla^2(\nabla\times\mathbf{H}) \end{aligned} \tag{2.14}$$

Here I have begun to utilize φ as the scalar potential generating the longitudinal displacement and $\mathbf{H}$ as the vector potential generating the shear displacement. In general, these potential symbols may themselves represent superpositions.

The foregoing clearly identifies the two isotropic bulk velocities:

$$c_s = \sqrt{\frac{\mu}{\rho}} = \sqrt{\frac{c_{44}}{\rho}} \tag{2.15}$$

$$c_\ell = \sqrt{\frac{\lambda+2\mu}{\rho}} = \sqrt{\frac{c_{11}}{\rho}} \tag{2.16}$$

These can then be substituted back into equation (2.9) to begin utilizing the bulk velocities as the primary characteristics of an isotropic material.

$$\sigma_{ij} = \rho\,(c_\ell^2 - 2c_s^2)\delta_{ij}u_{\ell\ell} + 2\rho\,c_s^2 u_{ij} \tag{2.17}$$

2.4 Free Boundary Conditions

The problem investigated herein is defined by taking the stress normal to the surfaces to be zero. I begin to transition toward a dimensionless form of this situation by first dividing equation (2.17) by $\rho\,c_s^2$. Recall that R has been designated to be the isotropic material property defined by c_ℓ/c_s. Then, in the coordinate system of the problem, the assertion that stress normal to the ith

surface at $x_i = \pm h_{x_i}$ is zero can be expressed in a form independent of ρ and c_s as

$$(R^2 - 2)\delta_{ij}u_{rr} + 2u_{ij} = 0 \quad \text{at } x_i = h_{x_i} \tag{2.18}$$

Since $u_{i,i} = u_{ii}$, the contraction u_{rr} in equation (2.18) is just $\nabla\cdot\mathbf{u}$. Since the divergence of a curl is zero, there is no shear contribution to this term. Then, since the longitudinal displacement is the gradient of the scalar potential φ, I can replace $\nabla\cdot\mathbf{u}$ with $\nabla^2\varphi$. Since φ is solution of a wave equation, we can finally replace $\nabla^2\varphi$ with $-k_\ell^2\varphi$ where $k_\ell = \omega/c_\ell$ at the eigenfrequency ω.

Finally, since each $u_i^{(\ell)} = \varphi_{,i}$ longitudinal strains are succinctly expressible in terms of the scalar potential in the form $u_{ij}^{(\ell)} = \varphi_{,ij}$.

Separating the shear and longitudinal parts and utilizing the foregoing, the boundary conditions at either of adjacent pairs of parallel surfaces can now be restated as a set of equalities relating shear and longitudinal contributions. To emphasize the essential structure independent of chosen surface, I will here use a surface-oriented notational scheme as follows: Denote with s the coordinate axes (i.e., y, or z) normal to these surfaces. With s as the starting coordinate, denote the next cyclic coordinate (right-handed progression) as p and the next after that as $\bar{p}$. In terms of this notation, the boundary conditions for any transverse surface can be written in the following form:

At $s = \pm h_s$ for all $p \in [-h_p, h_p]$ and all $\bar{p} \in [-h_{\bar{p}}, h_{\bar{p}}]$:

$$\begin{aligned} -\beta\ k_\ell^2\ \varphi + 2\varphi_{,ss} &= -2u_{ss}^{(s)} \\ \varphi_{,sp} &= -u_{sp}^{(s)} \\ \varphi_{,s\bar{p}} &= -u_{s\bar{p}}^{(s)} \end{aligned} \tag{2.19}$$

where $\beta = R^2 - 2$.

The shear displacement vector components in terms of **H** components are explicitly:

$$\mathbf{u^{(s)}} = \nabla\times \begin{bmatrix} H_x \\ H_y \\ H_z \end{bmatrix} = \begin{bmatrix} H_{z,y} - H_{y,z} \\ H_{x,z} - H_{z,x} \\ H_{y,x} - H_{x,y} \end{bmatrix} \tag{2.20}$$

The shear strain tensors are then:

$$\begin{aligned} u^{(s)}_{xx} = & \quad H_{z,xy} - H_{y,zx} \\ u^{(s)}_{yy} = & \quad H_{x,yz} - H_{z,xy} \\ u^{(s)}_{zz} = & \quad H_{y,zx} - H_{x,yz} \\ u^{(s)}_{xy} = & \quad \tfrac{1}{2}(H_{z,yy} - H_{y,yz} + H_{x,xz} - H_{z,xx}) \\ u^{(s)}_{yz} = & \quad \tfrac{1}{2}(H_{x,zz} - H_{z,zx} + H_{y,yx} - H_{x,yy}) \\ u^{(s)}_{zx} = & \quad \tfrac{1}{2}(H_{z,zy} - H_{y,zz} + H_{y,xx} - H_{x,xy}) \end{aligned} \tag{2.21}$$

with the order of differentiation having been arranged to highlight patterns and symmetries.

When tensors from equations (2.21) are substituted into surface manifestations of equation (2.19), the overall cyclic pattern of the subscripts is fully apparent and the boundary conditions can be resummarized completely in terms of potentials as follows:

At $s = \pm h_s$ for all $p \in [-h_p, h_p]$ and all $\bar{p} \in [-h_{\bar{p}}, h_{\bar{p}}]$:

$$\begin{aligned} -\beta\ k_\ell^2\ \varphi + 2\varphi_{,ss} &= -2(H_{\bar{p},p} - H_{p,\bar{p}})_{,s} \\ \varphi_{,sp} &= -\tfrac{1}{2}\left[(H_{\bar{p},pp} - H_{\bar{p},ss}) + (H_{s,s} - H_{p,p})_{,\bar{p}}\right] \\ \varphi_{,s\bar{p}} &= +\tfrac{1}{2}\left[(H_{p,\bar{p}\bar{p}} - H_{p,ss}) + (H_{s,s} - H_{\bar{p},\bar{p}})_{,p}\right] \end{aligned} \tag{2.22}$$

2.5 Nonseparability of Boundary Solutions

Utilizing equation (2.19), I will directly demonstrate that assuming the existence of separable potential functions in this scenario leads to an essential contradiction. The z-boundary and y-boundary conditions are explicitly

at $z = h$ for all $y \in [-h, h]$:

$$\begin{aligned} -\beta\ k_\ell^2\ \varphi + 2\varphi_{,zz} &= -2u_{zz}^{(s)} \\ \varphi_{,yz} &= -u_{yz}^{(s)} \\ \varphi_{,xz} &= -u_{xz}^{(s)} \end{aligned} \tag{2.23}$$

and at $y = h$ for all $z \in [-h, h]$:

$$\begin{aligned} -\beta\ k_\ell^2\ \varphi + 2\varphi_{,yy} &= -2u_{yy}^{(s)} \\ \varphi_{,yz} &= -u_{yz}^{(s)} \\ \varphi_{,xy} &= -u_{xy}^{(s)} \end{aligned} \tag{2.24}$$

Suppose that φ is separable wave function representable as the product of wave functions such as $p(kx)\psi_y(\eta_y y)\psi_z(\eta_z z)$. Since longitudinal strains are symmetrical in the derivatives taken to form them from φ, separability of φ will perpetuate to the longitudinal strains. The x dependencies of the potentials will be assumed to be common because that is the basis choice appropriate to propagating modes. I can thus anticipate that the x-dependencies of each boundary condition will cancel and will henceforth omit them from the remainder of this proof.

Assume also that $\mathbf{H}$ generates a $\mathbf{u^{(s)}}$ such that shear strains formed by combining their derivatives can be represented as separable products such as $u_{ij}^{(s)} =$

$u_{ij_y}(\sigma_y y) u_{ij_z}(\sigma_z z)$. Again, since $\mathbf{H}$ must be a wave function satisfying the boundary conditions, the separable products making up the strains it generates will ultimately also be wave equations for each direction.

If the shear components were individually separable, the best possible case would be that the shear strains which combine derivatives of these, nevertheless produce individually separable functions. If they do not, the shear strains will immediately be nonseparable and there will be no point in going further since the case will have been made. It is sufficient, therfore, to assume this best possible case.

Substituting these assumptions into equation (2.23) at $+h_z$:

$$\begin{aligned} -(\beta\, k_\ell^2 + \eta_z^2)\, \psi_y(\eta_y y)\, \psi_z(\eta_z h_z) &= -2k\, u_{zz_y}(\sigma_y y)\, u_{zz_z}(\sigma_z h_z) \\ \eta_y\, \eta_z\ \widetilde{\psi}_y(\eta_y y)\ \widetilde{\psi}_z(\eta_z h_z) &= -\sigma_y\, u_{yz_y}(\sigma_y y)\, u_{yz_z}(\sigma_z h_z) \\ k\, \eta_z\, \psi_y(\eta_y y)\, \widetilde{\psi}_z(\eta_z h_z) &= -u_{yz_y}(\sigma_y y)\, uyz_z(\sigma_z h_z) \end{aligned} \tag{2.25}$$

where the tilde's over functions conveys that they have been differentiated. It will not matter to the proof precisely what the separable product or its derivative is because the only facts essential to the result are that they are wave functions with definite wave numbers fixing their directional dependencies.

For the foregoing equations to be valid, they must each be invariant in y. By inspection, this will only be true if $\eta_y = \sigma_y$ including the possibility that they are both zero.

Given the cyclic symmetry connecting the boundary conditions at adjacent sides, substituting the same assumptions into equation (2.24) will similarly result in the symmetrical requirement that $\eta_z = \sigma_z$.

Separability of the potentials together with their satisfaction of wave equations implies that

$$\eta_y^2 + \eta_z^2 + k^2 = \frac{\omega^2}{c_\ell^2}$$

$$\sigma_y^2 + \sigma_z^2 + k^2 = \frac{\omega^2}{c_s^2}$$

If $\eta_y = \sigma_y$ and $\eta_y = \sigma_y$, this will require that $c_\ell = c_s$—which is impossible for real isotropic materials since it has already been pointed out that $c_\ell \geq \sqrt{2}\, c_s$. This calamity can be cured by (a) some superposition of solutions, (b) eliminating one side by extending it to infinity (the infinite plate scenario as solved by Rayleigh), or (c) special modal situations involving only shear ($\varphi = 0$). Each of these exceptions will be encountered in the sequel.

CHAPTER 3
THEORETICAL ASPECTS OF RECENT NUMERICAL METHODS

3.1 Stationary Lagrangian

Holland [21], Demarest [23], and Visscher et al. [25] each construct a theoretical justification for their numerical approximation methods by identifying solutions with the stationary points of a Lagrangian. More precisely, the starting point is stated to be Hamilton's principle, but the stationary points of the time-independent Lagrangian were asserted to be equivalent. Whether or not this is a valid substitution, however, depends on whether the function space within which the stationary functions are found is restricted to functions with a separable harmonic time dependence. (Demarest makes some attempt at varying a more general class of trial functions, but fails to properly factor the variation in his equation (4).)

This is not an inconsequential suspension of formality. Stationarity of the Lagrangian within a function space of elastic normal modes follows immediately from stationarity of the Hamiltonian precisely because the kinetic and potential functions have the same harmonic time dependency. Integrating any such function over time produces a factor of $1/2$ so that stating variations to be zero is equivalent to stating that variation of the time-independent part is zero.

$$\frac{\partial^2 \mathbf{L}}{\partial t^2} = -\omega^2 \mathbf{L} \;\Rightarrow\; \delta \int \mathbf{L}(r,t)dt = \frac{1}{2}\delta \mathbf{L}(r)$$

However, if the function space is not in some way restricted to those functions whose spatial parts are at least consistent with their having harmonic time dependencies, then it is elementary that stationarity of their spatial parts is not a

priori equivalent to the satisfaction of Hamilton's principle. The desired harmonic dependency will be manifest by satisfaction of a wave equation. After reviewing, below, the actual Lagrangian variational result, the point to be emphasized will be that restricting the set of trial functions to those which satisfy wave equations is indeed the most important prerequisite to validity of the result.

The Lagrangian variational analysis within each of the aforementioned numerical approximation attempts is well represented by that contained in the paper by Visscher et al. The reasoning is elegant and clear, and so it is helpful to reproduce it while making relevant observations.

The derivation by Visscher et al. takes place within a closed domain of arbitrary shape. It therefore is, in essence, a resonance problem. The same reasoning can be extended to propagating modes by considering such a mode virtually bounded within a rectangular waveguide by surfaces perpendicular to the sides and separated by a phase difference of 2π. If any surface integration is performed, reversal of the outward normals of these virtual surfaces would cancel their contributions to the overall surface integral insuring the same result as for a physically bounded region.

Visscher et al. pre-incorporate harmonic time dependence by writing the kinetic energy density as

$$\frac{1}{2}\rho\,\omega^2\,u_i\,u_i \tag{3.1}$$

They represent the potential energy density as

$$\frac{1}{2}c_{ijk\ell}\,u_{i,j}\,u_{k,\ell} \tag{3.2}$$

which can be shown to be equivalent to equation (2.3) by expanding the definitions of the strain tensors and exploiting symmetries of the elastic tensor. The

Lagrangian, then, is just the integral over the domain of the difference of these densities.

$$L = \frac{1}{2}\int_V \left(\rho\,\omega^2\,u_i\,u_i - c_{ijk\ell}\,u_{i,j}\,u_{k,\ell}\right) dV \tag{3.3}$$

I will assume, as do the authors, that the reader is able to complete the exercise by doing the algebra and applying the divergence theorem to find the conditions for the first variation of this Lagrangian to be zero. I will point out that their particular choice of how to represent the potential energy density appears not to have been accidental since it guides the algebra and makes the process rather more transparent. The result, as is expected from any Lagrangian minimization, is the equation of motion. However, the particular form of the surface term which must be zero is important and so I will write the full statement of the result to include it.

$$\delta L = 0 \Rightarrow \left[\int_V (\rho\,\omega^2\,u_i - c_{ijk\ell}\,u_{k,\ell j})\delta u_i\,dV + \int_S (c_{ijk\ell}\,u_{k,\ell})\delta u_i\,n_j\,dS_j\right] = 0 \tag{3.4}$$

where $\mathbf{n}$ is the outward normal surface vector and $n_j\,dS_j$ is the magnitude of the outward area vector. The parenthesized parts of the integrands are the elastic wave equation and a representation of stress. These parenthesized terms should be separately zero when the first variation of the Lagrangian is zero.

The same tensor algebra used to show that equation (2.3) is equivalent to Visscher's choice of representation for the potential energy density can be reversed to show that the integrand of the surface term in equation (3.4) can be replaced with

$$c_{ijk\ell}\,u_{k\ell}$$

which is just the right hand side of equation (2.1) by which stress is defined. Surface normal stress, by definition, is just

$$\sigma_{ij}\, n_j$$

and so the surface term in equation (3.4) is zero if the surface normal stress is zero.

With respect to the wave equation aspect of this condition, the assumption of harmonic time dependence guaranteed this in advance as can be seen from the fact that in section 2.3 I derive the same wave equation from that assumption and the application of Newton's Law. Lagrangian formalism is simply a substitute for Newton, but leads to the same result.

What is new is the appearance of the stress-free boundary condition as a prerequisite. There is nothing in section 2.3 that will directly permit me to infer that modal functions formulated to satisfy the elastic wave equation correspond to a stationary Lagrangian.

What Visscher et al. conclude from this outcome is that "... the displacement functions u, which are solutions to the elastic wave equation with free boundary conditions on S, are just those points in function space at which L is stationary." I want to emphasize yet again that this result is partly self-fulfilled by starting with an assumption of harmonic time dependence and that the satisfaction of the wave equation is actually an implicit starting point rather than a condition on the outcome.

In view of the foregoing, I draw a slightly different conclusion than that expressed by Visscher's team nearly a decade ago. What I see in the result is the following implication: Some superposition of functions which are guaranteed to satisfy the elastic wave equation will also satisfy free boundary conditions when their combination makes the Lagrangian stationary.

Having rephrased the conclusion it becomes easier to see how limitations can arise from the basis choices employed in actual numerical calcualtion. Visscher et al., like Demarest before them, choose a basis for computational convenience which does not in fact consist of functions which satisfy any wave equation. Resonances which involve a combination of components which are hyperbolic together with some having low periodicity can be accurately determined. It becomes more difficult to compute resonances for shapes, such as those having high aspect ratios, which require more components having extended periodicities. It is harder to assemble superpositions of waves from a basis devoid of wave solutions than it is from a basis made up of wave solutions.

In any case the paper continues with a logical inversion. It is claimed that finding a stationary point in some robust function space is paramount to finding a function that is a good approximation to a normal mode. If, of course, the basis is not made up of wave solutions, the stationary point must often simultaneously produce a wave solution, but failure to constrain the basis means that more is expected of the process than need be. It can only be said that concern for placing extra burdens on the process is lost in the fact that resonance problems usually converge nicely and give results which appear to be accurate for situations of practical interest [26].

That convergence of some kind does occur in the case of resonance problems is well documented in the literature cited. Indeed, the process has now been in used to determine elastic constants for a decade. I have personally observed the process and discussed it with Dr. Migliori. From his unpublished remarks, however, it is clear that the numerical process will sometimes fail to converge when attempting to match the actual eigenspectrum of objects with high aspect ratios against estimates generated by the numerical process. This difficulty appears not to have been

the subject of published methodical analysis, but I have personally experimented with rectilinear copper samples having roughly a 1x1x3 aspect and was unable to find elastic constants using which the numerical algorithm could produce a resonance spectrum that matched what was directly measurable. This author can therfore confirm that whatever its success, there are clear indications that the numerical algorithm premised on the Lagrangian formalism just reviewed is not reliable outside of some convenient set of situations despite remarkable accuracy when confined to that set of situations.

My purpose is not to bury the numerical techniques, nor even their theoretical underpinnings. Rather, I wish to praise them since they are extremely helpful and useful despite the limitations they operate under. It is important to repeat that whatever the precise limits of the numerical algorithms turn out to be, the fact is that they are demonstrably accurate within the domain they are typically deployed in. Moreover, the somewhat circular logic of the Lagrangian formulation carries within it the inspiration for a renewed investigation of whether analytical solutions are possible at least for the rectilinear geometry. Specifically, the Lagrangian technique highlights that the most promising basis ought to be one made up of functions which satisfy a consistent wave equation. By "consistent" wave equation, I simply mean any whose wave vector magnitude, and thus frequency, matches that of the normal mode in a propagating problem.

It is interesting that my observations on the limitations of Lagrangian minimization actually vindicates Holland's initial approach in that he purposely chose products of sines and cosines as basis functions. If Holland had found a way to implement his numerical approximation by separating his basis to expand the longitudinal and shear contributions separately, his algorithm would have been an appropriate numerical analog of the analytic approach which I develop herein.

3.2 Numerical Approximation

Equation (3.3) can be written as a difference of distinct functionals in the form

$$\mathbf{L} = T[\mathbf{u}] - V[\mathbf{u}] \tag{3.5}$$

I am now using $\mathbf{u}$ in the sense of time-independent displacement functions which inhabit a vector space equipped with the usual inner-product (i.e., vector dot product). T and V are linear functionals operating on the space of $\mathbf{u}$'s and I invoke a theorem of Functional Analysis (see Ref. [44] §30), usually taken for granted, that permits me to assume the existence of operators such that I can rewrite (3.5) in the form

$$\mathbf{L} = \langle \mathbf{u}|\mathbf{T}|\mathbf{u}\rangle - \langle \mathbf{u}|\mathbf{V}|\mathbf{u}\rangle \tag{3.6}$$

where the elements of a matrix representation of the operators are defined, as usual, by the results of their operation on a basis for the $\mathbf{u}$'s. This meticulous formality allows me to emphasize that this is a situation in which that basis is not actually known and that the situation is thus far abstract.

Actual numerical computation is accomplished using a finite and definite basis. Following Visscher et al., I will expand the displacements and estimate the operators using basis functions of the form $x^p y^q z^r$. This is just a computationally convenient realization of the more general use of products of functions of the coordinates such that each element is a separable function. These are not, in general, orthogonal, but they may be chosen to be linearly independent–as is the case here where simple powers of the coordinates are used. It will follow that a Rayleigh-Ritz like minimization of their combination is reducible to a generalized eigenvalue problem. One distinct and useful feature of this particular basis will be that its elements have definite parity.

Equation (3.2) which will form the basis for computing matrix elements of the potential energy operator, involves a mixing of different components of the displacement vector functions. I therefore expand the components of the displacement vector functions in bases of their own–rather than expand the displacement vectors in a basis of vector functions. Correspondingly, there are multiple (though perhaps not distinct) basis sets $\{\Phi_k^{(i)}\}$ $(i = x, y, z)$ and these result in an expansion of the form

$$\begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix} = \begin{bmatrix} \sum_k c_k^{(x)} \Phi_k^{(x)} \\ \sum_k c_k^{(y)} \Phi_k^{(y)} \\ \sum_k c_k^{(z)} \Phi_k^{(z)} \end{bmatrix}$$

In terms of these finite bases, approximations to the kinetic and potential operators have the following representations:

$$\mathbf{T}_{\lambda^{(i)}\gamma^{(j)}} = \int_{Vol} \rho \; \Phi_\lambda^{(i)} \; \Phi_\gamma^{(j)} \; \delta_{ij} \; dV \tag{3.7}$$

$$\mathbf{V}_{\lambda^{(i)}\gamma^{(k)}} = \int_{Vol} \partial_j \Phi_\lambda^{(i)} \; c_{ijk\ell} \; \partial_\ell \Phi_\gamma^{(k)} \; dV \quad \text{(sum over } j \text{ \& } \ell\text{)} \tag{3.8}$$

These matrix element definitions are in the same form as used by Visscher. They are elements of approximate, not exact, operators.

There is no assumption that basis functions individually satisfy the bulk wave equations of the elastic media, let alone the boundary conditions of the geometry. There is no reliance upon orthogonality properties of the basis functions in any geometry whatsoever.

One way to express that the Lagrangian is stationary, is to assert that the derivatives of equation (3.6) with respect to components of **u** are all zero. When the **u**'s are discrete, this certainly is equivalent to the following (functional differentiation

will lead to the same result):

$$\omega^2 \mathbf{T}|\mathbf{u}\rangle - \mathbf{V}|\mathbf{u}\rangle = 0 \tag{3.9}$$

This is just a statement of a generalized eigenvalue problem for which computer algorithms are widely available. The numerical algorithm generates a basis to some limit on the powers of the coordinates, calculates the matrix elements of the operators by integrating over the sample pursuant to equations (3.7) and (3.8), and then invokes standard software to calculate the eigenvalues and, if desired the eigenvectors.

This eigenvalue problem form of a minimization problem is expected to display a convergence of eigenvalues toward some limit as the order of the operators increases. It is common experience that this convergence is always of a decreasing nature so that the estimates are always upper bounds of the true values. However, when the basis functions used in the numerical procedures are merely products of powers of the coordinates and when the physical reality being modeled does not involve the converging eigenfunctions, or their derivatives, converging uniformally to zero at the boundaries, the expectation that eigenvalues are upper bounds may not be provable. A typical proof that a Rayleigh-Ritz procedure will consistently produce upper bounds depends upon the trial functions themselves vanishing at the boundaries [45]. In addition, although the numerical basis consists of elements that are linearly independent, they are not in general orthogonal as is assumed, for example, in a more elegant approach to a similar theorem [46] proposed by Peierls.

Of course, it may happen that some of the modes are well represented by functions that vanish at the boundaries. In Ch. 6 I will derive all of the $k = 0$ modes of a rectangular waveguide and a subset of these will be uncoupled and defined explicitly by the boundary zeros of sines and cosines. Powers of the coordinates

certainly converge rapidly to isolated sine and cosine functions–a feature endemic to numerical computations of their values. If the solutions consist of superpositions of sines and cosines, however, and if the boundary values are not well-represented, there are no such regularities to assist the rate of convergence and no vanishing of the wave functions or their derivatives to converge toward. While it is beyond the scope of this research to pursue these concerns further, I have attempted to point out that the choice of a non-orthogonal basis coupled with the fact that the boundary conditions do not generally involve vanishing of the basis elements, creates inherent difficulties for the efficacy of prevailing numerical methods. This strengthens the motivation for developing an analytical solution.

3.3 Partitioning the Problem into Parity Groups

Holland [21] carefully organizes his basis, made up of products of sines and cosines, into parity groups related to a coordinate system centered in his parallelepiped with axes normal to the faces. Demarest [23], and Visscher et al. [25] adhere to the same system. The same parity pattern classification system plays a pivotal role in my own analytical derivations which follow.

The importance of this classification scheme derives from the interplay of a rectilinear geometry with the symmetry properties of the elastic tensor (see section 2.3) within the potential energy operator defined in equation (3.8). The kinetic energy operator can be diagonalized independently, but careful examination of equation (3.8) shows that the potential energy operator will block diagonalize where the blocks are defined by parity patterns of any basis whose elements have definite parity. Among the consequences of this fact is that the eigenvalue problem stated in equation (3.9) can be divided into a distinct eigenvalue problem for each

block. From an analytical standpoint, however, the real value is that the symmetry classifications are an organizing principle for the entire analytical approach.

Although the parity classification scheme was first identified in the context of closed geometries, the potential energy operator does not change form in the case of a rectangular waveguide and so the classification scheme remains equally valid and applicable. The only distinction is that the parity along the long direction is often arbitrary as will be so indicated. When turning a waveguide into a resonator, the resonator modes symmetric along the former long axis will simply divide into even and odd groups. For the sake of comprehensiveness, I therefore exhibit the full parity pattern using the nomenclature of Holland.

Let E and O denote even and odd parity for functions relative to Cartesian coordinates aligned normal to the surfaces and along the central axis of the wire. An unspecified, but consistent, parity–as may be associated with the long axis of a waveguide–is denoted with a P. Parity complementation of an otherwise unspecified P (as, for example, resulting from differentiation) is denoted by $\bar{\mathrm{P}}$. Directional dependencies are implied by juxtaposition in a product–in the order x, y, z.

Symmetry patterns of displacement functions are then limited to only the families shown in Table 3.1–shown as column vectors of the x, y, z components.

The last two of these (flexural) are a degenerate pair in the case of square rectangular cross sections, but as pointed out by Nishiguchi [1], they are distinct for the general rectangular case.

Each of these parity patterns can be generated by a single product representing the parity pattern of the scalar potential whose gradient produces the longitudinal part of the displacement functions. The shear part must have a matching parity pattern and so the pattern of the scalar potential fixes that of the mode generally. In the case of Dilatation, for example, observe that the gradient of a product

Table 3.1: Mode Families in Rectangular Geometry Defined by Displacement Parity Patterns

$$\begin{bmatrix} \text{P} & \text{E} & \text{E} \\ \bar{\text{P}} & \text{O} & \text{E} \\ \bar{\text{P}} & \text{E} & \text{O} \end{bmatrix} \text{ Dilatation group: } \begin{cases} D, & \text{P=E} \\ F_1, & \text{P=O} \end{cases}$$

$$\begin{bmatrix} \text{P} & \text{O} & \text{O} \\ \bar{\text{P}} & \text{E} & \text{O} \\ \bar{\text{P}} & \text{O} & \text{E} \end{bmatrix} \text{ Torsion group: } \begin{cases} T, & \text{P=E} \\ S_1, & \text{P=O} \end{cases}$$

$$\begin{bmatrix} \text{P} & \text{E} & \text{O} \\ \bar{\text{P}} & \text{O} & \text{O} \\ \bar{\text{P}} & \text{E} & \text{E} \end{bmatrix} \text{ Flexion of z-x plane: } \begin{cases} S_2, & \text{P=E} \\ F_3, & \text{P=O} \end{cases}$$

$$\begin{bmatrix} \text{P} & \text{O} & \text{E} \\ \bar{\text{P}} & \text{E} & \text{E} \\ \bar{\text{P}} & \text{O} & \text{O} \end{bmatrix} \text{ Flexion of y-x plane: } \begin{cases} S_3, & \text{P=E} \\ F_2, & \text{P=O} \end{cases}$$

with parity pattern PEE automatically produces vectors with the parity pattern of the Dilatation group. Similarly, the gradient of POO will generate the Torsion group–and so on.

CHAPTER 4
MATHEMATICAL STRATEGY

4.1 Notation and Function Extension Issues

In section 3.3 I reviewed how, for rectangular geometries, the potential energy operator block diagonalized to partition the problem into independent solution families based on parity patterns of the basis. For this to be the case, it is only necessary that the basis functions individually display a definite parity in their separable parts. From this, any superposition of basis functions from the same family will also exhibit the same definite parity pattern even though the superposition becomes nonseparable.

My approach involves solving the boundary value problem in a way that takes advantage of the fact that solutions are partitioned by parity family. However, this does not imply that it is necessary to do a distinct derivation for each such family. Rather, as much as possible, each derivation will encompass all parity families in such a way that one algebraic result can be converted into a realization for each distinct parity family by a straightforward substitution. It thus is a distinct result that the solutions for each family are shown to be manifestations of a single theoretical result.

In order to be able to transcend distinctions between parity families for each derivation, it will be necessary to utilize a specialized notation. The notation will not violate other conventions of mathematical notation. It will sharpen rather than obscure important relationships. It will considerably shorten the expression of individual relationships, dramatically shorten derivations, and preclude redundant

derivations. Results expressed in this notation will unify the manifestations among parity families.

In its simplest characterization, the basis functions defined below will all be products of functions that are coincident with sines and cosines inside of the waveguide. It should be immediately pointed out that extending these functions to infinity is a tempting, but inadvisable option. It is true that vanishing of the density beyond the boundary would serve to keep the physical description realistic even if the displacement per se were to be so extended. However, there are distracting adverse mathematical consequences of indulging such an extension that should be avoided. One of these is that the Helmholtz Theorem, which is crucially relied upon to separate the displacement field into parts generated by a scalar potential and a vector potential, becomes problematic when the fields do not vanish. Either they should vanish totally at the boundary, or if extended, some kind of convergence factor should be inserted. This, however, leads to other complications. Ultimately, all such complications will be neatly avoided by a judicious choice of transform, but it remains the case that I will need the freedom to assume that the displacement fields have a behavior beyond the surfaces which will not contravene assumptions of the Helmholtz Theorem and at the same time I wish to avoid having to specify what that behavior is specifically.

Henceforth, the functions multiplied together to form displacement basis functions will themselves be defined as cosine or sine functions only between the surfaces. With respect to the stipulated coordinate system, these need to have definite parity, and so they will never have constant offsets to their phase at least in the transverse directions. It will prove an asset to intuition if I choose to simply

call such functions E and O in respect of their being either even or odd.

$$E_\eta^q = \begin{cases} \cos \eta q & |q| \leq h \\ \text{undefined} & |q| > h \end{cases} \tag{4.1}$$

$$O_\eta^q = \begin{cases} \sin \eta q & |q| \leq h \\ \text{undefined} & |q| > h \end{cases} \tag{4.2}$$

where h is the boundary limit and q is y, z, or some coordinate value (such as h).

To minimize notation, subscripts and superscripts can be used to indicate wave number and directional dependencies. Distinguishing x, y and z dependencies will sometimes be inferred by position if there is no risk of ambiguity.

$$\text{For example,} \quad E(\eta_x x)E(\eta_y y)E(\eta_z z) = E_{\eta_x}^x E_{\eta_y}^y E_{\eta_z}^z = E_x E_y E_z = EEE.$$

To implement generality in the derivations, I will need to denote functions of definite, but unspecified, parity by using a function variable. In general, the letter P, with appropriate subscripting to distinguish variables, will be used for this purpose. For example, $P_y P_z$ could take on the specific function values of OO, EE, OE, or EO.

Often, the mathematical structure of relationships will depend upon the relative parity of juxtaposed functions. I will accommodate this by denoting parity complementation of a function variable by placing a bar over it. For example, $P\bar{P}$ could be EO or OE.

Differentiation of E or O with respect to coordinate will be the dominant operation. Because of the sign change introduced by differentiation of a E within sample boundaries, it is not always correct to assume, for P_i as an abbreviation of $P(\eta_i x_i)$, that $P_{i,i} = \eta_i \bar{P}_i$. To nevertheless permit differentiations to be unambiguously specified at the highest level of abstraction, parity complementation

resulting from differentiation per se will be denoted by placing a tilde over a function variable. Specifically, $\widetilde{E}= -O$ while $\widetilde{O}= E$, and it is thus always correct to write $P_{i,i} = \eta_i \, \widetilde{P}_i$. Upon eventual substitution of E or O for P_i, the appropriate sign changes can be made. However, the symmetries of the problem ultimately result in the cancelation of sign distinctions and there are also cases of successive differentiation which invoke the parity-invariant rule $\widetilde{\widetilde{P}}= -P$.

The derivations will eventually reveal that the sign distinctions inchoate in notations like $\widetilde{P}$ are eliminated in the final results which can invariably be stated strictly in terms of parity variables and their simple compliments. Solutions specific to parity families will be realized simply by choosing E vs O assignments for at most two function variables (viz. P_y and P_z in the case of a waveguide with P_x being an additional variable only if the waveguide is capped to become a resonator). In order for this to be resolved in the derivations, the following additional notational device will be needed:

$$\begin{Bmatrix} x \\ y \end{Bmatrix}_{P_i} \equiv \begin{cases} x, & P_i = E \\ y, & P_i = O \end{cases} \tag{4.3}$$

Some simple examples that illustrate how this can be applied are:

$$\widetilde{\widetilde{P}}_y= \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_y} P_y \qquad \widetilde{P}_x= \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}_{P_x} \bar{P}_x \tag{4.4}$$

Therefore, differentiation of any P or $\bar{P}$ can be expressed without recourse to the tilde notation by using

$$P_{i,i} = \eta_i \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}_{P_i} \bar{P}_i \qquad \bar{P}_{i,i} = \eta_i \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_i} P_i \tag{4.5}$$

4.2 Defining the Basis and Superpositions

My key physical strategy is to assume that some superposition of tractable fundamental basis functions will assemble a tractable non-separable function that manages to satisfy the boundary conditions. While this is clearly an obvious approach to the predicament, I have been unable to find examples in the literature that reveal any attempt to actually apply it to the analytical solution of this problem. Though the assumption that some kind of superposition is needed can be inferred from various discussions, a failure to even realize the possibility is sometimes clearly evident (see, e.g., Ref. [3] §IV). I can only speculate that the unavailability of basis functions which themselves satisfy the boundary conditions has been viewed as such a serious departure from typicality that it has been more tempting to conclude unsolvability than to pursue superposition in spite of it.

A more courageous view has been exhibited by researchers looking for theoretical underpinnings of numerical approximation attempts. I therefore give credit for reinvigorating an analytic pursuit of superpositions to those who developed numerical approximation methods. These have been reviewed in the prior chapter in part because they create a framework in which superpositions and notions of how their elements should be structured take concrete shape. In fairness, it should thus be suggested that the results achieved herein are the result of adapting the progress made in numerical methods to a revisit of the presumably intractable analytic problem.

My own approach to formulating the basis functions is straightforward. I accept that the longitudinal and shear contributions should be expanded in separate bases. I depart from recent numerical approaches in that I require all basis functions to at least satisfy the bulk wave equation with respect to longitudinal or shear expansions. In fairness, it should be pointed out that Holland [21], by using

products of sine and cosine functions as trial functions, approached the problem similarly, but he did not segregate the trial functions into longitudinal and shear contributions. Neither did his successors, Demarest and the Visscher team, each of whom abandoned any efforts to constrain trial functions to those which satisfied a wave equation. Numerical approaches to date have sought an overall convergence of displacement arrangements guided by a variational principle and so the trend has been to impose increasing arbitrariness on the structure of the trial functions. Analytically, this is fruitless–or worse. The reason is that the trial bases used by Holland and his successors is not in an analytically "complete" representation.

The three key issues surrounding the construction of superpositions are (1) which fundamental basis functions to use, (2) how to represent the superpositions of these, and (3) how to transform the equations written in terms of the superpositions so that they can be solved as a finite set or solved by some recursive process operating on an infinite set. I have already indicated that I will expand both longitudinal and shear contributions in basis functions which individually represent wave solutions. What remains is to specify more concretely both the representation of these basis elements and the representation of their superpositions. In a subsequent section of this chapter, I will take up the transformation issue–also proposing a simple approach which has not appeared heretofore in the literature.

In derivations which follow, fundamental basis elements will always take the form of a separable product in the form $P_x P_y P_z$ where the P's (see preceding section on notation) stand for particular even or odd functions which, within the boundaries of the sample, are coincident with cosine and sine functions, respectively. The eight parity patterns corresponding to the eight mode families defined in table 3.1 correspond to the eight possible values of $P_x P_y P_z$. Note, however, that as long as each component of a superposition has a common parity pattern,

any superposition of them will exhibit the same parity pattern even though it is itself not a separable function. Also, since I am concerned with propagating modes characterized by translational invariance, P_x will naturally cancel out among the relations. (For a rectangular resonator, mode families could be generated by superimposing propagating modes of varying x-dependencies.)

Since the mode family is defined by a parity pattern of the overall displacement, the parity pattern of the shear components is constrained by the necessity that the parity pattern of $\nabla \times \mathbf{H}$ be identical to the parity pattern of $\nabla\varphi$. It is easily checked that for

$$\varphi \sim P_x P_y P_z$$

this constraint will be satisfied so long as the components of $\mathbf{H}$ are made up of fundamental elements which have parity patterns in terms of these as follows:

$$\begin{aligned} H_x &\sim P_x \bar{P}_y \bar{P}_z \\ H_y &\sim \bar{P}_x P_y \bar{P}_z \\ H_z &\sim \bar{P}_x \bar{P}_y P_z \end{aligned}$$

Fundamental basis elements (either longitudinal or shear) are linearly independent, in the intervals defined by the medium, for the obvious reason that sines and cosines with distinct wave numbers along any coordinate form a linearly independent set. Any linear combination of the basis elements must satisfy the bulk wave equation and this will be achieved so long as the elements individually satisfy "a" wave equation for the same magnitude wave number–which then factors. This is accomplished by requiring the wave numbers of terms in each basis element to satisfy the bulk dispersion relations:

$$k^2 + \eta_y^2 + \eta_z^2 = \frac{\omega^2}{c_\ell^2} \quad \text{for components of } \varphi \tag{4.6}$$

$$k^2 + \sigma_y^2 + \sigma_z^2 = \frac{\omega^2}{c_s^2} \quad \text{for components of each } H_i \tag{4.7}$$

This constraint reduces the degrees of freedom by one. I contemplate a mode with a fixed k, and choose one transverse direction as representing a degree of freedom while constraining the second according to equations (4.6, 4.7).

These functions can be treated as orthogonal with respect to each coordinate, but this will only be exploited obliquely in the sense that this property is deeply buried in the nature of the transform that will be applied to their superpositions.

A superposition can be modeled as discrete or continuous. On physical grounds, however, the discrete distribution is the correct one in this case. The physical reality being modeled consists of elastic waves in bounded media. Although the boundary conditions will be viewed through the lens of mathematical abstraction, the reality is that elastic waves refract at the boundaries and any superposition models the summation process that superimposes all of their reflections. Since there are a finite set of surfaces, each refraction is a discrete event. The superposition is thus a discrete summation.

Accordingly, the scalar potential will have the form

$$\varphi = P_x(k\,x) \sum_i d_i P_y(\eta_i\,y) P_z(\eta_i^*\,z) \tag{4.8}$$

where the $*$ superscript denotes a "longitudinal conjugation" defined by

$$\eta_i^* = \sqrt{\omega^2/c_\ell^2 - k^2 - \eta_i^2} \quad \text{positive root} \tag{4.9}$$

For the shear superpositions, the expansions take the form

$$H_x = P_x(k\,x) \sum_j a_j \bar{P}_y(\sigma_j\,y) \bar{P}_z(\sigma_j^+\,z)$$

$$H_y = \bar{P}_x(k\,x)\sum_j b_j P_y(\sigma_j\,y)\bar{P}_z(\sigma_j^+\,z)$$
$$H_z = \bar{P}_x(k\,x)\sum_j c_j \bar{P}_y(\sigma_j\,y)P_z(\sigma_j^+\,z) \tag{4.10}$$

where the + superscript denotes a "shear conjugation" defined by

$$\sigma_j^+ = \sqrt{\omega^2/c_s^2 - k^2 - \sigma_j^2} \quad \text{positive root} \tag{4.11}$$

Note that, in the foregoing, both η_i's and σ_i's can range over real and imaginary values within the same superposition.

There is a further, important, observation. The foregoing sums contain weighted terms that are solely products of sines and cosines–at least within their boundary domains. Since each such term has definite parity with respect to the sign of the wave number in its arguments, the effect of the sign of the wave number is always factorable, in the sense of

$$P_i(-k\,x_i) = \left\{ \begin{matrix} 1 \\ -1 \end{matrix} \right\}_{P_i} P_i(k\,x_i)$$

A series of terms which differ only in the combination of the signs of the wave numbers in the arguments will always factor into a single term with positive wave number arguments multiplying a sum and difference of coefficients. The combination of coefficients can always be absorbed into a single coefficient.

The result is that in each of the foregoing sums, I always choose a representation in which only positive (albeit real or imaginary) wave numbers are summed over, but for which some of the coefficients may be negative.

4.3 Dimensionless Representations

In section 2.1 the groundwork was laid for reporting the results of derivations in a dimensionless form. The specific scale factors and notation for this are as follows:

First, recall that of the rectangular half-widths, the h_z will be arbitrarily notated as the smallest one whenever $h_y \neq h_z$. The appearance of an unsubscripted h will refer to h_z. The dimensionless aspect ratio a will be h_y/h_z. Of the two isotropic velocities, c_s will used to rescale results and the ratio c_ℓ/c_s is, as noted, denoted R.

Frequency will be rescaled by the rule

$$\Omega = \frac{\omega}{\omega_o} \quad \text{where } \omega_o = \frac{c_s}{h} \tag{4.12}$$

This is a common rescaling in the historical literature of the problem and, in addition, it is common to report Ω in units of $\pi/2$.

Consistent with this scheme, wave numbers will be put into compatible dimensionless units by multiplying them by h. There is no absolute rule on the choice of symbols for longitudinal versus shear wave numbers in dimensionless units, but the general attempt will be to show dimensionless forms by converting Latin letters from lower to upper case and choosing distinct Greek symbols to convert existing Greek symbols to dimensionless form. For example, $K = kh$, $\alpha = \sigma\, h$ and $\beta = \eta\, h$ are common choices. The bulk wave dispersion relations in dimensionless terms would then take forms as follows:

$$\begin{aligned} \alpha_y^2 + \alpha_z^2 + K^2 &= \Omega^2 \\ \beta_y^2 + \beta_z^2 + K^2 &= \frac{\Omega^2}{R^2} \end{aligned} \tag{4.13}$$

In some derivations, a free wave-number variable, λ, will be used. In its dimensionless form, it will be denoted Λ.

The ease with which dimensionless forms of results can be written from inspection of dimensional ones will become apparent as examples appear.

4.4 Derivation of Rayleigh-Lamb Equation

In the Introduction, it was pointed out that in 1889 Lord Rayleigh was able to derive a transcendental "frequency equation" which defines the propagating modes of an isotropic infinite plate. This equation is now universally referred to as the "Rayleigh-Lamb Equation" and the dispersion patterns it generates are often called "Lamb Waves". Derivations, and introductions designed to promote reader-completed derivations, appear often in the literature, but most of these expositions are considerably more cumbersome than the derivation about to be demonstrated (see Rayleigh [8], Miklowitz [10], and Ch. 10 of Auld [38]). Nevertheless, the form of this equation is of fundamental importance in the results to follow and it will be helpful to demonstrate how it can be re-derived succinctly. Besides producing a needed result, this exercise will provide a clarifying example of the devised specialized notation at the same time it introduces the basic pattern for novel derivations which follow.

The physical scenario consists of an isotropic elastic material sandwiched between infinite planes at $z = \pm h$. The surfaces are stress free. A plane-wave system propagates along the x direction. At any x position, there are no variations along the y directions.

Referring to boundary conditions (2.22), $s \to z$, $p \to x$, $\bar{p} \to y$:

At $z = \pm h_z$ for all y and all x:

$$\begin{aligned} -\beta\ k_\ell^2\ \varphi + 2\varphi_{,zz} &= -2(H_{y,x} - H_{x,y})_{,z} \\ \varphi_{,yz} &= \tfrac{1}{2}\left[(H_{x,yy} - H_{x,zz}) + (H_{z,z} - H_{y,y})_{,x}\right] \\ \varphi_{,xz} &= -\tfrac{1}{2}\left[(H_{y,xx} - H_{y,zz}) + (H_{z,z} - H_{x,x})_{,y}\right] \end{aligned} \tag{4.14}$$

However, since $\partial_y \to 0$, the foregoing will simplify dramatically. In addition, the simplest vector potential that will generate a shear wave with no y displacement, is just:

$$\mathbf{H} = \begin{bmatrix} 0 \\ H_y \\ 0 \end{bmatrix}$$

and so I can set $H_x = H_z = 0$. The boundary conditions above now collapse to

$$\begin{aligned} -\beta\ k_\ell^2\ \varphi + 2\varphi_{,zz} &= -2H_{y,xz} \\ \varphi_{,zx} &= -\tfrac{1}{2}(H_{y,xx} - H_{y,zz}) \end{aligned} \tag{4.15}$$

Assuming that I will not need superpositions, individual basis elements that produce potentials able to satisfy the wave equation are (according to my already reasoned basis characterization above) simply

$$\begin{aligned} \varphi &= D\, P_x(k\,x)\, P_y(\eta y)\, P_z(\eta^* z) \\ H_y &= A\, \bar{P}_x(k\,x)\, P_y(\sigma y)\, \bar{P}_z(\sigma^* z) \end{aligned}$$

With D and A as unknown constants. But if $\partial_y \to 0$, it must be that $\eta = \sigma = 0$, and $P_y(0) = E(0) = 1$ and so the potentials can be further simplified to

$$\begin{aligned} \varphi &= D\, P_x(k\,x)\, P_z(\eta^* z) \\ H_y &= A\, \bar{P}_x(k\,x)\, \bar{P}_z(\sigma^+ z) \end{aligned} \tag{4.16}$$

Substituting equations (4.16) into (4.15) and performing the differentiations at $z = \pm h$ produces the following simultaneous equations:

$$\begin{aligned} -D\,(\beta k_\ell^2 + 2(\eta^*)^2)\,P_x(k\,x)\,P_z(\eta^*\,h) &= -2A\,k \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_x} P_x(k\,x)\,\sigma^+ \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_z} P_z(\sigma^+ h) \\ D\,k \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}_{P_x} \bar{P}_x(k\,x)\,\eta^* \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}_{P_z} \bar{P}_z(\eta^*\,h) &= A\,\frac{1}{2}(k^2 - (\sigma^+)^2)\,\bar{P}_x(k\,x)\bar{P}_z(\sigma^+\,h) \end{aligned} \quad (4.17)$$

At this point it is trivial to divide one equation into the other to eliminate the unknown coefficients as well as all of the x-dependencies. The result is

$$\frac{(\beta k_\ell^2 + 2(\eta^*)^2)\,P_z(\eta^*\,h)}{k\,\eta^*\,\bar{P}_z(\eta^*\,h)} = \frac{4\,k\,\sigma^+\,P_z(\sigma^+ h)}{(k^2 - (\sigma^+)^2)\,\bar{P}_z(\sigma^+\,h)} \quad (4.18)$$

It is gratifying to notice that the two conflicting sign contingencies for P_x simply each resolve to an assured minus sign–and the P_z sign contingencies neutralize each other. This is a pattern that will repeat itself in the more involved derivations.

The result is already in the form of a "frequency equation," but besides some rearrangement, there is a final analytical step to be performed. Since the wave systems are made "plane"by virtue of $\eta = \sigma = 0$, the conjugations are:

$$\begin{aligned} (\eta*)^2 &= \omega^2/c_\ell^2 - k^2 \\ (\sigma^+)^2 &= \omega^2/c_s^2 - k^2 \end{aligned} \quad (4.19)$$

Recall that $\beta = R^2 - 2$ with $R = c_\ell/c_s$. Applying these to the parenthetical on the left of equation (4.18), it can be restated more usefully.

$$(\beta k_\ell^2 + 2(\eta^*)^2) = (R^2 - 2)\frac{\omega^2}{R^2 c_s^2} + 2(\frac{\omega^2}{R^2 c_s^2} - k^2) = -\left(2k^2 - \frac{\omega^2}{c_s^2}\right) = -(k^2 - (\sigma^+)^2)$$

The details of such steps are indulged here because they are prototypical of rearrangements that recur in subsequent derivations, but which will not be hereafter presented in detail. With this particular re-expression substituted, and following some rearrangement of terms, equation (4.18) can be written in the form

$$\frac{P_z(\eta^* h)\,\bar{P}_z(\sigma^+ h)}{\bar{P}_z(\eta^* h)\,P_z(\sigma^+ h)} = -\frac{4\,k^2\,\eta^*\,\sigma^+}{(k^2-(\sigma^+)^2)^2} \tag{4.20}$$

Equation (4.20) is, in fact, the Rayleigh-Lamb frequency equation in a representation which encapsulates both its alternative forms. The so-called symmetric form follows from setting $P_z = E = \cos$ in which case the left hand side becomes $\tan(\eta^* h)/\tan(\sigma^+ h)$. If $P_z = O = \sin$, the left hand side becomes $\tan(\sigma^+ h)/\tan(\eta^* h)$ and the equation is said to be in its antisymmetric form. Accordingly, the Rayleigh-Lamb equation is often written in the form

$$\frac{\tan(\eta^* h)}{\tan(\sigma^+ h)} = -\left[\frac{4\,k^2\,\eta^*\,\sigma^+}{(k^2-(\sigma^+)^2)^2}\right]^{\pm 1} \tag{4.21}$$

In this representation, the symmetric and antisymmetric forms correspond to the exponent on the right being positive or negative respectively. The names given these forms obviously match the parity of P_z in my specialized notation, but that is not why they were so-named. If the displacement patterns corresponding to these equations are mapped, it is readily seen that the symmetric form corresponds to "dilatations" in which the plate surfaces are either extended or indented together at each x position. In the antisymmetric case, the sides of the plate are alternately extending or indenting–giving rise to a ripple effect. Indeed, the "antisymmetric" solutions are flexural in nature.

Finally, the Rayleigh-Lamb equation in dimensionless form (see section 4.3) would be

$$\frac{P_z(\beta)\,\bar{P}_z(\alpha)}{\bar{P}_z(\beta)\,P_z(\alpha)} = -\frac{4\,K^2\,\beta\,\alpha}{(K^2-\alpha^2)^2} \tag{4.22}$$

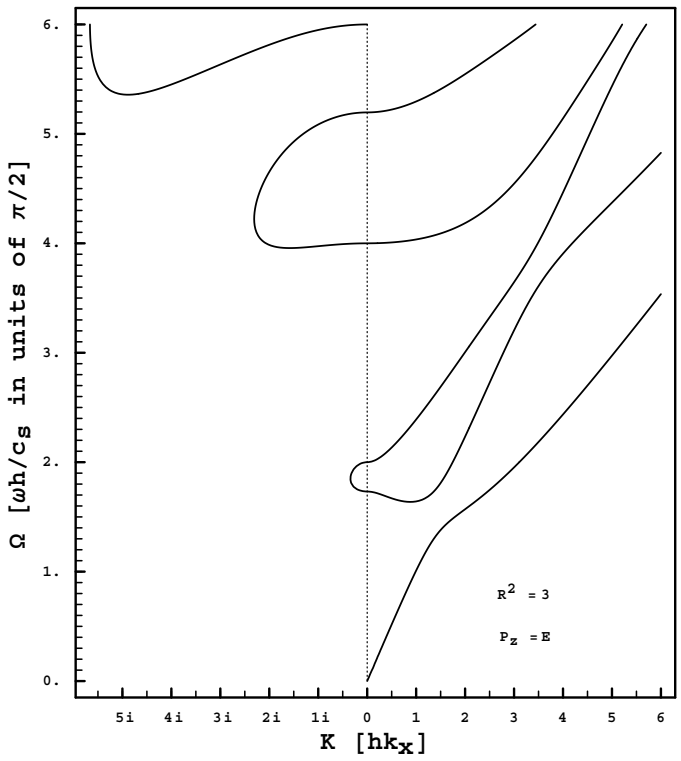

Figure 4.1: Rayleigh-Lamb Dilatational Modes. In the historical development of the solution, these are referred to as the "symmetric" modes.

The roots of this equation can be mapped as dispersion curves using contour plotting. There will be a set of branches for $P_z = E$ and another for $P_z = O$. A rearrangement of the equations is necessary to preclude fatal divergences in the numerical computation involved in the plotting. In terms of the dimensionless representation, numerical plotting is based on finding the roots of

$$\begin{Bmatrix} \cos(\beta)\frac{\sin(\alpha)}{\alpha} \\ \frac{\sin(\beta)}{\beta}\cos(\alpha) \end{Bmatrix}_{P_z} [(K^2-\alpha^2]^2 + 4 \begin{Bmatrix} \beta \\ \alpha \end{Bmatrix}_{P_z} K^2\,\bar{P}_z(\beta)\,P_z(\alpha) = 0 \tag{4.23}$$

Figure 4.1 shows branches of the dilatational or "symmetric" modes resulting from substituting $P_z = E = \cos$ into dimensionless Rayleigh-Lamb equation. Figure 4.2 shows the branches of the flexural or "antisymmetric" modes resulting from substituting $P_z = O = \sin$ into the equation.

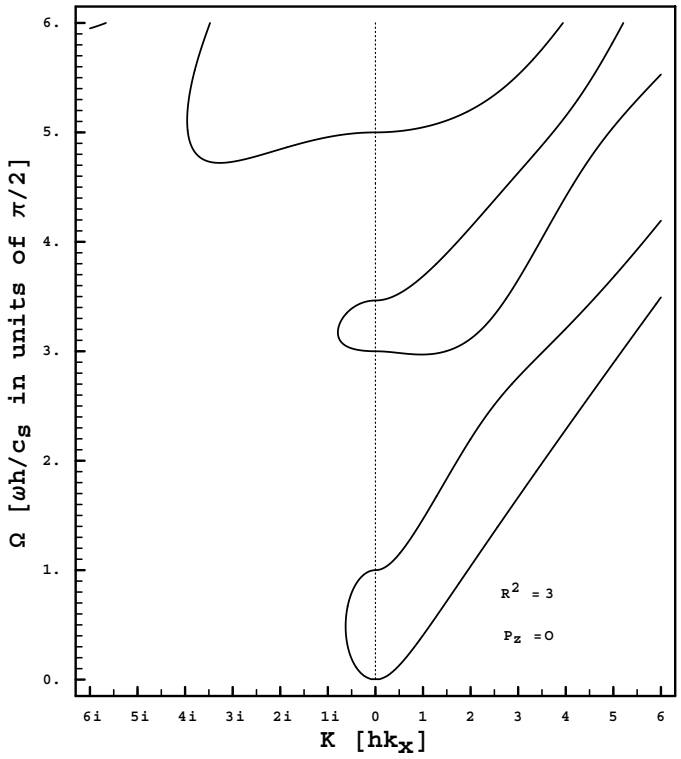

Figure 4.2: Rayleigh-Lamb Flexural Modes. In the historical development of the solution, these are referred to as the "antisymmetric" modes.

The Rayleigh-Lamb equation defines coupled modes of the infinite plate, but there are also a set of uncoupled modes that can propagate in this geometry. Recall that surface coupling is an interaction between longitudinal and that polarization of shear waves having displacement components perpendicular to the surface. I can set $\varphi = 0$ in the boundary equations (4.14) and contemplate plane waves that only have displacement parallel to the $z = \pm h$ surfaces. This can be realized simply by

setting $H_x = H_y = 0$. The boundary conditions then reduce simply to

$$\begin{aligned} 0 &= H_{z,zy} \\ 0 &= H_{z,zx} \qquad \text{at } z = \pm h \end{aligned} \tag{4.24}$$

Shear waves with polarization resulting in displacements solely parallel to encountered surfaces are often called SH waves (for "shear, with displacements horizontal to the surfaces") in contradistinction to SV waves (for "shear, with dispacements vertical to the surfaces") that are coupled to longitudinal waves at surfaces.

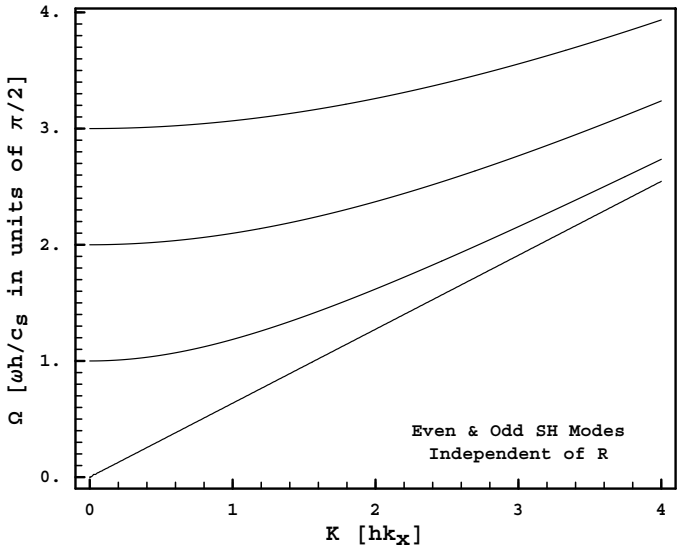

Figure 4.3: Infinite Plate SH Modes. Shear waves with displacements parallel to the surfaces, and which vanish there, form an uncoupled propagating system in an infinite plate. The lowest even and odd subbands are shown together.

Now, H_z, by my chosen basis representation, must have the form

$$H_z = B\,\bar{P}_x(k\,x)\bar{P}_y(\sigma y)\,P_z(\sigma^+ z)$$

Thus, it is easy to see that the nontrivial solutions for this SH system follow simply by setting the z-derivative of $P_z(\sigma^+ z)$ at $z = \pm h$ to zero.

$$\bar{P}_z(\sigma^+ h) = 0 \Rightarrow \left\{ \begin{matrix} \sin(\sqrt{\Omega^2 - K^2}) \\ \cos(\sqrt{\Omega^2 - K^2}) \end{matrix} \right\}_{P_z} = 0 \tag{4.25}$$

I will make comparison to these SH solutions in the sequel. Subbands of this solutions set are shown in Figure 4.3. They are analagous to torsional modes of a waveguide whose dominant displacement pattern are also characterized by displacements parallel to the surfaces.

4.5 How to Transform Superpositions

The re-derivation of the Rayleigh-Lamb equation was an exercise in organizing the problem into an algebraic form amenable to the simple elimination of unknown constants. The three independent boundary value equations constitute at most three constraints. Absent superpositions, the scaling constants for the potentials will constitute one degree of freedom for the scalar potential and as many additional degrees of freedom as there are distinct components of the vector potential to resolve. However, at each surface only one directional coordinate will be fixed and so there is possibly one additional degree of freedom to be resolved with respect to the other. The Rayleigh-Lamb scenario reduces enough degrees of freedom to balance the constraints. Specifically, restricting the scenario to plane waves eliminates the directional degree of freedom at each surface and reduces the number of constants from a maximum of four to a manageable two. With the nontrivial constraints reduced to the same number, a solution follows.

In the rectangular waveguide case, the directional degree of freedom at each surface persists by the fact that plane waves no longer suffice. Barring some fortuity,

the constraint equations will go to six–three for each surface–with no eliminations. A question arises: how many independent vector potential components must there be and thus how many degrees of freedom due to them? Two directional degrees of freedom plus the scalar plus at least one vector potential component makes the minimum degrees of freedom to be four. If I couple the adjacent sides, the independent constraints may be reduced and if I include additional vector potential components I can increase the degrees of freedom. But as I have demonstrated (section 2.5) abstractly, nonseparability compels the introduction of superpositions and a new problem arises of how to resolve their expansion coefficients. The goal of devising a transformation is to deal with this latter complication.

Once transforms are applied to a superposition substituted into a rectangular system, it will be seen that there are elementary Rayleigh-Lamb relationships between components of the coupled potentials. This result might be anticipated qualitatively by noticing that many dispersion subbands revealed by numerical approximations of the rectangular case bear uncanny resemblance to Rayleigh-Lamb dispersions of a plate. (This was noticed somewhat by Nishiguchi [1] §3).

Given that I have explicitly chosen a representation of the potentials as sums of products of sines and cosines (over only positive wave numbers), the boundary conditions are readily visualized as equalities among sums of exponentials. A Fourier transform is then naturally suggested. The domain of the transformed terms is, however, finite, and so it is necessary to either periodically extend the functions beyond the boundary or to modify the definition of the functions so that they vanish or converge towards zero beyond the boundary. A simple periodic extension would have the desirable consequence of producing delta functions under a Fourier transform corresponding to each term–but only if the wave numbers were real and the coordinate was being mapped to real transform variables–or

imaginary and being mapped to imaginary transform variables. However, it is to be expected that a given potential is represented by a sum containing both real and imaginary wave numbers with the result that a standard Fourier transform would diverge. The inescapability of this expectation comes from examining Rayleigh-Lamb solutions for which the fundamental modes distinctly involve transverse wave numbers which are pure imaginary. In addition, as anticipated in section 4.1, a periodic extension would, at least abstractly, render the assumptions supporting application of the Helmholtz Theorem invalid.

Preservation of the ability to transform into delta functions, despite divergence of the Fourier transform and other issues, can be secured by a simple expedient. Instead of the general Fourier transform, I use a "simplified" transform which is valid only on exponentials (though without regard to whether the wave numbers are real or imaginary) and which avoids the divergence problem by virtue of the details of its defined domain. Moreover, since the domain of functions to be transformed is explicitly constrained by the physical boundaries of the sample, there is nothing illogical or restrictive in defining the applicable domain for the transform to include only exponentials defined on the coordinate intervals that measure the sample and thus there need not be any concern over periodic extension. The transform to be used then has a simple definition determined by the element mapping

$$ae^{i\eta q} \mapsto 2a\delta(\gamma - \eta) \tag{4.26}$$

$$-h_q \leq q \leq h_q \ (q = x, y, z) \qquad \gamma, \eta \in \Re \cup \Im \qquad a \in \mathcal{C}$$

The factor of 2 is a convenience to dispose of the factor $\frac{1}{2}$ in the exponential representation of the sine and cosine.

It is almost self-evident that the transform is 1-1 between the set of exponentials and set of delta functions so restricted. Since a can be zero, an additive identity exists on each side and we have an isomorphism between two groups. Because the range of η and γ is defined to encompass both real and pure imaginary values, the transform operates without difficulty on any combination of exponentials with real or imaginary wave numbers.

It is troublesome to write down an integral form of this transform which smoothly adapts to whether the argument is real versus imaginary and which limits itself to the coordinate boundaries. Of course, the underlying mechanism is a trivial Fourier transform. Fortunately, since the domain of functions is strictly limited and the element mapping from that domain to the transform domain is clear and unambiguous, the transform can be performed without difficulty. It is interesting, therefore, that the inverse transform can be trivially written down in an integral form that is not troublesome and which applies adaptively to transforms involving imaginary delta arguments as well as real ones. One example would be

$$f(y) = \mathcal{T}^{-1}_{y\to\lambda}\{f(\lambda)\} = \frac{1}{2}\left[\int_{\lambda=-\infty}^{\lambda=+\infty} e^{i\lambda y} d\lambda + \int_{\lambda=-i\infty}^{\lambda=+i\infty} e^{i\lambda y} d\lambda\right]$$

Of course, this is only valid for $f(\lambda)$ that are produced by the "simplified" transform in the first place.

The left or right hand side of any boundary value constraint will involve one or more sums in the following general form (which omits common factors of $P_x(k\,x)$ which are also subject to differentiation):

$$\sum_j a_j\, f(\rho_j, \rho_j^\dagger)\, P_{q_1}(\rho_j\, q_1) P_{q_2}(\rho_j^\dagger\, q_2) \tag{4.27}$$

Here, I have generalized the various cases:

q_1, q_2 stand for distinct coordinates y or z;

P_1, P_2 are distinct function variables, or derivatives of them: P_y or P_z;

ρ will be an η or σ for representations of the scalar or component of the vector potential respectively;

$\rho^\dagger$ is an abbreviation for the conjugate wave number based upon the applicable velocity: $\rho^* = \sqrt{\omega^2/c_\ell^2 - k^2 - \rho^2}$ $\quad$ $\rho^+ = \sqrt{\omega^2/c_s^2 - k^2 - \rho^2}$;

$f(\rho_j, \rho_j^\dagger)$ will be a prefactor resulting from one or more derivatives taken. These will always be a single product or sums of products of ρ_j and/or $\rho_j^\dagger$. To characterize the effect of transforms on the boundary expressions, it will be sufficient to contemplate $f(\rho_j, \rho_j^\dagger)$ as being a single such product since a sum of such terms can be distributed to produce sums of summations.

By summarizing below how the "simplified" transform affects boundary value terms generically defined by equation (4.27), it will be possible to immediately write down the transforms of the actual boundary conditions without elaboration. First, with $y \to \lambda$ chosen to make the example concrete, consider the general effect of the transform on a function variable.

$$\mathcal{T}_y\{P_y(\eta\, y)\} = \begin{Bmatrix} 1 \\ -i \end{Bmatrix}_{P_y} \left[\delta(\lambda - \eta) + \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_y} \delta(\lambda + \eta) \right] \tag{4.28}$$

I can now write down the transform of equation (4.27) with respect to $q_1 \to \gamma$. Note that γ stands for whichever transform dimension variable is matched to q_1. My convention henceforth will be that $y \to \lambda$ and $z \to \mu$. In the rendering of transforms, q_1 and q_2 could be either y or z, though always distinct in a given case.

While the reader is presumed capable of writing down Fourier transforms of sines and cosines on his/her own, the usual results are somewhat simplified and adjusted in this case by virtue of the stipulation that superpositions will always be chosen to involve only positive (albeit possibly imaginary) values of whatever wave number variable ρ designates. The result, therefore, is that only one of the two delta terms survive in each case and the specific form of the result depends upon the sign of the transform variable in a way that can be neatly summarized. So, for convenience I list the results in detail:

$\mathcal{T}_{q_1\to\gamma}\left\{\sum_j a_j\, f(\rho_j,\rho_j^\dagger)\, P_{q_1}(\rho_j\, q_1)P_{q_2}(\rho_j^\dagger\, q_2)\right\} =$

$$\sum_j a_j\, f(\gamma,\gamma^\dagger) \begin{Bmatrix} 1 \\ -i \end{Bmatrix}_{P_{q_1}} \delta(\gamma-\rho_j)\, P_{q_2}(\gamma^\dagger\, q_2) \quad \text{for } \gamma>0 \tag{4.29}$$

$$\sum_j a_j\, f(0,0^\dagger) \begin{Bmatrix} 1 \\ -i \end{Bmatrix}_{P_{q_1}} \begin{Bmatrix} 2 \\ 0 \end{Bmatrix}_{P_{q_1}} \delta(\rho_j)\, P_{q_2}(0^\dagger\, q_2) \quad \text{for } \gamma=0 \tag{4.30}$$

$$\sum_j a_j\, f(|\gamma|,\gamma^\dagger) \begin{Bmatrix} 1 \\ -i \end{Bmatrix}_{P_{q_1}} \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_{q_1}} \delta(|\gamma|-\rho_j)\, P_{q_2}(\gamma^\dagger\, q_2) \quad \text{for } \gamma<0 \tag{4.31}$$

Similarly, I can now write down the transform of equation (4.27) with respect to q_2. There are again the same three cases depending upon the sign or γ.

$\mathcal{T}_{q_2\to\gamma}\left\{\sum_j a_j\, f(\rho_j,\rho_j^\dagger)\, P_{q_1}(\rho_j\, q_1)P_{q_2}(\rho_j^\dagger\, q_2)\right\} =$

$$\sum_j a_j\, f(\gamma^\dagger,\gamma) \begin{Bmatrix} 1 \\ -i \end{Bmatrix}_{P_{q_2}} \delta(\gamma-\rho_j^\dagger)\, P_{q_1}(\gamma^\dagger\, q_1) \quad \text{for } \gamma>0 \tag{4.32}$$

$$\sum_j a_j\, f(0^\dagger,0) \begin{Bmatrix} 1 \\ -i \end{Bmatrix}_{P_{q_2}} \begin{Bmatrix} 2 \\ 0 \end{Bmatrix}_{P_{q_2}} \delta(0-\rho_j^\dagger)\, P_{q_1}(0^\dagger\, q_1) \quad \text{for } \gamma=0 \tag{4.33}$$

$$\sum_j a_j \, f(\gamma^\dagger, |\gamma|) \begin{Bmatrix} 1 \\ -i \end{Bmatrix}_{P_{q_2}} \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_{q_2}} \delta(|\gamma| - \rho_j^\dagger) \, P_{q_1}(\gamma^\dagger \, q_1) \quad \text{for } \gamma < 0 \tag{4.34}$$

Having been meticulous in motivating, then justifying, and now demonstrating the effects of this "simplified" transform. It can be drastically simplified in practice with the following observations:

The prior stipulation that expansions will only be over positive, though possibly including imaginary, wave numbers has allowed the results of transforms in the context of the problem to be more easily summarized in terms of single delta functions. One additional fact can now literally trivialize the use of this transform in practice. Namely, the fact that any given derivation takes place in the context of a specific parity pattern guarantees that within any derivation the parity pattern on each side of a boundary equation will be identical. This parity agreement guarantees that when either coordinate is transformed, the P function transformed on each side will be the same. That being the case, all of the prefactors which depend on the parity of P will cancel between the sides in all cases where that P function is common to all additive terms. In general, this condition nearly always fulfilled. Moreover, with those distinctions gone, an examination of equations (4.29-4.34) will reveal that replacing γ with $|\gamma|$ throughout is sufficient to cover all cases. Therfore, as a practical matter, the only rule that will be needed is just

> Replace every $P_q(\rho_q \, q)$ to be transformed with $\delta(|\gamma| - \rho_q)$.
>
> It will not matter whether ρ_q is a wave number or its conjugate.
>
> It will not matter whether ρ_q is real or imaginary (it is guaranteed to be positive).

The triviality of this rule is a direct consequence of having imposed a meticulous series of specific choices. It is neither fortuitous nor could it have been readily anticipated that it would reduce to this.

Once a boundary equation that has had a superposition substituted into it becomes transformed in this way, it becomes an equality in terms of the transform variable. Given the behavior of delta functions, each value of the transform variable will, on each side of the summation, select out either a specific term of the sum, or be identically zero. This collapses the equality of functions of sums into a constraint between components of sums from each side. Correlating transforms over y with those over z is a remaining issue, but how this must be done will be developed in the derivations which follow.

CHAPTER 5
DERIVING NORMAL MODES OF PROPAGATION

5.1 General Considerations

The full boundary conditions revealed generically by equations (2.22) infer that all three components of the shear vector potential are mixed together in satisfying the stress-free surface boundary constraint. If, however, I restrict attention to just those modes which propagate, it is not immediately clear whether all three vector components are needed. Some inspiration can be drawn from the Rayleigh-Lamb derivation of section 4.4. That derivation involves an infinite plate bounded in the z directions and only requires the H_y component of the vector potential. The intimation is that bounding also in the y directions might simply invoke the need for the H_z component, but there is no a priori reason to expect to need an H_x component as well.

The foregoing motivates an attempt to find the essential relationship between H_x and the other vector components which participate in satisfying the boundary conditions along the surface of a waveguide in which normal modes propagate in the x direction. Unless some constraint can be found that eliminates components or establishes some dependency among them, there will be more degrees of freedom in the propagating problem then constraints available to resolve them. I thus proceed to investigate this relationship among components in a way that is independent of the boundary conditions per se so as to confidently narrow the approaches used in solving the boundary problems.

5.2 Acoustic Poynting Vector of a Normal Mode

Propagating modes carry energy. In analogy with electrodynamics, there will be a vector that indicates both the direction and magnitude of the energy flux. This vector, the acoustic Poynting vector, describes a physical reality whose invariances under manipulation of the coordinate system constrain its mathematical form. The components of **H** participate in the construction of a mode via antisymmetries arising from the fact that shear displacement is $\nabla \times \mathbf{H}$. Moreover, satisfying the boundary conditions ostensibly involves mixing up the components of **H**. This suggests I look for constraints on the relationship between components of **H** that might be required to maintain invariances of the Poynting vector while preserving the antisymmetries built into the shear contributions to the mode. Some familiarity with the acoustic Poynting Vector reveals it to be a complicated object when expressed in terms of strain and this provokes a curiosity over whether it may harbor such constraints. It is difficult to articulate further what is, in the end, an intuition that this might be so. The intuition will ultimately be justified by the result.

A representation of the acoustic Poynting Vector itself can be readily derived. The Poynting Vector will be denoted herein as **J** since P's have been extensively used for another purpose. It its defined by the property that integrating it over a surface S produces the energy flux through that surface.

$$E_{,t} = \int_S J_i n_i ds \tag{5.1}$$

where **n** is the surface normal vector. Components of the force density at a point on the surface are

$$F_j = -\sigma_{ij} n_i \tag{5.2}$$

These are the internal forces which are a response to strain. Though inconsequential in what follows, the minus sign which appears above propagates to one in the expression for **J** where it is likely to seem counterintuitive.

Force density times displacement is just work density and thus the measure of energy transport per unit volume. Taking the force density as constant over infinitesimal displacements, the time rate of change of $F_j u_j$ is just $F_j u_{j,t}$. The result is that power density is simply the scalar product of force density and velocity–a result familiar from elementary mechanics. Applying this to equations (5.1) and (5.2), the expression for flux and Poynting vector components it infers are

$$E_{,t} = \int_S (-\sigma_{ij} n_j) u_{i,t}\, dS \quad \Rightarrow \quad J_i = -\sigma_{ij}\, u_{j,t} \tag{5.3}$$

I shall be concerned only with invariance of directionality and symmetry properties and, since only normal modes are relevant, I will dispense with the time derivative and ignore the minus sign.

For an isotropic material, dividing equation (2.17) by material constants c_s and ρ expresses the proportionality of stress to strain independent of the material.

$$\sigma_{ij} \sim -\beta \delta_{ij} \sum_r u_{rr} \; + 2u_{ij} \tag{5.4}$$

where $\beta = R^2 - 2$. From the discussions following equation (2.18), I can replace the invariant sum with $\nabla^2 \varphi = -k_\ell^2 \varphi$. σ_{ij} is symmetric in its indices. Substituting this result into the representation for **J** in equations (5.3), I can express the proportionality of the time-independent part of J_i to a function of the undifferentiated scalar potential, total strain, and displacement.

$$J_i \sim \beta\, k_\ell^2\, \varphi\, u_i - 2u_j\, u_{ij} \tag{5.5}$$

u_{ij} separates into shear and longitudinal parts and the longitudinal part is immediately expressible in terms of the scalar potential using

$$u_{ij}^{(\ell)} = \varphi_{,ij} \tag{5.6}$$

whereas applying definition (2.2) to $u_{ij}^{(s)}$ and replacing the displacements thereof by curls of the shear vector potential requires the more involved substitution

$$u_{ij}^{(s)} = \tfrac{1}{2}(\epsilon_{i\alpha\beta}H_{\beta,\alpha j} + \epsilon_{j\alpha\beta}H_{\beta,\alpha i}) \tag{5.7}$$

Correspondingly, components of the remaining displacement term can be expanded in terms of potentials by

$$u_j = \epsilon_{jab}H_{b,a} + \varphi_{,j} \tag{5.8}$$

With the foregoing substitutions, the $u_j u_{ij}$ on the right side of equation (5.5) expands to

$$\begin{aligned} u_j u_{ij} = \ & \tfrac{1}{2}(\epsilon_{jab}\epsilon_{i\alpha\beta}H_{b,a}H_{\beta,\alpha j} + \epsilon_{jab}\epsilon_{j\alpha\beta}H_{b,a}H_{\beta,\alpha i}) \\ & + \tfrac{1}{2}(\epsilon_{iab}H_{b,aj} + \epsilon_{jab}H_{b,ai})\varphi_{,j} \\ & + \epsilon_{jab}H_{b,a}\varphi_{,ij} + \varphi_{,j}\varphi_{,ij} \end{aligned} \tag{5.9}$$

After a tedious amount of tensor algebra, the first term on the right of equation (5.9) reduces to

$$2H_{[a,i]}H_{[b,a],b} + 2H_{[a,b]}H_{a,bi} \tag{5.10}$$

The remaining terms do not simplify in useful ways at this level of expression. Collecting the foregoing results, equation (5.5) can now be fully expressed in terms

of the scalar potential and components of the vector potential by

$$\begin{aligned} J_i \;\sim\; & (\beta k_l^2 \varphi)(\epsilon_{iab} H_{b,a} + \varphi_{,i}) \\ & -4(H_{[a,i]} H_{[b,a],b} + H_{[a,b]} H_{a,bi}) \\ & -(\epsilon_{iab} H_{b,aj} + \epsilon_{jab} H_{b,ai})\varphi_{,j} \\ & -2\epsilon_{jab} H_{b,a} \varphi_{,ij} - 2\varphi_{,j}\varphi_{,ij} \end{aligned} \tag{5.11}$$

The actual physical direction of energy propagation should be invariant under exchange of the transverse coordinates and so, using equation (5.11), I show that in order for this to be true, the y and z components of $\mathbf{H}$ cannot be allowed to mix with the x component.

Specifically, since the Poynting vector is composed, in part, of curls of $\mathbf{H}$, any exchange of y and z will interact with the handedness of the coordinate system to require an appropriate antisymmetry. To be consistent with this, the wave number in the x direction must also be taken to change sign with any $y \leftrightarrow z$ interchange and so all derivatives with respect to x will also be required to change sign. In view of the translational invariance of the modes, this is simply a reflection of the need to change k, the common wave number in the x direction, to $-k$ to accommodate inversion of the x direction.

The full expansion of the Poynting vector reveals a complex mixing of x versus y, z components of $\mathbf{H}$, but I shall show that failure to separate solutions to avoid this mixing results in failures in the $y \leftrightarrow z$ symmetries of some of the terms and therefore implies that solutions must be built distinctly from cases in which $H_x = 0$ versus those in which, alternatively $H_y = H_z = 0$. It should also be noted that this will turn out to be consistent with an analysis of the $k = 0$ case to follow. In the $k = 0$ case all derivatives with respect to x vanish and the boundary conditions

naturally take a form which reflects a decoupling of the x versus y, z components of $\mathbf{H}$.

The first additive term on the right hand side of equation (5.11) contributes a term to the x component of the Poynting vector that is proportionate to

$$\epsilon_{xab} H_{b,a} + \varphi_{,x} = 2H_{[z,y]} + \varphi_{,x} \tag{5.12}$$

which is fully antisymmetric under $y \leftrightarrow z$ once we incorporate the rule that $\varphi_{,x} \rightarrow -\varphi_{,x}$. Moreover, this antisymmetry is preserved even if either one of the terms vanishes.

The second term on the right hand side of equation (5.11) contributes a term to the x component of the Poynting vector equal to

$$\begin{aligned} -4 \Big\{ \Big[& H_{[x,y]} (H_{[x,y],x} + H_{[y,z],z}) \\ & + H_{[x,z]} (H_{[x,z],x} + H_{[z,y],y}) \Big] \\ & + 2 H_{[y,z]} H_{[y,z],x} \Big\} \end{aligned} \tag{5.13}$$

and before sign changes due to differentiation by x, this expression is totally symmetric under $y \leftrightarrow z$. Expansion of the antisymmetric parts entails production of the pair

$$\ldots - H_{y,x} H_{x,yx} \ldots - H_{z,x} H_{x,zx} \ldots \tag{5.14}$$

and thus does not uniformally assemble odd numbers of x differentiations with symmetric yz terms and so the needed antisymmetry is not fully realized unless some terms vanish. If H_x is set to zero, then it is easily checked that antisymmetry will be realized, to wit:

$$- \Big[H_{y,x} H_{y,xx} - H_{y,x} H_{[y,z],z} + H_{z,x} H_{z,xx} - H_{z,x} H_{[z,y],y} + 8 H_{[y,z]} H_{[y,z],x} \Big] \tag{5.15}$$

Similarly, if H_x does not vanish, but H_y and H_z vanish simultaneously, then, again, the resulting expansion will become antisymmetric under $y \leftrightarrow z$ once the x-derivatives are considered, to wit:

$$-\left[H_{x,y}H_{x,yx} + H_{x,z}H_{x,zx}\right] \tag{5.16}$$

The third term on the right of equation (5.11) contributes the following term to the x component of the Poynting vector:

$$\begin{aligned} -2\Big[&H_{[z,y],x}\varphi_{,x} + H_{[z,y],x}\varphi_{,y} + H_{[z,y],z}\varphi_{,z} \\ &+H_{[z,y],x}\varphi_{,x} + \left(H_{[x,z],x}\varphi_{,y} + H_{[x,y],x}\varphi_{,z}\right)\Big] \end{aligned} \tag{5.17}$$

Under $y \leftrightarrow z$ this is totally antisymmetric when the derivatives of x are considered. Moreover, if H_x vanishes or, alternatively H_y and H_z vanish together, the antisymmetry of the result is preserved.

Finally, the last two terms on the right of equation (5.11) contribute the following terms to the x component of the Poynting vector:

$$\begin{aligned} -4\Big[&H_{[z,y]}\varphi_{,xx} + \left(H_{[x,z]}\varphi_{,xy} - H_{[x,y]}\varphi_{,xz}\right)\Big] \\ &-2\left[\varphi_{,x}\varphi_{,xx} + \left(\varphi_{,y}\varphi_{,xy} + \varphi_{,z}\varphi_{,xz}\right)\right] \end{aligned} \tag{5.18}$$

It is easily checked that the desired antisymmetry is preserved and that, again, the vanishing either of H_x alone or H_y and H_z together does not change this result.

The conclusion is that, when assembling a propagating normal mode, the shear contribution must be made out of components for which all the H_x's are zero, or for which all the H_y and H_z parts are zero. In considering how to represent the shear superpositions of a propagating normal mode, there is no case in which a

superposition for H_x will be mixed with ones for H_y and H_z. This removes at least one degree of freedom from the problem.

5.3 Propagating Modes Involving H_y, H_z Shear

5.3.1 Deriving the Frequency Equations

The ease with which the Rayleigh-Lamb solution is derived inspires a derivation that follows the same pattern. Armed with the conclusion that H_x cannot even be accommodated in a normal mode solution that also includes H_y and H_z components, I proceed to derive the spectrum of propagating modes with $H_x = 0$.

Accordingly, from equations (2.22), the boundary conditions at $z = h_z$ with $s \rightarrow z$, $p \rightarrow x$, $\bar{p} \rightarrow y$, and $H_x = 0$ become

$$\begin{aligned} -\beta\, k_\ell^2\, \varphi + 2\varphi_{,zz} &= 2H_{y,xz} \\ \varphi_{,zx} &= -\tfrac{1}{2}\left[(H_{y,xx} - H_{y,zz}) + H_{z,zy}\right] \\ \varphi_{,zy} &= +\tfrac{1}{2}(H_{z,zx} - H_{y,yx}) \end{aligned} \tag{5.19}$$

As a reminder, under the basis rules devised for this problem (see section 4.2), the representations of potentials will have the following forms:

$$\varphi = P_x(k\,x) \sum_i d_i\, P_y(\eta_i\, y)\, P_z(\eta_i^*\, z) \qquad \text{with } \eta_i^* = \sqrt{\omega^2/c_\ell^2 - k^2 - \eta_i^2}$$

$$\mathbf{H} = \bar{P}_x(k\,x) \begin{bmatrix} 0 \\ \sum_j a_j P_y(\sigma_j\, y) \bar{P}_z(\sigma_j^+\, z) \\ \sum_j b_j \bar{P}_y(\sigma_j\, y) P_z(\sigma_j^+\, z) \end{bmatrix} \qquad \text{with } \sigma_j^+ = \sqrt{\omega^2/c_s^2 - k^2 - \sigma_j^2}$$

Substituting into the first boundary condition of equations (5.19) I obtain

$$\sum_i d_i \,(\beta k_\ell^2 + 2(\eta_i^*)^2) P_y(\eta_i y)\, P_z(\eta_i^* \, h_z) = \\ 2k \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_x} \sum_j \sigma_j^+ \, P_y(\sigma_j \, y) \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_z} P_z(\sigma_j^+ \, h_z) \tag{5.20}$$

The "simple" transform devised in section 4.5 is then applied so that $P_y(\eta_i y) \rightarrow \delta(|\lambda| - \eta_i)$ and $P_y(\sigma_j \, y) \rightarrow \delta(|\lambda| - \sigma_j)$. By choosing a value λ_o of the transform variable λ such that $\lambda_o \in \{\eta_i\} \cap \{\sigma_j\}$, the sums on both sides collapse leaving the following equality:

$$d_o \,(\beta k_\ell^2 + 2(\lambda_o^*)^2)\, P_z(\lambda_o^* \, h_z) = 2k \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_x} \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_z} a_o \, \lambda_o^+ \, P_z(\lambda_o^+ \, h_z) \tag{5.21}$$

This provides one constraint on possible combinations of (ω, λ_o) at a given value of k. Here and in subsequent steps, $\lambda_o \equiv |\lambda_o|$.

Repeating this process for the second boundary condition in equations (5.19), the transformed version of the second constraint becomes

$$d_o \, k \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}_{P_x} \lambda_o^* \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}_{P_z} \bar{P}_z(\lambda_o^* \, h_z) = \\ \frac{1}{2} \bar{P}_z(\lambda_o^+ \, h_z) \left[a_o(k^2 - (\lambda_o^+)^2) - b_o \, \lambda_o(\lambda_o^+) \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_y} \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}_{P_z} \right] \tag{5.22}$$

Substituting into the third boundary condition in equations (5.19), the transformed result is

$$d_o\,\lambda_o\lambda_o^*\begin{Bmatrix}-1\\1\end{Bmatrix}_{P_y}\begin{Bmatrix}-1\\1\end{Bmatrix}_{P_z}\bar{P}_z(\lambda_o^*\,h_z)=$$
$$\frac{1}{2}k\begin{Bmatrix}1\\-1\end{Bmatrix}_{P_x}\left[b_o\,\lambda_o^+\begin{Bmatrix}-1\\1\end{Bmatrix}_{P_z}-a_o\,\lambda_o\begin{Bmatrix}-1\\1\end{Bmatrix}_{P_y}\right]\bar{P}_z(\lambda_o^+\,h_z) \tag{5.23}$$

Equations (5.22) and (5.23) can be reconciled into one constraint by finding a relationship between a_o and b_o that renders them equivalent. Dividing the two equations will eliminate the transcendental terms and some common factors, leaving

$$\frac{-k^2}{\lambda_o}=\frac{a_o\begin{Bmatrix}-1\\1\end{Bmatrix}_{P_y}(k^2-(\lambda_o^+)^2)+b_o\begin{Bmatrix}-1\\1\end{Bmatrix}_{P_z}\lambda_o\lambda_o^+}{b_o\begin{Bmatrix}-1\\1\end{Bmatrix}_{P_z}\lambda_o^+-a_o\begin{Bmatrix}-1\\1\end{Bmatrix}_{P_y}\lambda_o} \tag{5.24}$$

Solving this for the required relationship between coefficients yields

$$b_o\begin{Bmatrix}-1\\1\end{Bmatrix}_{P_z}=a_o\begin{Bmatrix}-1\\1\end{Bmatrix}_{P_y}\left[\frac{\lambda_o\lambda_o^+}{k^2+\lambda_o^2}\right] \tag{5.25}$$

Substitution of equation (5.25) into either equation (5.22) or (5.23) to eliminate b_o will produce the following result:

$$d_o\begin{Bmatrix}-1\\1\end{Bmatrix}_{P_x}\lambda_o^*\,\bar{P}_z(\lambda_o^*\,h_z)=\frac{1}{2}a_o\,k\left[\frac{(k^2+\lambda_o^2)-(\lambda_o^+)^2}{k^2+\lambda_o^2}\right]\bar{P}_z(\lambda_o^+\,h_z) \tag{5.26}$$

Having reduced the system to two equations with two unknown coefficients, those coefficients can be eliminated by dividing equation (5.26) into equation (5.21). First, however, inspired by a similar step in the Rayleigh-Lamb derivation, the parenthetical on the left side of equation (5.21) can be manipulated and found to have a convenient equivalent representation which coincides with a subexpression in equation (5.26).

$$(\beta k_\ell^2 + 2(\lambda_o^*)^2) = -((k^2 + \lambda_o^2) - (\lambda_o^+)^2)$$

Substituting this and then performing the division followed by the usual rearrangement produces the following frequency equation:

$$\frac{P_z(\lambda_o^* h_z)}{\bar{P}_z(\lambda_o^* h_z)} \frac{\bar{P}_z(\lambda_o^+ h_z)}{P_z(\lambda_o^+ h_z)} = \frac{-4\,\lambda_o^*\,\lambda_o^+\,(k^2 + \lambda_o^2)}{[(k^2 + \lambda_o^2) - (\lambda_o^+)^2]^2} \tag{5.27}$$

Comparing with equation (4.20) this is seen to be a Rayleigh-Lamb equation with the quantity $\sqrt{k^2 + \lambda_o^2}$ playing the role of the magnitude of the propagation vector. This equation is parameterized by λ_o and is valid, at a given k, only for certain combinations of λ_o and ω. Consider, nevertheless, what happens when $k \gg \lambda_o$ (or, effectively $\lambda_o \to 0$ for finite k).

$$\begin{aligned} \lambda_o^* = \sqrt{\omega^2/c_\ell^2 - k^2 - \lambda_o^2} \quad &\to \quad k^* = \sqrt{\tfrac{\omega^2}{c_\ell^2} - k^2} = \eta \quad \text{(in equation (4.21))} \\ \lambda_o^+ = \sqrt{\omega^2/c_s^2 - k^2 - \lambda_o^2} \quad &\to \quad k^+ = \sqrt{\tfrac{\omega^2}{c_s^2} - k^2} = \sigma \quad \text{(in equation (4.21))} \end{aligned} \tag{5.28}$$

Therefore, at large k, it is expected that equation (5.27) becomes more like the classic Rayleigh-Lamb equation (4.21) for plane waves in an infinite plate. At this intermediate point, the constraint of boundary conditions at the $y = h_y$ surface

has not yet been imposed, but the implication nevertheless is that for large k the system looks substantially like a Rayleigh-Lamb system. This implies that as $k \gg 0$ the effect from the sides tend to decouple and the propagation looks increasingly like simple decoupled Rayleigh-Lamb plane-wave propagation.

It now becomes important to focus on a simple observation. The foregoing step produced an equation in terms of the transform variable λ whose permissible values (denoted λ_o) come from the set of values in $\{\eta_i\}$ and $\{\sigma_j\}$ (what mathematicians would call the "support" for φ and H_y, H_z). There could be more elements of those support sets than possible values of λ_o and these additional values may be uniquely connected to the transform variable μ applied to boundary conditions at the adjacent surface. However, connecting the $y = h_y$ and $z = h_z$ surfaces will be possible only to the extent that some of the values which the two transform variables take on are indeed shared selections from $\{\eta_i\}$ and $\{\sigma_j\}$.

From equations (2.22), the boundary conditions at $y = h_y$ with $s \to y$, $p \to z$, $\bar{p} \to x$, and $H_x = 0$ will be

$$\begin{aligned} -\beta\, k_\ell^2\, \varphi + 2\varphi_{,yy} &= 2H_{z,xy} \\ \varphi_{,yx} &= \tfrac{1}{2}\left[(H_{z,xx} - H_{z,yy}) + H_{y,yz}\right] \\ \varphi_{,zy} &= -\tfrac{1}{2}(H_{y,yx} - H_{z,zx}) \end{aligned} \tag{5.29}$$

The transformed first boundary condition from equations (5.29), before application of the delta functions, will be

$$\begin{aligned} -\sum_i d_i\,(\beta k_\ell^2 + 2\eta_i^2) P_y(\eta_i\, a h_z)\,\delta(\eta_i^* - \mu) = \\ 2k \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_x} \sum_j \sigma_j \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_y} P_y(\sigma_j\, a h_z)\,\delta(\sigma_j^+ - \mu) \end{aligned} \tag{5.30}$$

To link the adjacent surface conditions, I rely upon the fact that each root of equation (5.27) necessarily corresponds to the existence of specific elements of the support sets $\{\eta_i\}$ and $\{\sigma_j\}$. In fact, if λ_o is, with some value of ω, a root of equation (5.27), it is solely because $\exists\eta_o \in \{\eta_i\}$ and $\exists\sigma_o \in \{\sigma_j\}$ such that $\lambda_o = \eta_o = \sigma_o$. Now, the left sum in equation (5.30) a will necessarily encounter $\lambda_o = \eta_o$.

Suppose, then, that I contemplate the value of μ corresponding to $\eta_o = \lambda_o$ for which the delta function on the left of equation (5.30) is nonzero. Obviously, it will be λ_o^*. If, for that value of μ, the right hand side of equations (5.30) is not trivially zero for the same value of μ, there must exist some σ_1 such that $\sigma_1^+ = \mu = \lambda_o^*$.

In the alternative, I could first contemplate a value of μ for which the argument of the delta function on the right of equation (5.30) is zero. Obviously, it will be λ_o^+. If, for that value of μ, the left hand side of equations (5.30) is not trivially zero for the same value of μ, there must exist some η_1 such that $\eta_1^* = \mu = \lambda_o^+$.

The upshot of this reasoning is that I can connect the transforms of the two sets of boundary conditions by contemplating simultaneous (ω, λ_o) roots of both of them. To write the transformed boundary conditions for the adjacent side in terms of λ_o, I require either that

$$\mu \rightarrow \lambda_o^*$$

so that, by operation the delta function, the wave number variables take on values:

$$\begin{aligned}
\eta_i &\rightarrow \eta_o = \lambda_o \\
d_i &\rightarrow d_o \quad \text{which will be eliminated} \\
\eta_i^* &\rightarrow \lambda_o^* \\
\sigma_j &\rightarrow \sigma_1 \quad \text{i.e. presumed to exist} \\
a_i &\rightarrow a_1 \quad \text{unknown, but to be eliminated}
\end{aligned}$$

$$\begin{aligned} \sigma_j^+ &\rightarrow \lambda_o^* \\ \sigma_j &= (\lambda_o^*)^+ \end{aligned} \tag{5.31}$$

or I require that

$$\mu \rightarrow \lambda_o^+$$

so that, by operation the delta function, the wave number variables alternatively take on values:

$$\begin{aligned} \sigma_j &\rightarrow \sigma_o = \lambda_o \\ b_j &\rightarrow b_o \text{ which will be eliminated} \\ \sigma_j^* &\rightarrow \lambda_o^+ \\ \eta_i &\rightarrow \eta_1 \text{ i.e. presumed to exist} \\ a_i &\rightarrow a_1 \text{ unknown, but to be eliminated} \\ \eta_i^* &\rightarrow \lambda_o^+ \\ \eta_i &= (\lambda_o^+)^* \end{aligned} \tag{5.32}$$

I will name the first of these alternatives (i.e., equations (5.31)) "L-Conjugation" since it is premised on equating the longitudinal conjugation of the z-surface solution with the y-surface solution. The second alternative (i.e., equations (5.32)) "S-Conjugation" since it is premised on equating the shear conjugation of the z-surface solution with the y-surface solution.

Because I have stipulated that all members of the support sets are positive and because only positive square roots are used, all of the preceding mappings are guaranteed to be unambiguous. The reader is invited to review the definitions of conjugation denoted with * and + superscripts which were introduced in connection with equations (4.8) and (4.10). From those definitions, it can be noted that

these conjugations have the property

$$(\lambda^*)^* = \lambda \qquad (\lambda^+)^+ = \lambda$$

A detailed expansion of the final relation in equations (5.31) is

$$(\lambda_o^*)^+ = \sqrt{\frac{\omega^2}{c_s^2} - (\lambda_o^*)^2} = \sqrt{(\lambda_o^+)^2 + \lambda_o^2 - (\lambda_o^*)^2} \tag{5.33}$$

and of the final relation in equations (5.32) is

$$(\lambda_o^+)^* = \sqrt{\frac{\omega^2}{(R^2 c_s^2)} - (\lambda_o^+)^2} = \sqrt{(\lambda_o^*)^2 + \lambda_o^2 - (\lambda_o^+)^2} \tag{5.34}$$

These expansions illustrate the general rule that L-Conjugation and S-Conjugation solutions are related by straightforward substitutions of variables. It will thus be sufficient to complete details of the ongoing derivation for the L-Conjugate case and then state the analogous results for the S-Conjugate case. Applying L-Conjugation to equation (5.30), and after collapsing the sums, the result is

$$-d_o\,(\beta k_\ell^2 + 2\lambda_o^2)\,P_y(\lambda_o\, a h_z) = 2k \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_x} \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_y} b_1\,(\lambda_o^*)^+\,P_y[(\lambda_o^*)^+\, a h_z] \tag{5.35}$$

Similarly, the second and third boundary conditions (5.29) under L-Conjugation can eventually be put into the following forms:

$$d_o\, k \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}_{P_x} \lambda_o \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}_{P_y} \bar{P}_y(\lambda_o\, a h_z) =$$

$$\tfrac{1}{2}\bar{P}_y[(\lambda_o^*)^+\, a h_z] \left[-b_1[k^2 - ((\lambda_o^*)^+)^2] + a_1\,\lambda_o^*(\lambda_o^*)^+ \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}_{P_y} \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_z} \right] \tag{5.36}$$

$$d_o\,\lambda_o\lambda_o^*\begin{Bmatrix}-1\\ \\ 1\end{Bmatrix}_{P_y}\begin{Bmatrix}-1\\ \\ 1\end{Bmatrix}_{P_z}\bar{P}_y(\lambda_o\,ah_z)=$$

$$-\frac{1}{2}k\begin{Bmatrix}1\\ \\ -1\end{Bmatrix}_{P_x}\left[a_1\,(\lambda_o^*)^+\begin{Bmatrix}-1\\ \\ 1\end{Bmatrix}_{P_y}-b_1\,\lambda_o^*\begin{Bmatrix}-1\\ \\ 1\end{Bmatrix}_{P_z}\right]\bar{P}_y[(\lambda_o^*)^+\,ah_z] \qquad (5.37)$$

Dividing equations (5.36) and (5.37) to put one a_1 in terms of b_1, I obtain

$$a_1\begin{Bmatrix}-1\\ \\ 1\end{Bmatrix}_{P_y}=b_1\begin{Bmatrix}-1\\ \\ 1\end{Bmatrix}_{P_z}\left[\frac{(\lambda_o^*)^+\lambda_o^*}{k^2+((\lambda_o^*)^+)^2}\right] \qquad (5.38)$$

It may be noted that equation (5.38) is not identical to (5.25). This highlights the fact that a_1 and b_1 are expected to be distinct from a_o and b_o. If equation (5.38) is substituted into either equation (5.36) or equation (5.36) to eliminate a_1, the result is identical, to wit:

$$d_o\begin{Bmatrix}-1\\ \\ 1\end{Bmatrix}_{P_x}\lambda_o\,\bar{P}_y(\lambda_o\,ah_z)=-\frac{1}{2}b_1\,k\left[\frac{(k^2+(\lambda_o^*)^2)-((\lambda_o^*)^+)^2}{k^2+(\lambda_o^*)^2}\right]\bar{P}_y[(\lambda_o^*)^+\,ah_z] \qquad (5.39)$$

In what is by now a ritual, the parenthetical on the left of equation (5.35) can be manipulated to show its relation to the numerator in the fraction on the right hand side of (5.39).

$$(\beta k_\ell^2+2\lambda_o^2)=-[(k^2+(\lambda_o^*)^2)-((\lambda_o^*)^+)^2] \qquad (5.40)$$

Having reduced the unknown coefficients to only d_o and b_1, equation (5.39) can be divided into equation (5.35) to eliminate them and produce the frequency equation

associated with the $y = h_y$ side.

$$\frac{P_y(\lambda_o\, a h_z)}{\bar{P}_y(\lambda_o\, a h_z)} \frac{\bar{P}_y[(\lambda_o^*)^+\, a h_z]}{P_y[(\lambda_o^*)^+\, a h_z]} = \frac{-4\,\lambda_o\,(\lambda_o^*)^+\,(k^2 + (\lambda_o^*)^2)}{[(k^2 + (\lambda_o^*)^2) - ((\lambda_o^*)^+)^2]^2} \tag{5.41}$$

Once again, this is a Rayleigh-Lamb equation. Here, the quantity $(k^2 + (\lambda_o^*)^2)$ plays the role of the square of the magnitude of propagation.

Equation (5.41) is manifested in the classic Rayleigh-Lamb structure which shows the consistency of the result with preceding derivations. However, in this case, that consistent manifestation masks an interesting feature. Namely, all k dependency in equation (5.41) cancels internally. To see this, expand the terms which ostensibly appear to show a k dependency and observe that k is eliminated.

$$(\lambda_o^*)^+ = \sqrt{\frac{\omega^2}{c_s^2} - \left[\frac{\omega^2}{c_\ell^2} - k^2 - \lambda_o^2\right]} = \sqrt{\frac{\omega^2}{c_s^2}\left(1 - \frac{1}{R^2}\right) + \lambda_o^2}$$

$$(k^2 + (\lambda_o^*)^2) = k^2 + \left[\frac{\omega^2}{c_\ell^2} - k^2 - \lambda_o^2\right] = \frac{\omega^2}{R^2 c_s^2} - \lambda_o^2$$

Equations (5.27) and (5.41) define (ω, λ_o) root systems for independent equations. The coincidences of these simultaneous constraints define values of ω that constitute the eigenspectrum with respect to the L-Conjugation case. There is a set of common roots for each distinct value of k. Plotting the k versus ω dispersions requires methodical plotting at different values of k, but, in principle, the spectrum (for L-Conjugation) has been analytically and precisely specified.

The entire derivation is also invariant under $y \leftrightarrow z$ exchange–though this might not be immediately apparent. The key to realizing that it must be so is to realize that relabeling the directions must also be applied to the representations of the potentials as superpositions and that with directions relabeled the aspect ratio is

defined by $h_z = ah_y$. The reader can easily verify that the results of derivation will be equivalent to the foregoing.

Repeating the same sequence of steps from equation (5.30) to the present point, using S-Conjugation defined by equations (5.32), the S-Conjugate analog to equation (5.41) is

$$\frac{P_y((\lambda_o^+)^*\, ah_z)}{\bar{P}_y((\lambda_o^+)^*\, ah_z)} \frac{\bar{P}_y[\lambda_o\, ah_z]}{P_y[\lambda_o\, ah_z]} = \frac{-4\,\lambda_o\,(\lambda_o^+)^*\,(k^2 + (\lambda_o^+)^2)}{[(k^2 + (\lambda_o^+)^2) - \lambda_o^2]^2} \tag{5.42}$$

5.3.2 Interpretation

Had anyone been insightful enough to anticipate that coupled modes of a rectangular waveguide could be characterized by a coincidence of Rayleigh-Lamb solutions, proving that it was so would have remained as elusive as history shows the main problem to have been. Moreover, there are aspects of the result which, had it been somehow forseen as possible, would have argued against believing it. The main impediment would have been that there is an intrinsic interference built into the result which precludes boundary satisfaction at both surfaces without the rest of the superposition. Although the elegant-looking result involves a coincidence with the full superposition–it is still not the full superposition. What is intriguing is that I do not need to have a full description of the superposition in order to find the eigenspectrum.

The derivation tells us that any superposition of shear and longitudinal components that satisfy boundary conditions at adjacent sides must include a particular combination of components that, in a partial sense, mimic a Rayleigh-Lamb wave system. The L-Conjugation and S-Conjugation cases are merely two different ways of realizing this. Figures 5.1 and 5.2 illustrate the essence of these alternatives. They imply two recipes for building the superpositions.

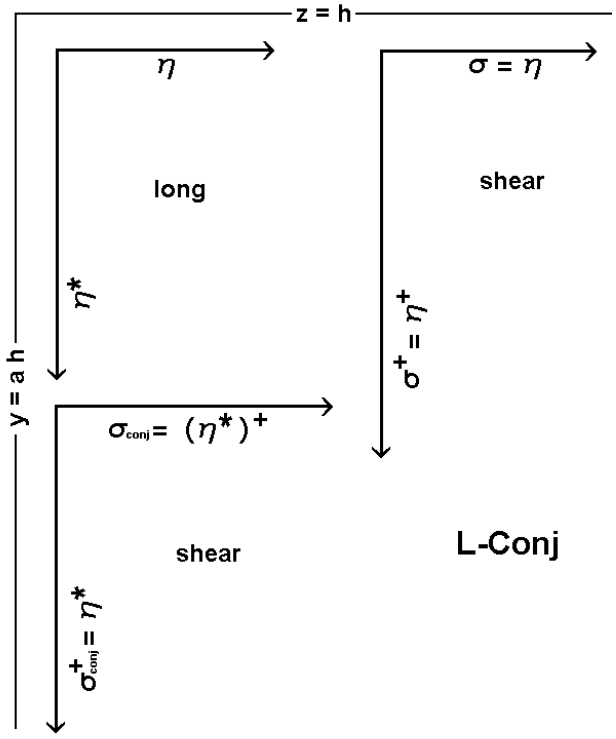

Figure 5.1: Illustrating the L-Conjugation Case for Modal Solutions. Shown are the relationships among longitudinal and shear wave vector components of the defining physical waves which must be among those making up the total superposition needed for a solution. The common x directional component (k) is normal to the page.

The recipe implied by L-Conjugation begins with a longitudinal wave at a desired k value. Conceptually, one could imagine starting out with some η and some ω close to a modal solution. Add a shear wave with the same fixed k. The polarization of this initial shear contribution is such that shear displacement is not parallel to the sides. Now adjust ω until the Rayleigh-Lamb equation is satisfied with respect to the $z = \pm h_z$ sides. There will be a range of ω's for which Rayleigh-Lamb can be satisfied at these parallel sides. For each possible ω the bulk dispersion relations will fix η^* and σ^+ wave vector components pointing along the z directions. Now add a second ("conjugate") shear wave at the same k and close

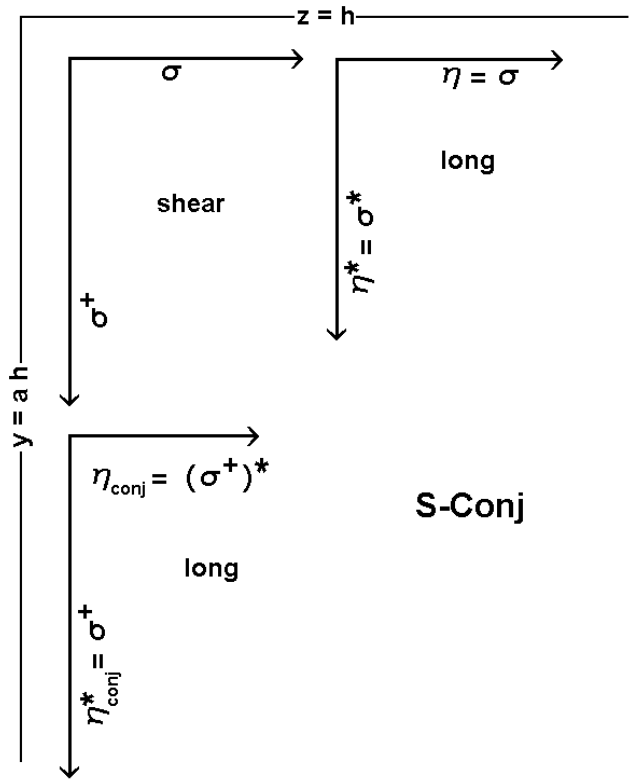

Figure 5.2: Illustrating the S-Conjugation Case for Modal Solutions. Shown are the relationships among longitudinal and shear wave vector components of the defining physical waves which must be among those making up the total superposition needed for a solution. The common x directional component (k) is normal to the page.

to the ω implied by the process so far. The polarization of this shear wave should also result in a displacement not parallel to the sides. The initial σ^+_{conj} is set equal to η^*. Just as the first shear wave had a wave vector component common with the foundation longitudinal wave along the y direction, this conjugate shear wave has a wave vector component common with the foundation longitudinal wave along the z direction. Bulk dispersion will fix the value of σ_{conj}. Now adjust ω over the range of values that continue to satisfy Rayleigh-Lamb at the $z = \pm h_z$ surfaces until one is found for which Rayleigh-Lamb is also satisfied for the foundation longitudinal wave and the just-added conjugate shear wave at the $y = \pm h_y$ surfaces. When

such an ω is found, it defines an eigenfrequency of a propagating system at the set value of k. There will be as many such ω's as there are subbands.

The recipe implied by S-Conjugation is procedurally the same as the recipe for L-Conjugation except that the foundation wave is a shear wave instead of a longitudinal one and two longitudinal ones are added instead of two shear waves.

The recipes each focus on combining waves in total disregard of their reflections at adjacent surfaces. Satisfaction of Rayleigh-Lamb incorporates reflections only at parallel surfaces. The reflections from adjacent surfaces being ignored constitute the rest of the superpositions. What the derivation shows is that to find the eigenspectrum, one can ignore these reflections. To build a complete description of the wave function, however, these reflections must be included. Incorporating the ignored reflections is not trivial. There are two principle complications. The first is that shear and longitudinal waves scatter into each other upon reflection and the scattering amplitude ratios are non trivial transcendental relations even when the surface environment is free of other interactions–which they are not. Secondly, from plotting the roots of the equation (5.27) paired alternately with equation (5.41) or equation (5.42) there appear to be cases where, for a given eigenfrequency, there are multiple root coincidences and it is not clear whether each is an independent foundation from which the reflections can be taken as emanating, or whether the multiple root coincidences are, in some sense, resonances of each other and building reflections from only one of them is sufficient.

Since the amplitude of reflection into either shear or longitudinal is bounded by unity, successive reflections must progressively dissipate in amplitude and the implication is that each foundation set constitutes a first-order characterization of the entire wave function.

In spite of the practical difficulties involved in building a complete description of the wave function that fully describes the superposition, the implication of the derivation remains a strong one. Namely, no matter how complicated the details of the wave function are, the dispersion is precisely defined by the behavior of only one overlapping set of components which must exist as a dominant part of the full superposition.

5.3.3 Mode Dispersions

The most important physical feature which can now be exhibited are the mode bands and particularly the dispersions. As the terms L-Conjugation and S-Conjugation will now appear more frequently and together, the contractions "L-Conj" and "S-Conj" will begin to be used routinely.

Equation (5.27) paired alternately with equation (5.41) or equation (5.42) defines simultaneous transcendental relationships which determine the propagating normal modes of an elastic isotropic rectangular waveguide. Extracting the actual subband dispersions is accomplished by substituting successive K values into a dimensionless form of equation (5.27), contour plotting its root system, and superimposing that root system on top of the root system plotted for equation (5.41) or (5.42). It is of some help that the latter equations are K-independent and need only be plotted once in dimensionless form.

The subbands resulting from L-Conjugation (viz. equations (5.27) and (5.41) combined) have a strong similarity to the standard Rayleigh-Lamb bands. This can be seen by methodically plotting the lowest subbands over a range of K values and superimposing them on the lowest subbands of the standard Rayleigh-Lamb solution. This is shown for dilatational modes in Fig. 5.3 where the dispersion for the first three subbands of these propagating waveguide modes are shown against

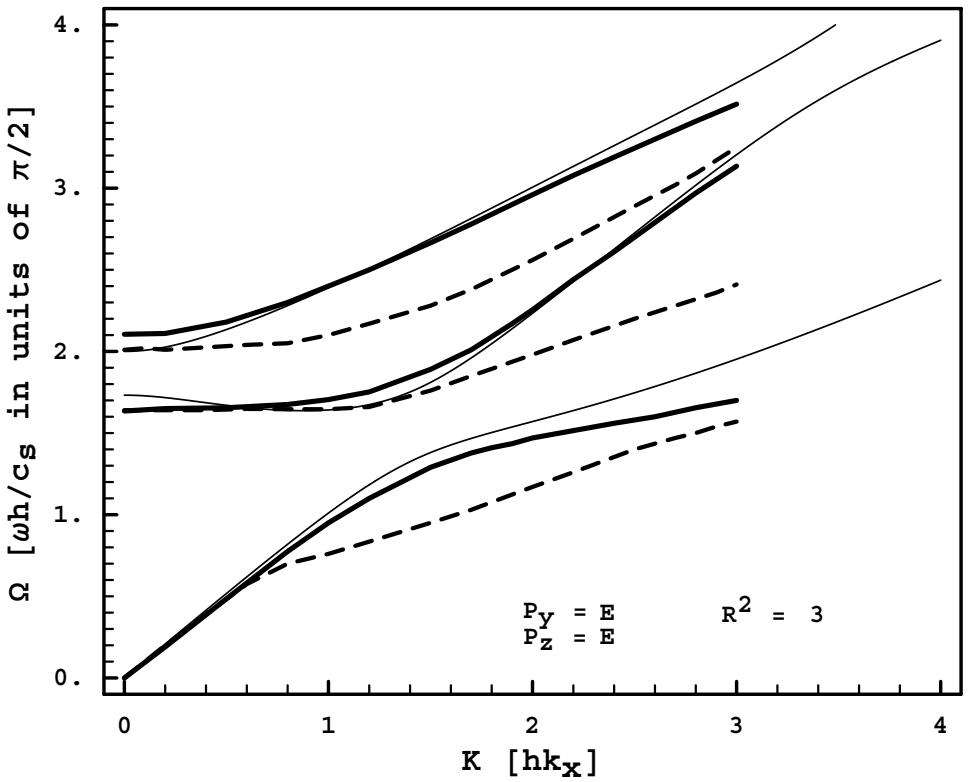

Figure 5.3: Plotted K-Dispersion for Propagating Dilatational Modes - L-Conjugate Case. Heavy lines correspond to waveguide modes (solid lines are square cross section, dashed lines are 1x2 cross section) and thin reference lines are Rayleigh-Lamb infinite plate "symmetric" solutions.

the background of the first three Rayleigh-Lamb branches. Waveguide modes for the square cross section (solid lines) are shown together with modes corresponding to a 1:2 cross-sectional aspect ratio (dashed lines). (L-Conj and S-Conj dilatational modes in combination are compared with numerical mode computations in Fig. 6.4.)

The dilatational subbands resulting from S-Conjugation are not intrinsically similar to infinite plate modes of the same parity pattern. Figure 5.4 shows the dispersion of the lower S-Conj subbands for the dilatational family against the background of the first three Rayleigh-Lamb branches. Waveguide modes for the square cross section (solid lines) are shown together with modes corresponding to a 1:2 cross-sectional aspect ratio (dashed lines).

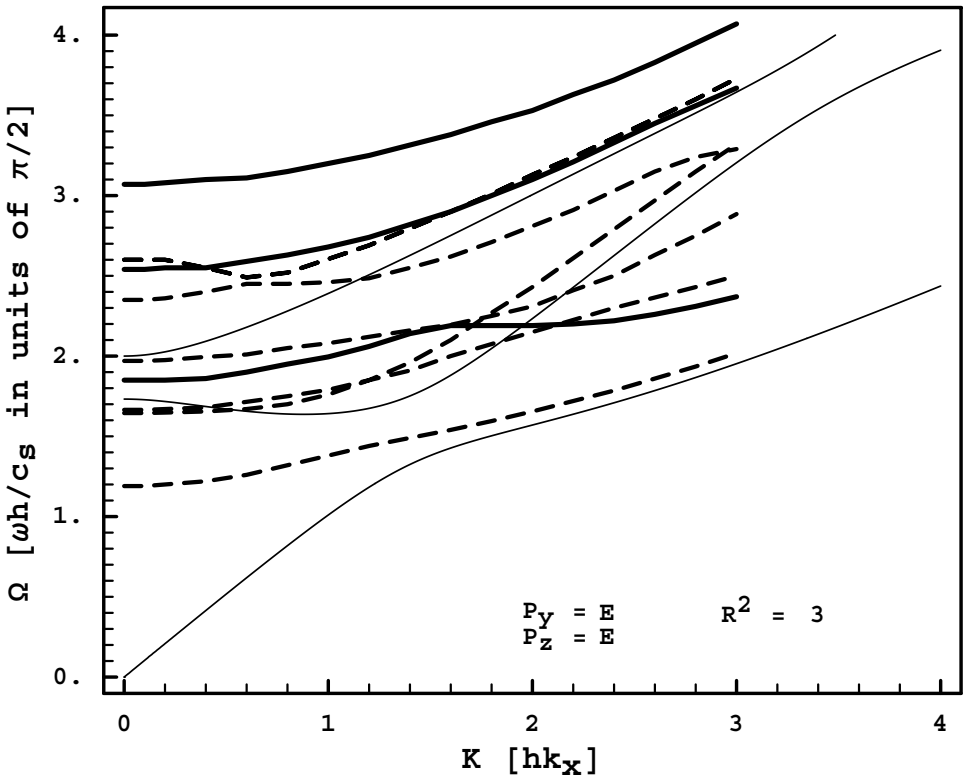

Figure 5.4: Plotted K-Dispersion for Propagating Dilatational Modes - S-Conjugate Case. Heavy lines correspond to waveguide modes (solid lines are square cross section, dashed lines are 1x2 cross section) and thin reference lines are Rayleigh-Lamb infinite plate "symmetric" solutions. For combined L-Conj and S-Conj dilatational modes compared to numerical results, please see Fig. 6.4.

Figure 5.5 shows the dispersion for L-Conjugated flexural solutions against a background of Rayleigh-Lamb flexural modes. The dashed lines correspond to a 1:2 cross-sectional aspect ratio where the flexing motion of the fundamental modes is of the same plane $(y-x)$ in both the Rayleigh-Lamb infinite plate and rectangular waveguide scenarios.

Although no torsion of an infinite plate is possible, the displacement pattern of low-order torsion modes (viz. dominantly parallel to the surfaces) is analogous to the uncoupled SH modes (viz. also parallel to the surfaces, but vanishing there) of the infinite plate derived at the end of section 4.4. Figure 5.6 shows torsional modes of a square waveguide against a background of infinite plate SH modes.

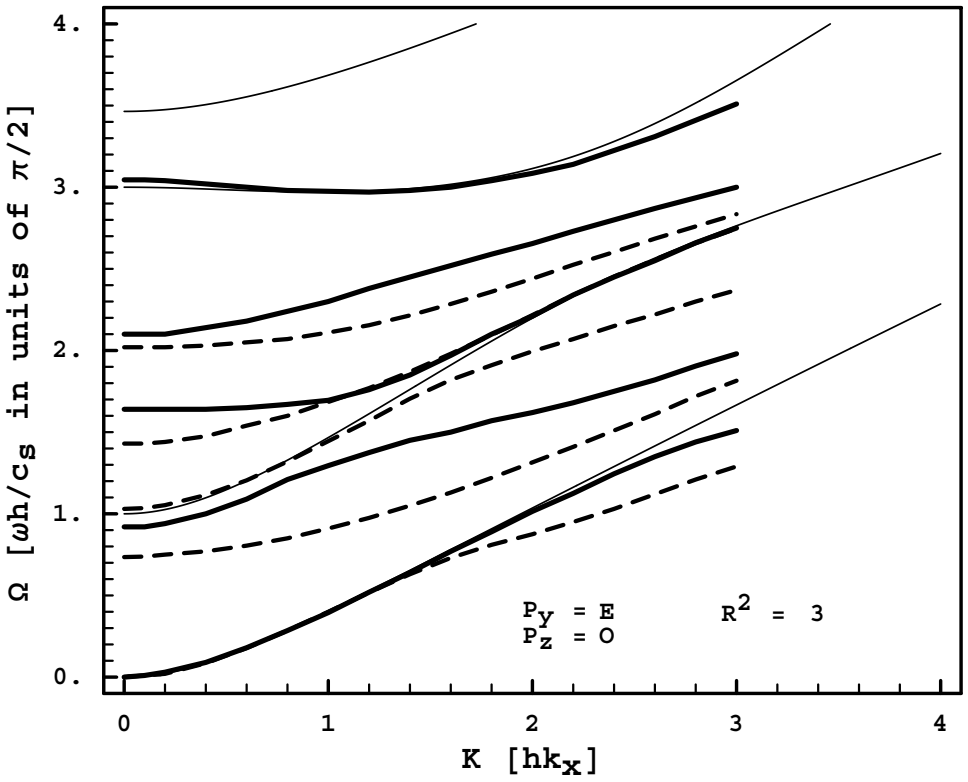

Figure 5.5: Plotted K-Dispersion for Propagating Flexural Modes - L-Conjugate Case. Heavy lines correspond to waveguide modes (solid lines are square cross section, dashed lines are 1x2 cross section) and thin reference lines are Rayleigh-Lamb infinite plate "antisymmetric" solutions. "Flexing" is of the $y - x$ plane.

Note that this figure combines in one display both L-Conjugate and S-Conjugate solutions and shows them against a background of combined even and odd SH plate modes. This figure does not attempt to show torsion modes other than for a square cross section.

It is surprising that there are more than one fundamental torsion mode–i.e., more than one torsion mode that goes to zero as $k \to 0$. The torsional dispersion for circular cross sections are analytically known and have a single fundamental mode (see Refs. [38,5,10]). While the rectangular cross section breaks the azimuthal symmetry of the circular cross section, it was not anticipated that consequent new modes would be fundamental (cf., assumptions in Ref. [47]).

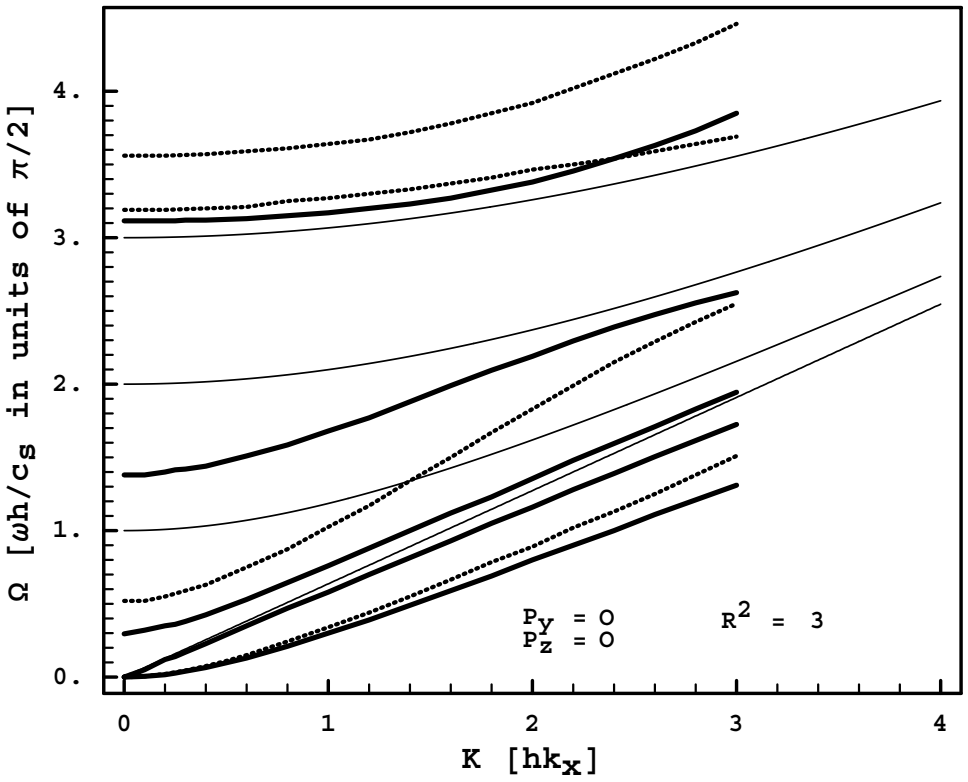

Figure 5.6: Plotted K-Dispersion for Propagating Torsional Modes - L-Conjugate and S-Conjugate Cases. Heavy lines correspond to square waveguide modes (solid lines are L-Conj and dotted lines are S-Conj) and thin reference lines are uncoupled SH modes of an infinite plate.

Two of the torsional mode branches are quadratic near $k = 0$ and their corresponding high density of states at low frequencies make them particularly interesting. It is worthwhile to extract some greater structural understanding of these mode branches and simultaneously demonstrate that the present theory does allow more than just the eigenspectrum to be exhibited.

As explained below, mode dispersions are obtained by meticulously plotting root families of the derived transcendental frequency equations and tracking how the frequencies corresponding to their intersections change as k is varied. This is accomplished using root plots like those displayed liberally throughout, but Figs. 5.7 and 5.8 show an actual progression of such intersections for the L-Conj and S-Conj cases of torsional modes for very small k. I will now show that these root

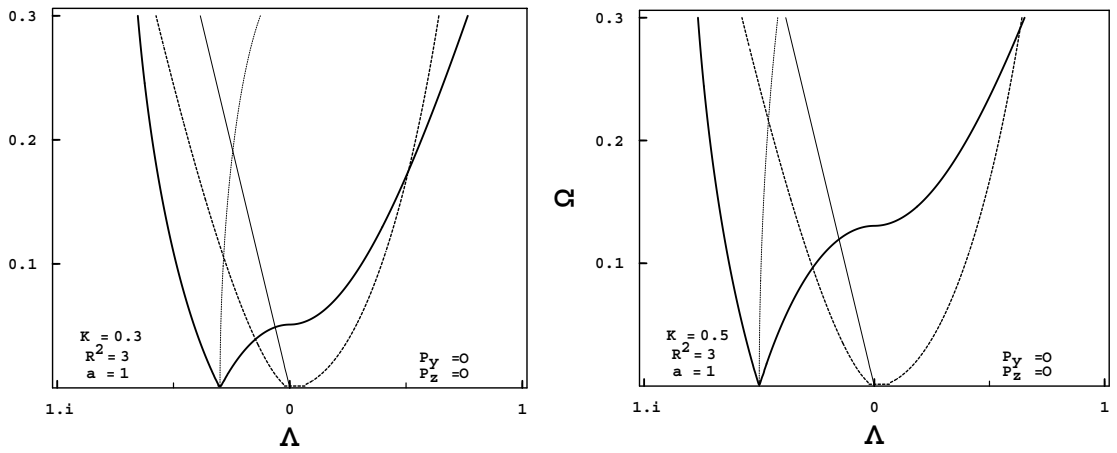

Figure 5.7: Mapping Fundamental Torsional Modes - L-Conjugate. Intersections of heavy solid and heavy dashed lines define two fundamental L-Conj torsional modes. Intersections to the right of dotted reference line correspond to imaginary Λ^* in the foundation longitudinal wave. Intersections to the right of the thin connected reference line correspond to real $(\Lambda^*)^+$. Correlate with Figure 5.1

plots are able to provide additional useful information about the structure of the foundation waves that characterize the modes (see Figs. 5.1, 5.2 and accompanying text).

The root plots show the zero contour in dimensionless Λ-Ω space corresponding to $f(\lambda, \omega) = 0$, where f is one of the frequency equations. In the same space, I can generate contours for other conditions. For example, the contours corresponding to $\lambda^+ = 0$, $\lambda^* = 0$, $(\lambda^+)^* = 0$, and $(\lambda^*)^+ = 0$ can also be plotted. These quantities are each a square root of some sum of pure real terms–one of which is λ^2. Now, as λ changes, the zeros of these terms always lie between a pure real and pure imaginary value. Plotting the zero contours of square of these terms always partitions the space between regions for which the term is imaginary or real. For terms containing $+\lambda^2$, the pure real region will be clockwise from the zero contour and for terms containing $-\lambda^2$, the pure real region will be counterclockwise from the zero contour. Hence, for a given root intersection, I can always establish the nature of λ from the horizontal axis and by superimposing a zero contour of the square of any of

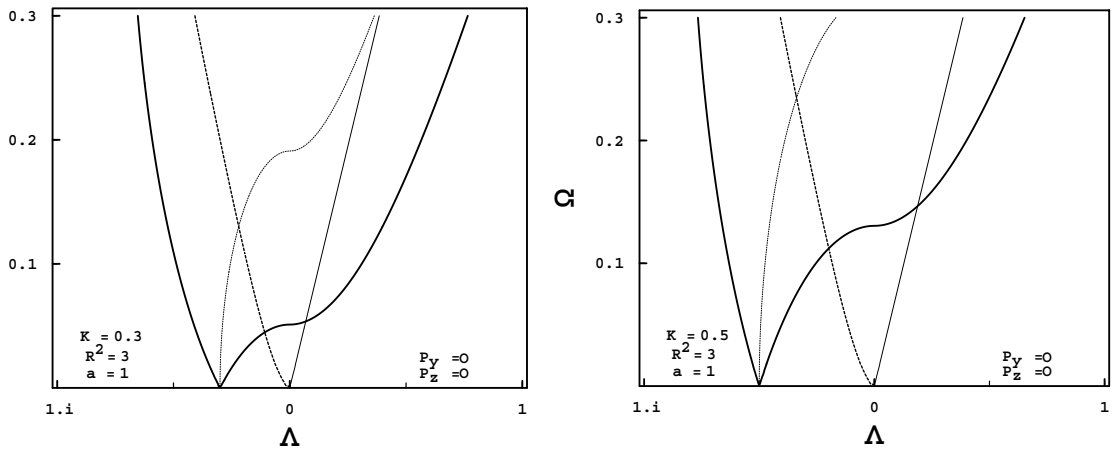

Figure 5.8: Mapping Fundamental Torsional Mode - S-Conjugate. Intersections of heavy solid and heavy dashed lines defines the fundamental S-Conj torsional mode. Intersections to the right of dotted reference line correspond to imaginary Λ^+ in the foundation shear wave. Intersections to the right of the thin connected reference line correspond to real $(\Lambda^+)^*$. Correlate with Figure 5.2

the foregoing terms, I can determine the nature of those terms with respect to the mode defined by that intersection as well.

Applying this technique to the torsional modes as they change with k (as shown in Figs. 5.7 and 5.8) I can now distinguish these modes in terms of the nature of the wave components of their L-Conj and S-Conj characteristic components. In terms of the nomenclature of Figs. 5.1, 5.2, the distinguishing characteristics among the fundamental torsional modes can be summarized as follows:

L-Conj Linear Branch η is real; η^* and $\eta^+ = \sigma^+$ are imaginary; $\sigma_{conj} = (\eta^*)^+$ is real; $\sigma^+_{conj} = \eta^*$ is imaginary.

L-Conj "Quadratic" Branch η is imaginary; η^* and $\eta^+ = \sigma^+$ are imaginary; $\sigma_{conj} = (\eta^*)^+$ is imaginary; $\sigma^+_{conj} = \eta^*$ is imaginary.

S-Conj "Quadratic" Branch σ is imaginary; σ^+ and $\sigma^* = \eta^*$ are imaginary; $\eta_{conj} = (\sigma^+)^*$ is imaginary; $\eta^+_{conj} = \sigma^+$ is imaginary.

Obviously, the thing that distinguishes the branches which are quadratic close to zero k is the fact that all of their transverse wave components are imaginary.

In this section I have not shown extensive examples of the contour plots of root families for all of the different parity families. I have chosen instead to show mode dispersions derived from them since these are the results of physical interest. However, a robust sampling of such contour plots for the $k = 0$ case are shown in section 6.4.2 and the interested reader is recommended to examine them in that context.

5.3.4 Comments on the Process of Mapping Mode Dispersions

Extracting the K vs Ω branch structure is presently a labor-intensive process. If it were not for the rapid evolution of affordable computing, each contour plot itself would be temporally expensive to produce, but in the course of this research, desktop computing power has increased by a factor of approximately forty and so the plots themselves can be rendered efficiently by a competent use of commercial software. Consequently, inspecting the overlaid plots and visually identifying the intersections has become the rate-limiting step.

The precision with which eigenfrequencies are determined depends upon the plotting precision. The best contour plots which appear in this dissertation were created by evaluating the contour space at 400x400 points. At that point density, each contour plot takes between two to five minutes to compute (viz. 16,000 equation computations) using a 600 MHz Pentium computer. The overlays must be rendered at high resolution and features must be zoomable so that an operator can visually identify the intersection point with adequate precision. If three digits of precision are desired, this process presently consumes between one and two hours per dispersion branch per 20 points.

In principle, the process of extracting the eigenspectrum and K-dispersions can be automated or at least rendered more "computer assisted." The contour plots are already stored as Postscript files which encapsulate a procedural description of the intended image and so it is conceivable that specialized software could be developed to accept multiple inputs of this type and then follow a custom algorithm to project the intersection points and then output them. Unlike numerical approximation methods, there is no theoretical limit on the accuracy of this process and there is no "estimate" per se which degrades in higher regions of the spectrum.

5.4 Disposing of the $H_x \neq 0$ Possibility

If H_y and H_z are taken to be zero and solutions are sought for $k > 0$ using H_x alone, then no systematic solution sets will be found. Rather than detail the attempts at derivation which ultimately fail to produce solutions, I will identify the underlying difficulties.

From equations (2.22), the boundary conditions at $z = h_z$ with $s \to z$, $p \to x$, $\bar{p} \to y$, $H_y = H_z = 0$ and $H_x \neq 0$ will be

$$\begin{aligned} -\beta\, k_\ell^2\, \varphi + 2\varphi_{,zz} &= 2H_{x,yz} \\ \varphi_{,zx} &= \tfrac{1}{2} H_{x,xy} \\ \varphi_{,zy} &= \tfrac{1}{2}(H_{x,yy} - H_{x,zz}) \end{aligned} \tag{5.43}$$

For the boundary conditions at $y = h_y$ with $s \to y$, $p \to z$, $\bar{p} \to x$, equations (2.22) take the following form:

$$\begin{aligned} -\beta\, k_\ell^2\, \varphi + 2\varphi_{,yy} &= -2H_{x,yz} \\ \varphi_{,yx} &= -\tfrac{1}{2} H_{x,xz} \end{aligned}$$

$$\varphi_{,zy} = \tfrac{1}{2}(H_{x,yy} - H_{x,zz}) \tag{5.44}$$

The superpositions to be substituted are again φ, defined as before, and H_x, which has the form

$$H_x = P_x(kx) \sum_j a_j \bar{P}_y(\sigma_j\, y)\, \bar{P}_z(\sigma_j^+\, z)$$

An early sign of difficulty can be seen immediately by noticing that in the previous derivation based on H_y, H_z, there were the same number of boundary conditions, but two more degrees of freedom for them to resolve. In the previous derivation, the transforms of H_y and H_z exhibited one each unknown expansion coefficient at each surface and the two bottom conditions were applied to resolve their relationship and thus eliminate one of them. Here, H_x has no companion and at the transform level there will be only one unknown coefficient per surface, but still two applicable independent boundary conditions. This imbalance will result in a constraint on the relationships among variables instead of the elimination of unknown constants. Ultimately, this leads to the following constraints:

$$\begin{aligned} \text{L-Conj case:} \quad \lambda^+ &= 0 \\ (\lambda^*)^+ &= 0 \end{aligned} \tag{5.45}$$

$$\begin{aligned} \text{S-Conj case:} \quad \lambda^+ &= 0 \\ \lambda &= 0 \end{aligned} \tag{5.46}$$

When these are substituted back into the boundary conditions, a frustrating series of contradictions arise justifying the conclusion that there are no non-trivial solution sets.

If the particular case of $k = 0$ is considered, one of its effects will be to eliminate entirely one of the boundary conditions and thus restore the balance between boundary constraints and unknown coefficients in the transformed equations. In the next chapter, this will be explored and it will be shown that while $H_x \neq 0$ does not generate any propagating modes within the wave guide, it does generate $k = 0$ transverse resonances which turn out to be the $k \rightarrow 0$ limit of the propagating solutions when H_y and H_z exist.

CHAPTER 6
$\mathbf{K} = 0$ MODES OF A RECTANGULAR WIRE

6.1 $\mathbf{k} = 0$ Boundary Conditions

As $k \to 0$, propagating modes approach cutoff limits and there are also non propagating modes resonating transversely in the waveguide. An ostensibly three dimensional problem is reduced to two dimensions for which the boundary conditions and representation of the potentials simplify significantly. Specifically, as $k \to 0$, the stipulated common x-dependency results in the vanishing of all derivatives with respect to x Therefore, $\varphi_{,x} \to 0$, and the curl of the vector potential $\mathbf{H}$ and strain components derived from its components simplify for the same reason, yeilding the following:

$$\mathbf{u^{(s)}} = \nabla\times \begin{bmatrix} H_x \\ H_y \\ H_z \end{bmatrix} \to \begin{bmatrix} H_{z,y} - H_{y,z} \\ H_{x,z} \\ -H_{x,y} \end{bmatrix} \tag{6.1}$$

$$\begin{aligned} u^{(s)}_{zz} &\to -H_{x,yz} \\ u^{(s)}_{yy} &\to H_{x,yz} \\ u^{(s)}_{xz} &\to \tfrac{1}{2}(H_{z,yz} - H_{y,zz}) \\ u^{(s)}_{xy} &\to \tfrac{1}{2}(H_{z,yy} - H_{y,yz}) \\ u^{(s)}_{yz} &\to \tfrac{1}{2}(H_{x,zz} - H_{x,yy}) \end{aligned} \tag{6.2}$$

Applying equation (6.2) and $\varphi_{,x} = 0$ to equation (2.22) at $z = \pm h_z$ and $y = \pm h_y$ surfaces produces the boundary conditions at $k = 0$:

at $z = h$ for all $y \in [-h, h]$

$$\begin{aligned} -\beta\ k_\ell^2\ \varphi + 2\varphi_{,zz} &= \quad 2H_{x,yz} \\ \varphi_{,yz} &= \ -\tfrac{1}{2}(H_{x,zz} - H_{x,yy}) \\ \varphi_{,xz} = 0 &= \ \tfrac{1}{2}(H_{z,yz} - H_{y,zz}) \end{aligned} \tag{6.3}$$

and at $y = h$ for all $z \in [-h, h]$

$$\begin{aligned} -\beta\ k_\ell^2\ \varphi + 2\varphi_{,yy} &= \quad -2H_{x,yz} \\ \varphi_{,yz} &= \ -\tfrac{1}{2}(H_{x,zz} - H_{x,yy}) \\ \varphi_{,xy} = 0 &= \ \tfrac{1}{2}(H_{z,yy} - H_{y,yz}) \end{aligned} \tag{6.4}$$

6.2 Uncoupled (Separable) $\mathbf{k} = 0$ Modes

The fact that the $k = 0$ boundary equations (6.3) and (6.4) include one condition independent of φ, suggests that I look for general solutions to the case $\varphi = 0$. These will be solutions for which there is no coupling of the shear and longitudinal parts and the superpositions representing components of the shear potential can collapse to single terms. Since all derivatives with respect to x are already expressed in the form of the boundary conditions, the common x dependency can be suppressed entirely. In respect of the representational scheme specified in section 4.2, the **H** components will have the following form:

$$H_x = A\bar{P}_y(\sigma_a\, y)\bar{P}_z(\sigma_a^+\, z)$$

$$H_y = BP_y(\sigma_b\, y)\bar{P}_z(\sigma_b^+\, z)$$

$$H_z = C\bar{P}_y(\sigma_c\, y)P_z(\sigma_c^+\, z) \tag{6.5}$$

Substituting into boundary conditions (6.3) and (6.4) at $\varphi = 0$ produces the following surface-paired conditions (recall my convention that $h_y = ah_z$):

$$\left.\begin{aligned} 0 &= \sigma_a\sigma_a^+ P_y(\sigma_a y)\, P_z(\sigma_a^+\, h_z) \\ 0 &= \sigma_a\sigma_a^+ P_y(\sigma_a\, ah_z)\, P_z(\sigma_a^+\, z) \end{aligned}\right\} A \neq 0 \tag{6.6}$$

$$\left.\begin{aligned} 0 &= ((\sigma_a^+)^2 - \sigma_a^2)\bar{P}_y(\sigma_a y)\, \bar{P}_z(\sigma_a^+\, h_z) \\ 0 &= ((\sigma_a^+)^2 - \sigma_a^2)\bar{P}_y(\sigma_a\, ah_z)\, \bar{P}_z(\sigma_a^+\, z) \end{aligned}\right\} A \neq 0 \tag{6.7}$$

$$\begin{aligned} C\sigma_c\sigma_c^+ \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_y} P_y(\sigma_c\, y) \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}_{P_z} \bar{P}_z(\sigma_c^+\, h_z) &= -B(\sigma_b^+)^2 P_y(\sigma_b\, y)\bar{P}_z(\sigma_b^+\, h_z) \\ C\sigma_c^2\bar{P}_y(\sigma_c\, ah_z)P_z(\sigma_c^+\, z) = -B\sigma_b\sigma_b^+ \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}_{P_y} \bar{P}_y(\sigma_b\, ah_z) \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_z} P_z(\sigma_b^+\, z) \end{aligned} \tag{6.8}$$

There are two independent cases defined by the alternatives of $A \neq 0$ versus $B, C \neq 0$.

Because sines and cosines have no common zeros, the only nontrivial solution for the $A \neq 0$ cases comes from $\sigma_a = \sigma_a^+ = \sigma$ with $P_y(\sigma\, h_z) = P_z(\sigma\, ah_z) = 0$. For the same reason, this eliminates solutions when $P_y \neq P_z$–which means simply that there are no solutions of this type for the flexural family of modes.

In the $B, C \neq 0$ cases, equations (6.8) harbor some subtleties. First, consider the special case of $\sigma_b = \sigma_c$. It can be readily established that algebraically this satisfies both equations simultaneously. However, in the cutoff limit, the only contribution to displacement of H_y and H_z is to define $u_x^{(s)} = 2H_{[z,y]}$. After carefully considering the cases of $P_y = P_z$ and $P_y = \bar{P}_z$, it can be shown that B/C ratio

that results from equating the sigmas also forces $u_x^{(s)}$ to vanish everywhere. So, $\sigma_b = \sigma_c$ turns out to be a null solution.

Then examine equations (6.8) and notice that if either B or C are zero, that the surviving σ will parameterize solutions defined by $\bar{P}_z(\sigma^+ h_z) = \bar{P}_y(\sigma\, a h_z) = 0$ regardless of which constant is zero.

If B and C are nonzero simultaneously, I can divide the equations to eliminate the coefficients. Then it appears possible to find roots defined by $\bar{P}_z(\sigma_a^+ h_z) = \bar{P}_z(\sigma_b^+ h_z) = 0$ or, independently, by $\bar{P}_y(\sigma_a\, a h_z) = \bar{P}_y(\sigma_b\, a h_z) = 0$. This, however, is illusory since in the limits $z \to h_z$ and $y \to h_y$ the ratios of these will approach nonzero constants. Apparently, solutions when both constants are nonzero are generated purely by the simultaneity of $\bar{P}_z(\sigma_a^+ h_z) = \bar{P}_y(\sigma_a\, a h_z) = 0$ and $\bar{P}_z(\sigma_b^+ h_z) = \bar{P}_y(\sigma_b\, a h_z) = 0$–which is just a superposition of independent solutions.

Using dimensionless units (see section 4.3), the worked out uncoupled solutions can be summarized.

First, the $A \neq 0$ case is restricted to the dilatational and torsional mode families and yields the following results:

dilatational family ($P_y = P_z = E$)

$$\Omega = \sqrt{(2p-1)^2 + \frac{(2q-1)^2}{a^2}}\left(\frac{\pi}{2}\right) \quad p, q = 1, 2, 3 \ldots \tag{6.9}$$

Subject to the following constraint:

$$q(integer) = \frac{1 + (2p-1)a}{2} \quad (a = 1 \Rightarrow p = q)$$

torsional family $(P_y = P_z = O)$

$$\Omega = 2\sqrt{p^2 + \frac{q^2}{a^2}}\left(\frac{\pi}{2}\right) \quad p, q = 0, 1, 2 \ldots \tag{6.10}$$

Subject to the following constraint: $q(integer) = ap$

This group of modes is fragile in the sense they can only exist for special exact values of a–starting with $a = 1$. These are an enumeration of so-called Lamé modes which correspond to shear waves polarized in a plane normal to the surfaces at an incidence of 45^o ($\sigma = \sigma^+$). At that precise incidence angle, it is well known (see e.g., Vol II Ch. 9 Ref. [38]) that the scattering amplitude for vertically polarized (SV) shear waves to reflect as longitudinal waves is zero–consistent with the $\varphi = 0$ condition stipulated.

Secondly, the $B, C \neq 0$ cases for each mode family yields the following:

dilatational family $(P_y = P_z = E)$

$$\Omega = 2\sqrt{p^2 + \frac{q^2}{a^2}}\left(\frac{\pi}{2}\right) \quad p, q = 0, 1, 2 \ldots \tag{6.11}$$

torsional family $(P_y = P_z = O)$

$$\Omega = \sqrt{(2p-1)^2 + \frac{(2q-1)^2}{a^2}}\left(\frac{\pi}{2}\right) \quad p, q = 1, 2, 3 \ldots \tag{6.12}$$

flexural family $(P_y \neq P_z)$

$$\Omega = \sqrt{(2p-1)^2 + \frac{q^2}{4a^2}}\left(\frac{\pi}{2}\right) \quad p = 1, 2, 3 \ldots, q = 0, 1, 2 \ldots \tag{6.13}$$

$$\Omega = \sqrt{\frac{(2p-1)^2}{a^2} + \frac{q^2}{4}}\left(\frac{\pi}{2}\right) \quad p = 1, 2, 3 \ldots, q = 0, 1, 2 \ldots \tag{6.14}$$

The degeneracy of the flexural modes for a square cross section is broken when $a \neq 1$ and the asymmetry of the affected equation augments the resulting distinction. For the other families, the fact that p, and q have the same enumeration obviates restating the equation with a juxtaposed.

While the $A \neq 0$ case involves vertically polarized (SV) shear surface incidence, the $C, B \neq 0$ case involves only $u_x^{(s)} \neq 0$–shear waves polarized parallel to the surface (SH waves). SH waves reflect entirely as themselves at all angles (provided they go to zero at the surface) and thus do not involve coupling to longitudinal waves. (again, see Ch. 9 Ref. [38]) These are the two dimensional extensions of the SH solutions for an infinite plate described at the end of section 4.4.

6.3 Uncoupled Modes not the Limit of Propagating Modes

The foregoing uncoupled solutions at $k = 0$ are not continuously connected to any set of propagating modes. This conclusion is in conflict with the results of numerical methods described in section 3.2.

One way to realize that uncoupled $k = 0$ modes do not connect continuously to propagating modes begins with a physical argument. These uncoupled solutions are uncoupled either because none of their shear displacements have components perpendicular to the surfaces, or because their vertical components arise at the one exception to coupling which occurs when the wave vector angle of incidence is precisely 45^o. If I perturb either situation by adding a wave vector component in the x direction–however small–I will necessarily induce some displacement component vertical to the surfaces that will result in a partial refraction of longitudinal waves. In other words, perturbation by adding $k > 0$ requires $\varphi \neq 0$. If φ cannot be identically zero, then the boundary constraints no longer simplify and the derivation of propagating modes, already presented, defines the solutions. The

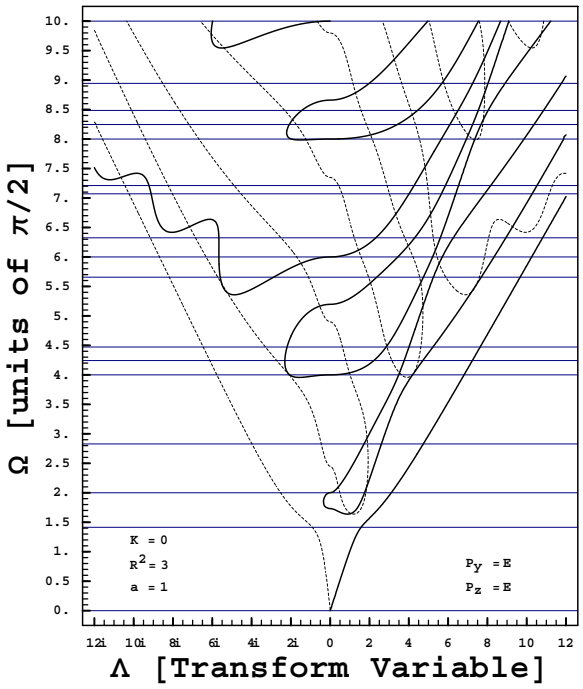

Figure 6.1: $k = 0$ Propagating Modes Compared to $k = 0$ Uncoupled Modes– L-Conj Dilational Modes in Square Cross Section. Intersections of curves define propagating modes. Horizontal lines at eigenvalues of uncoupled $k = 0$ (SH and Lamé) modes do not generally correspond to $k \to 0$ limits of propagating modes.

$k \to 0$ limits of the solution set for propagating modes, as shown in the dispersion plots within section 5.3.3, does not correspond to the solutions for uncoupled modes.

That the $k \to 0$ limit of propagating modes are not coincident with the uncoupled solutions can be seen directly by superimposing the eigenspectrum of uncoupled modes onto the $k = 0$ root coincidences of any parity family. Figures 6.1 and 6.2 show L-Conj and S-Conj root intersections for the dilatational system of a square waveguide with horizontal lines at values of the uncoupled SH and Lamé uncoupled eigenvalues. As the figures show, there are no systematic agreements between the uncoupled eigenvalues and the root system intersections. The same disparity occurs for all of the torsional and flexural modes, though figures showing these are omitted.

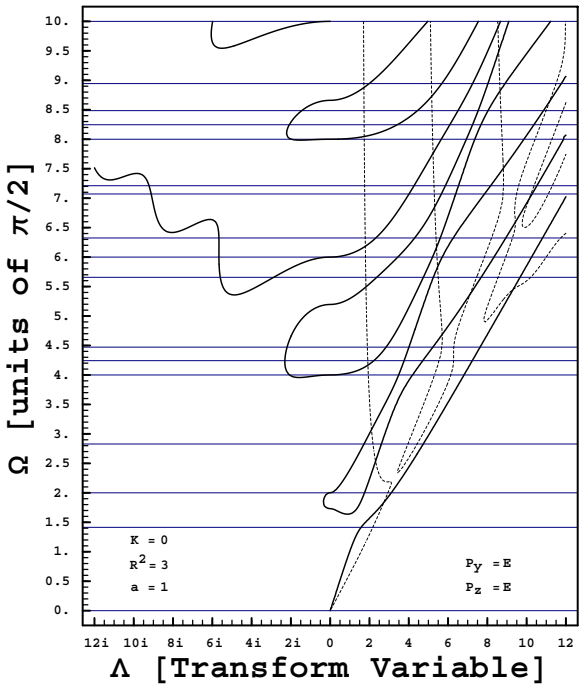

Figure 6.2: $k = 0$ Propagating Modes Compared to $k = 0$ Uncoupled Modes–S-Conj Dilational Modes in Square Cross Section. Intersections of curves define propagating modes. Horizontal lines at eigenvalues of uncoupled $k = 0$ (SH and Lamé) modes do not generally correspond to $k \to 0$ limits of propagating modes.

Numerical solutions of waveguide modes which apply the numerical procedure discussed in section 3.2 do show a continuous connection between uncoupled solutions and propagating ones. This can be seen, for example, by comparing the small-k part of the numerically generated subbands to the uncoupled eigenvalues. Figure 6.3 shows that some of the numerically produced mode families converge to uncoupled values as $k \to 0$. Nishiguchi's application [1] of the numerical procedure–using a basis formed from powers of the coordinates–shows the same result.

The inability of the numerical process to distinguish nonpropagating modes from propagating ones was actually an impediment to developing an analytic method since numerous efforts to contrive an analytic solution that agreed with

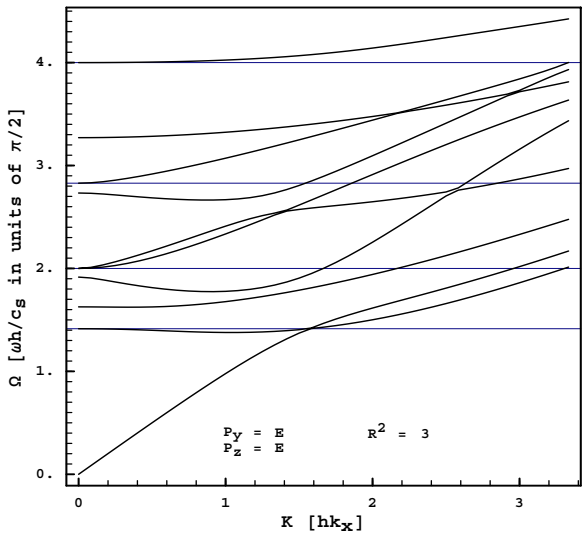

Figure 6.3: Numerically Approximated Modes Compared to $k = 0$ Uncoupled Modes –Dilational Modes in Square Cross Section. Dilatational subbands computed using the powers of coordinates as a basis for numerical method developed by Visscher et al. and utilized by Nishiguchi. Horizontal lines at eigenvalues of uncoupled $k = 0$ (SH and Lamé) modes each correspond (incorrectly) to $k \to 0$ limit of a propagating modes.

numerical results were unfruitful. This disparity can now be understood and a significant limitation of the numerical process finally appreciated.

At $k = 0$, the numerical method converges–at least at the lower part of the spectrum–to modes consistent with that value of k. These will include, among others, the uncoupled $k = 0$ modes. In fact, I have found that the modes produced by the numerical process which appear to coincide with uncoupled $k = 0$ modes, agree with uncoupled modes determined analytically to at least six significant digits. These uncoupled modes involve zeros of simple sines and cosines and, as long as the periodicity is low (i.e., $\eta h/\pi < 2$) powers of the coordinates will converge strongly to such solutions. The numerical process is based on finding the eigenvalues of a finite-order matrix. In this case k is just a parameter of the calculation. As $k > 0$, the eigenvalues will be perturbed, but for $k \ll 1$ only slightly. But the finite order of the matrix does not change with k and

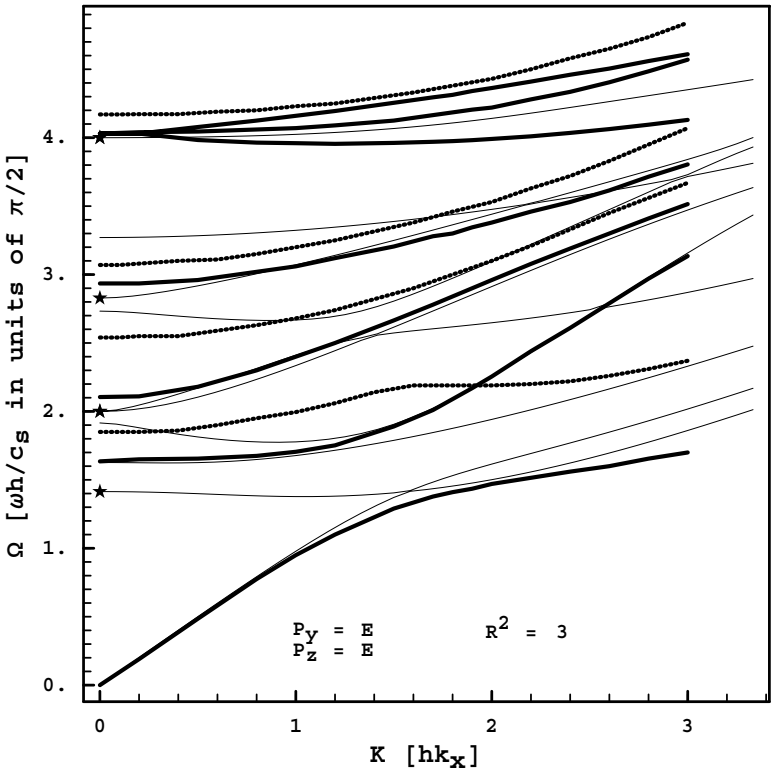

Figure 6.4: Analytically Derived Propagating Modes compared to Numerically Approximated Modes –Dilational Modes in Square Cross Section. Stars on Ω axis denote uncoupled $k = 0$ modes. Thick lines are analytic modes (solid are L-Conj, broken S-Conj) and thin lines are from numerical method as used by Nishiguchi.

the relaxation is not absolutely consistent with stress-free boundary conditions, but only approaches it as best the finite basis will allow. The fixed finite order precludes entirely new modes from suddenly appearing and the numerical process makes it appear as if the eigenfunctions are small perturbations from their $k = 0$ forms. What is actually happening, however, is that combinations of the finite basis elements approximating the propagating functions is begining to be mixed into the eigenfunctions which do not propagate giving an illusion of continuity. Eventually, as k continues to grow, the subbands which appeared to start out from the uncoupled $k = 0$ modes, take on more of the characteristics of the true propagating modes. This is consistent with Figure 6.4 which shows L-Conj and

S-Conj Dilatation modes of a square wire against a background of numerically approximated subbands.

6.4 Coupled (nonseparable) $\mathbf{k} = 0$ Modes

6.4.1 Derivation–All Parity Families

If the right sides of boundary equations (6.3) and (6.4) are set to zero, it is quickly seen that absent a superposition of some kind, no parity pattern of φ will be able to satisfy the equations. Looking at the derivatives to be taken, the impossibility of common zeros of sines and cosines of the same argument is decisive.

Further progress can only be achieved by facing the nonseparability issue and substituting superpositions of basic functions. However, this can now be accomplished using a somewhat abbreviated version of the process used to derive propagating modes.

Substituting representations from equations (4.8) thru (4.10) into the boundary conditions (6.3) at $z = h_z$, I obtain:

$$\begin{aligned} &-P_x \sum_i d_i(\beta k_\ell^2 + 2(\eta_i^*)^2) P_y(\eta_i\, y) P_z(\eta_i^*\, h_z) = \\ &2P_x \sum_j a_j\, \sigma_j \sigma_j^+ \left\{ \begin{matrix} 1 \\ -1 \end{matrix} \right\}_{P_y} P_y(\sigma_j\, y) \left\{ \begin{matrix} 1 \\ -1 \end{matrix} \right\}_{P_z} P_z(\sigma_j^+\, h_z) \end{aligned} \tag{6.15}$$

$$\begin{aligned} &P_x \sum_i d_i\, \eta_i \eta_i^* \left\{ \begin{matrix} -1 \\ 1 \end{matrix} \right\}_{P_y} \bar{P}_y(\eta_i\, y) \left\{ \begin{matrix} -1 \\ 1 \end{matrix} \right\}_{P_z} \bar{P}_z(\eta_i^*\, h_z) = \\ &-\tfrac{1}{2} P_z \sum_j a_j\, (-(\sigma_j^+)^2 + \sigma_j^2) \bar{P}_y(\sigma_j\, y) \bar{P}_z(\sigma_j^+\, h_z) \end{aligned} \tag{6.16}$$

Applying the simple transform from section 4.5 and denoting with λ_o any of the possible values of the transform variable which simultaneously match elements of both $\{\eta_i\}$ and $\{\sigma_j\}$, the sums collapse and the relationships become: (Reminder: under the transform, $\eta_i \rightarrow |\lambda|$ etc., but to avoid clutter, I will omit the absolute sign notations so that $\lambda_o \equiv |\lambda_o|$. Recall also, that all square roots are assumed positive. Both λ_o and any square roots shown could, however, be positive imaginary.)

$$-d_o(\beta k_\ell^2 + 2(\lambda_o^*)^2)P_z(\lambda_o^* h_z) = 2a_o\,\lambda_o\lambda_o^+ \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_y} \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_z} P_z(\lambda_o^+ h_z) \tag{6.17}$$

$$d_o\,\lambda_o\lambda_o^* \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}_{P_y} \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}_{P_z} \bar{P}_z(\lambda_o^* h_z) = \tfrac{1}{2}a_o((\lambda_o^+)^2 - \lambda_o^2)\bar{P}_z(\lambda_o^+) \tag{6.18}$$

Inspired again by similarities with the Rayleigh-Lamb derivation (see section 4.4), I can manipulate the parenthetical on the left of equation (6.17) to show that

$$(\beta k_\ell^2 + 2(\lambda_o^*)^2) = -((\lambda_o^+)^2 - \lambda_o^2)$$

I then divide (6.17) by (6.18), apply the foregoing manipulation, and rearrange terms to produce

$$\frac{P_z(\lambda_o^* h_z)}{\bar{P}_z(\lambda_o^* h_z)}\frac{\bar{P}_z(\lambda_o^+ h_z)}{P_z(\lambda_o^+ h_z)} = \frac{-4\,\lambda_o^2\,\lambda_o^*\,\lambda_o^+}{(\lambda_o^2 - (\lambda_o^+)^2)^2} \tag{6.19}$$

In this equation, ω is defined implicitly by λ_o^* and λ_o^+ and so roots of this equation define the valid combinations of ω and λ_o.

Comparing equation (6.19) to equation (4.20) reveals it to be a Rayleigh-Lamb solution with λ_o playing the role of the wave number in the propagation direction. In the $k \rightarrow 0$ limit this makes even more physical sense than it did when

encountered in deriving the propagating modes. Before I impose adjacent sides, the physical situation is precisely the Rayleigh-Lamb scenario. If wave activity survives the cutoff limit, it must be limited to the transverse direction–which is y at the $z = h_z$ boundary. The transform of the y-dependency has mapped to a transform variable that is, in effect, a wave number in that direction. If we are in a cutoff mode, then there should be no non-uniform displacement in the x direction–which is precisely the same condition as requiring plane waves. We already know that the Rayleigh-Lamb equation defines a solution in this set of circumstances and so the result is consistent.

To complete the derivation, I take the same representations of φ and H_x and insert them into the boundary equations (6.4) at the $y = a\,h_z$ surface. In order to see clearly the logic used to link the transformed boundary conditions between adjacent surfaces, I will start by writing down the situation following the transformation of the conditions, but before the delta functions have been used to transform wave number variables or the sums have been collapsed.

$$\sum_i d_i(\beta k_\ell^2 + 2\eta_i^2) P_y(\eta_i\, a h_z)\, \delta(\eta_i^* - \mu) =$$
$$2\sum_j a_j\, \sigma_j \sigma_j^+ \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_y} P_y(\sigma_j\, a h_z) \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}_{P_z} \delta(\sigma_j^+ - \mu) \tag{6.20}$$

$$\sum_i d_i\, \eta_i \eta_i^* \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}_{P_y} \bar{P}_y(\eta_i\, a h_z) \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}_{P_z} \delta(\eta_i^* - \mu) =$$
$$\tfrac{1}{2}\sum_j a_j\, ((\sigma_j^+)^2 - \sigma_j^2) \bar{P}_y(\sigma_j\, a h_z)\, \delta(\sigma_j^+ - \mu) \tag{6.21}$$

In the course of derving the propagating modes, I have developed the logic for linking the transformed boundary value constraints at the $z = \pm h_z$ surface with the transformed boundary constraints at the adjacent $y = \pm h_y$ ($h_y = ah_z$) surface. (See section 4.4) The same logic applies here and, as was done in the propagating mode case, I show details only for the L-Conj case and simply state the analagous result for the S-Conj case.

Applying the mappings (5.31) to equations (6.20) and (6.21) and collapsing the sums, I obtain the following:

$$d_o(\beta k_\ell^2 + 2\lambda_o^2)P_y(\lambda_o\, ah_z) = 2a_1\, \lambda_o^*(\lambda^*)^+ \left\{ \begin{matrix} 1 \\ -1 \end{matrix} \right\}_{P_y} P_y((\lambda_o^*)^+\, ah_z) \left\{ \begin{matrix} 1 \\ -1 \end{matrix} \right\}_{P_z} \tag{6.22}$$

$$d_o\, \lambda_o\lambda_o^* \left\{ \begin{matrix} -1 \\ 1 \end{matrix} \right\}_{P_y} \bar{P}_y(\lambda_o\, ah_z) \left\{ \begin{matrix} -1 \\ 1 \end{matrix} \right\}_{P_z} = \tfrac{1}{2}a_1\, [(\lambda_o^*)^2 - ((\lambda_o^*)^+)^2]\, \bar{P}_y((\lambda_o^*)^+\, ah_z) \tag{6.23}$$

As has been the pattern thus far, the parenthetical on the left of equation (6.22) can be shown to be the negative of the bracketed difference on the right hand side of equation (6.23). With this applied, dividing the two equations and rearranging terms produces the following:

$$\frac{P_y(\lambda_o\, ah_z)}{\bar{P}_y(\lambda_o\, ah_z)} \frac{\bar{P}_y((\lambda_o^*)^+\, ah_z)}{P_y((\lambda_o^*)^+\, ah_z)} = \frac{-4(\lambda_o^*)^2\lambda_o(\lambda_o^*)^+}{[(\lambda_o^*)^2 - ((\lambda_o^*)^+)^2]^2} \tag{6.24}$$

Comparing equation (6.24) to equation (4.20) reveals that it also is a Rayleigh-Lamb solution, this time with λ_o^* playing the role of the wave number in the propagation direction. This too, in isolation, makes physical sense. If I ignore the adjacent sides, the physical situation can be interpreted as a Rayleigh-Lamb scenario. The transform of the z-dependency has mapped to a transform variable that is, in effect, a wave number in that direction. If we are in a cutoff mode and

pretend the adjacent sides do not exist, then there should be no non-uniform displacement in the x direction–which is the same condition as requiring plane waves. Again, the Rayleigh-Lamb equation defines a solution in this set of circumstances and so the result is consistent.

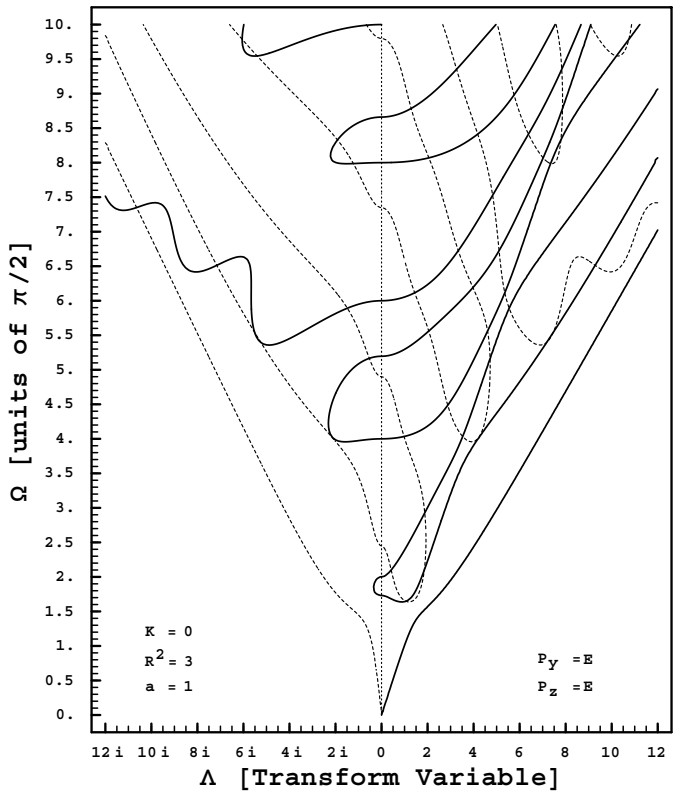

Figure 6.5: Plotted Root Families for L-Conj Dilatational Cutoff Modes–Square Cross Section. Solid curves represent roots of equation (6.19) and dashed curves represent roots of equation (6.24).

Applying mappings (5.32) to equations (6.20) and (6.21) leads to the following S-Conj solution analagous to equation (6.24):

$$\frac{P_y((\lambda_o^+)^*\, ah_z)}{\bar{P}_y((\lambda_o^+)^*\, ah_z)}\frac{\bar{P}_y(\lambda_o\, ah_z)}{P_y(\lambda_o\, ah_z)} = \frac{-4(\lambda_o^+)^2\lambda_o(\lambda_o^+)^*}{[(\lambda_o^+)^2 - \lambda_o^2]^2} \tag{6.25}$$

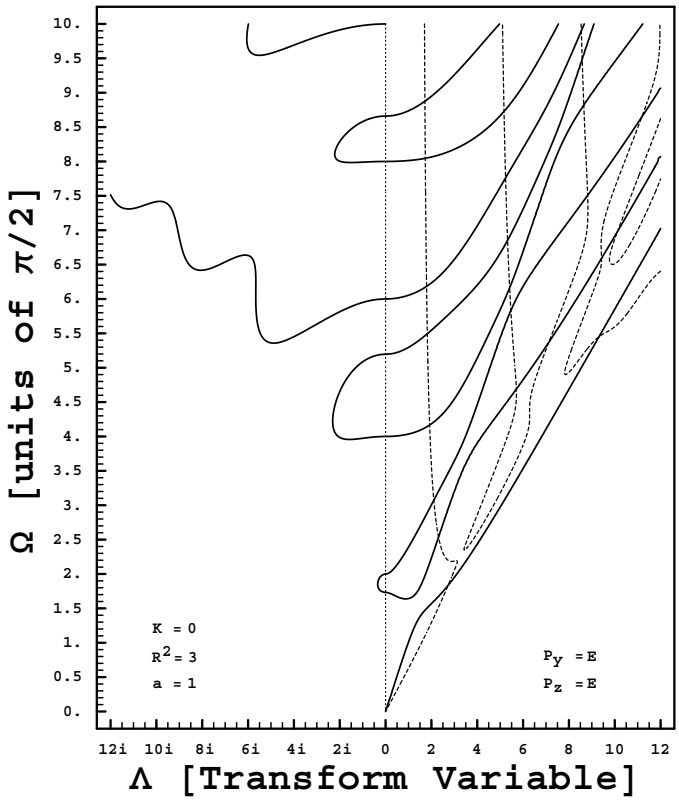

Figure 6.6: Plotted Root Families for S-Conj Dilatational Cutoff Modes–Square Cross Section. Solid curves represent roots of equation (6.19) and dashed curves represent roots of equation (6.25).

Since the foregoing solutions were premised upon $H_y = H_z = 0$ with $H_x \neq 0$, it is counterintuitive that these solutions would be a special case of the propagating modes for which $H_x = 0$. That is, however, precisely the case. It can be checked even by sight that equation (6.19) is precisely equation (5.27) with $k = 0$, that equation (6.24) is precisely equation (5.41) with $k = 0$, and finally, that equation (6.25) is precisely equation (5.42) with $k = 0$.

6.4.2 Manifestations By Parity Family

In the same manner as Rayleigh-Lamb dispersion curves were plotted (see section 4.4) the root systems of the two transcendental equations (6.19) and (6.24) can be found by contour plotting. Where the root systems intersect, eigenvalues of

the coupled $k = 0$ modes will be found. These solutions will be made specific to the three parity families by making appropriate substitutions of sine or cosine functions into P_y and P_z. To preserve generality of the result, graphical results will be shown in dimensionless units (see section 4.3). As was the case for plotting Rayleigh-Lamb solutions, a rearrangement of the equations is necessary to preclude fatal divergences in the numerical computation involved in the plotting. For example, equation (6.19), in dimensionless form, is rearranged for computation into the following form:

$$\begin{Bmatrix} \cos(\Lambda^*)\frac{\sin(\Lambda^+)}{\Lambda^+} \\ \frac{\sin(\Lambda^*)}{\Lambda^*}\cos(\Lambda^+) \end{Bmatrix}_{P_z} [(\Lambda^+)^2 - \Lambda^2]^2 + 4 \begin{Bmatrix} \Lambda^* \\ \Lambda^+ \end{Bmatrix}_{P_z} \Lambda^2 \, \bar{P}_z(\Lambda^*) \, P_z(\Lambda^+) = 0 \tag{6.26}$$

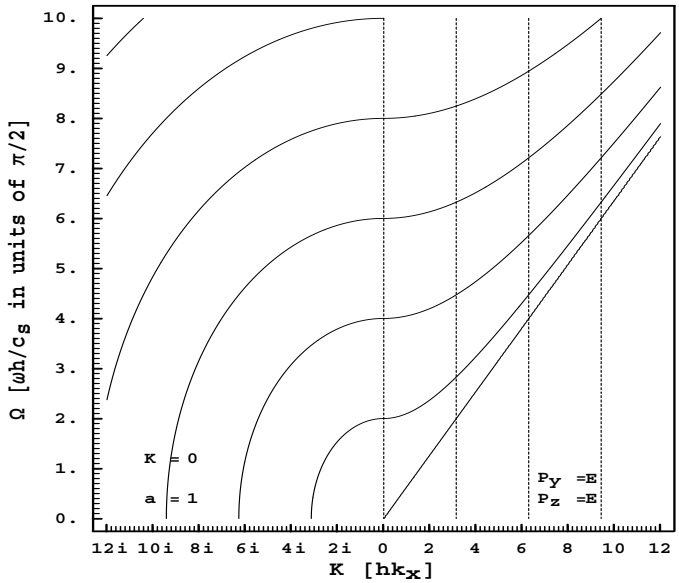

Figure 6.7: Root Families for Uncoupled Shear Modes–Square Cross Section. This figure illustrates how the root families would appear for S-Conj solutions if all coupling to longitudinal waves were to be removed.

Figure 6.5 shows the result of substituting the dilatational parity pattern into the foregoing solutions in the case of a square cross section. Intersections of the

root systems occur at values of Ω which are eigenvalues of the problem. S-Conj modes have the same root system arising from the $z = \pm h_z$ boundary as do L-Conj solutions (viz. equation (6.19)), but the root system arising from the $y = \pm h_y$ boundary shifts from equation (6.24) to equation (6.25). Figure 6.6 shows the intersecting root systems for S-Conj dilatational modes in a square cross section. Unlike the L-Conj cases, these only involve values of Λ which are real. The conjugate root system (i.e., roots of equation (6.25)) also displays greater regularity than L-Conj modes except near the maximal values of λ_o.

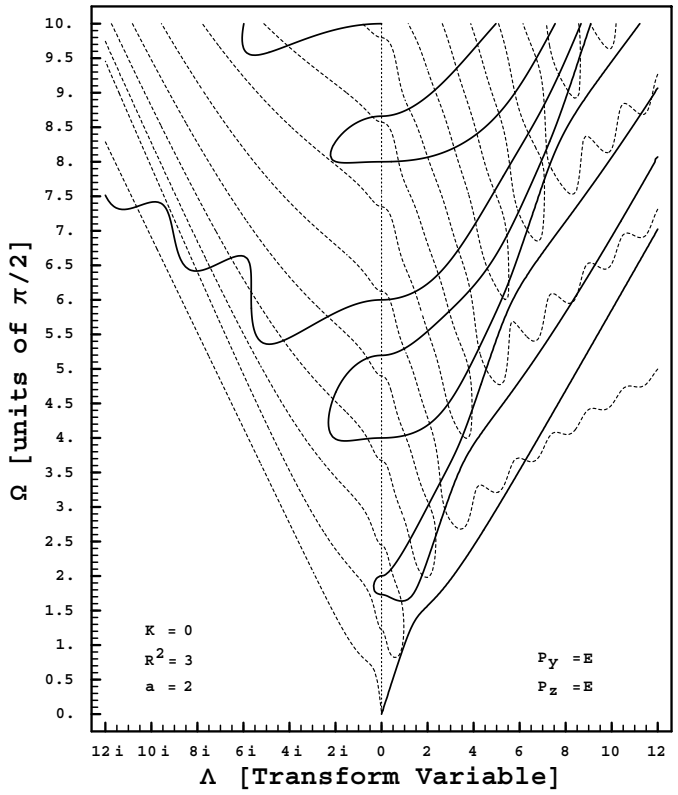

Figure 6.8: Plotted Root Families for L-Conj Dilatational Cutoff Modes–1x2 Cross Section. Solid curves represent roots of equation (6.19) and dashed curves represent roots of equation (6.24).

The S-Conj modes other than along the edge defined by maximal values of Λ, appear to reflect weaker coupling of shear to longitudinal. To see this, compare the small Λ–large Ω area of the root contours in Figure 6.6 to the root contours

within the same region of an uncoupled system in Figure 6.7 for which the only boundary condition is that the derivative of the wave function (i.e., EE family or $\cos(\sigma\, y)\cos(\sigma^{+}\, z)$ within the boundaries) in the surface normal direction is zero at the surfaces.

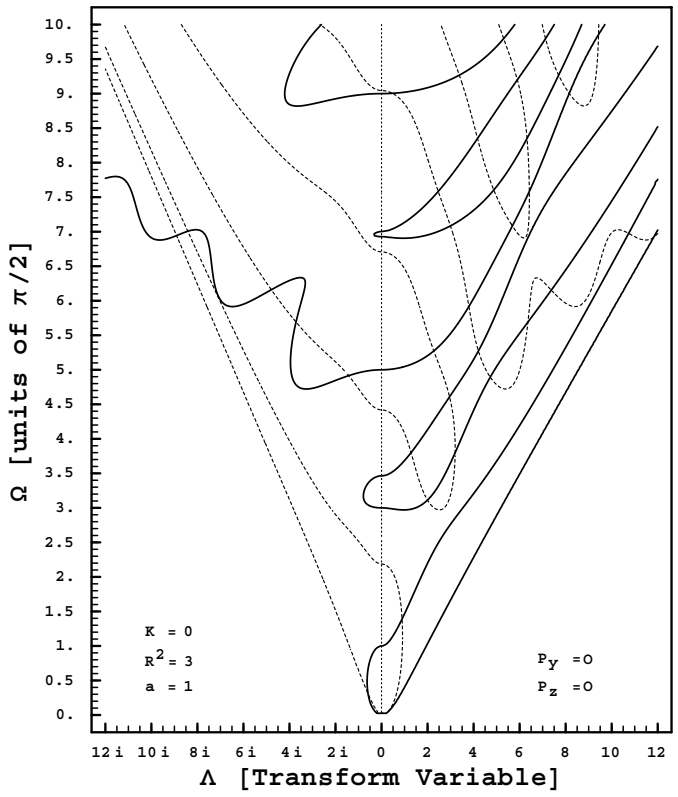

Figure 6.9: Plotted Root Families for L-Conj Torsional Cutoff Modes–Square Cross Section. Solid curves represent roots of equation (6.19) and dashed curves represent roots of equation (6.24).

Figure 6.8 shows the result of transitioning from a square cross section to one with an aspect ratio of 2:1. As the aspect ratio changes, the root system at h_z does not change since h_y is defined relative to this surface. As the aspect ratio increases, the root system associated with the h_y system has more branches in the same space and the total number of modes increases.

Figure 6.9 shows the result of substituting the Torsional parity pattern into the foregoing solutions in the case of a square cross section. Intersections of the root

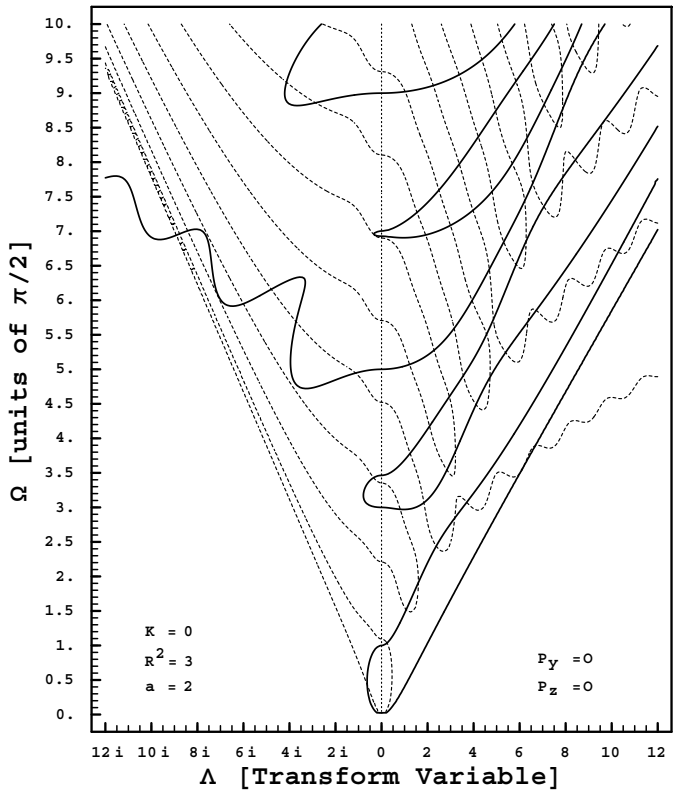

Figure 6.10: Plotted Root Families for L-Conj Torsional Cutoff Modes–1x2 Cross Section. Solid curves represent roots of equation (6.19) and dashed curves represent roots of equation (6.24).

systems occur at values of Ω which are eigenvalues of the problem. Figure 6.10 shows the result of transitioning from a square cross section to one with an aspect ratio of 2:1.

Figure 6.11 shows the result of substituting the Flexural parity pattern into the foregoing solutions in the case of a square cross section. Intersections of the root systems occur at values of Ω which are eigenvalues of the problem.The root structure of the square case is mapped in Fig. 6.11.

The flexural modes are degenerate in a square geometry. In the low frequency fundamental modes, one can imagine flexing of the $x - z$ or $x - y$ planes which are geometrically indistinguishable when the cross section is square. However, as the aspect ratio increases from unity, the flexure of those same planes is across

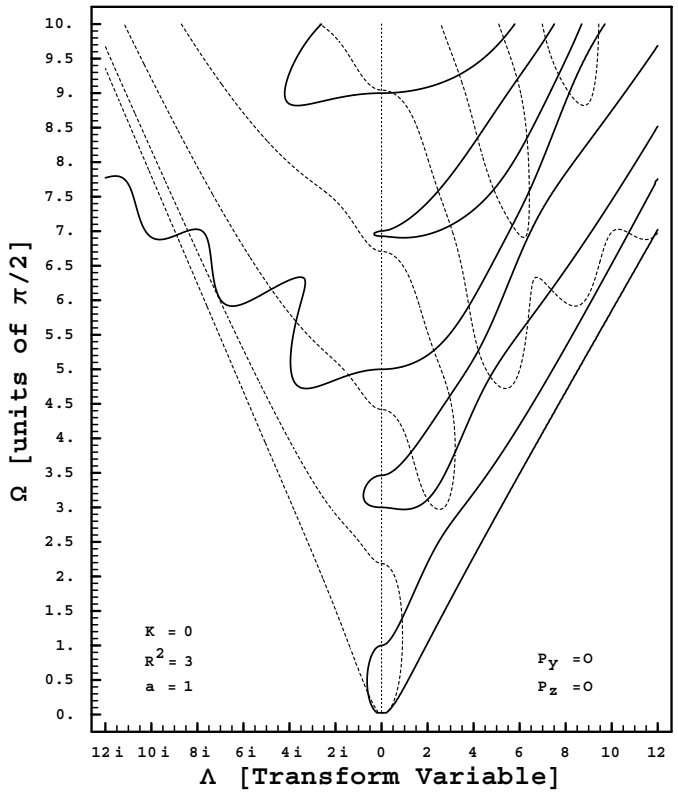

Figure 6.11: Plotted Root Families for L-Conj Flexural Cutoff Modes–Square Cross Section. Solid curves represent roots of equation (6.19) and dashed curves represent roots of equation (6.24) Transposing the P_y and P_z parity assignments will not change the mode set in the square geometry.

differing widths and juxtaposing the parity assignments of P_y and P_z will produce a different set of modes. This can be discerned by comparing Figures 6.12 and 6.13.

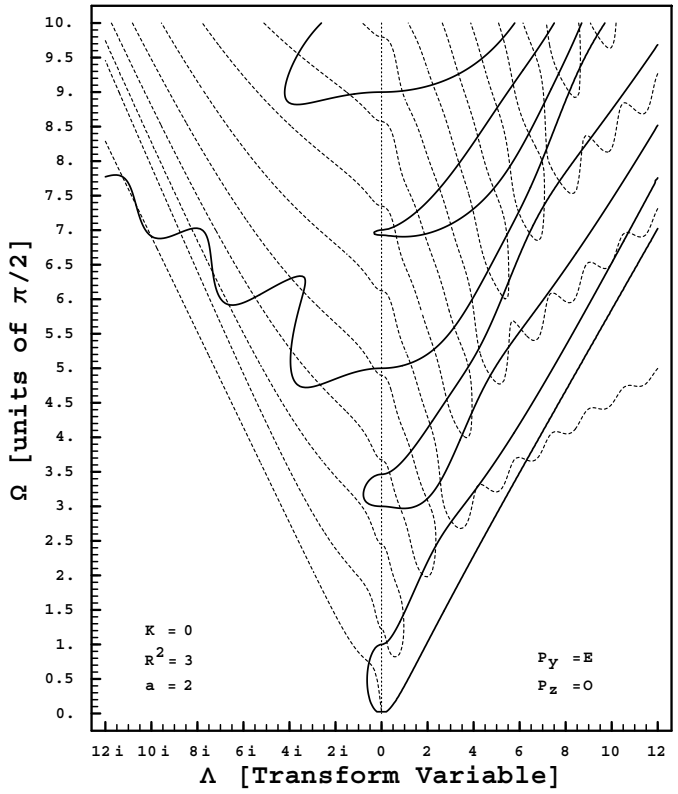

Figure 6.12: Plotted Root Families for L-Conj Flexural Cutoff Modes–1x2 Cross Section. Solid curves represent roots of equation (6.19) and dashed curves represent roots of equation (6.24). These mode families are generated by substitutions $P_y = E$ and $P_z = O$.

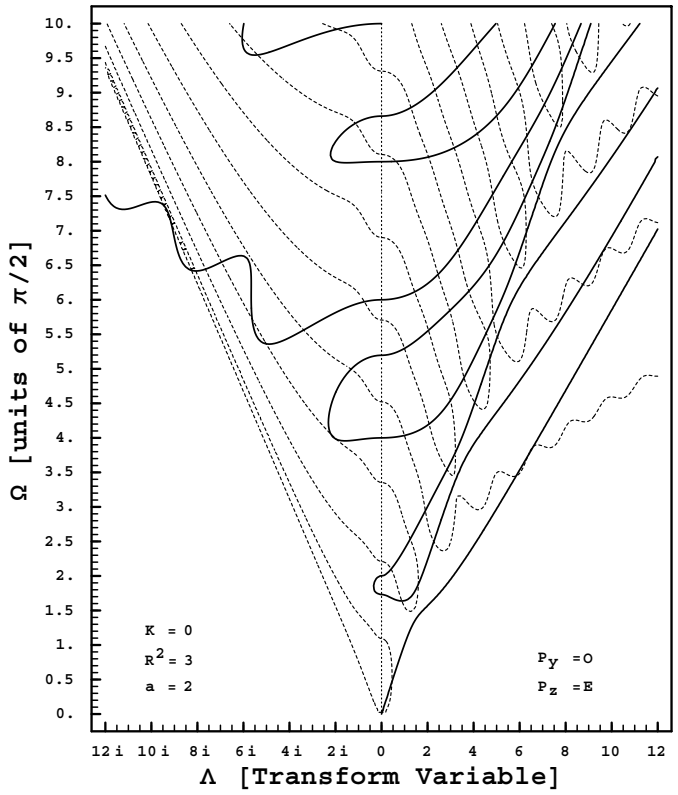

Figure 6.13: Plotted Root Families for L-Conj Flexural Cutoff Modes–1x2 Cross Section. Solid curves represent roots of equation (6.19) and dashed curves represent roots of equation (6.24). These mode families are generated by substitutions $P_y = O$ and $P_z = E$. The modes defined by root intersections are distinct from those generated by $P_y = E$, $P_z = O$.

CHAPTER 7
FRACTAL PHASE SPACE OF COUPLED MODES AT $K \to 0$

7.1 Motivation: Low Temperature Heat Conductance

Rego and Kirczenow have shown [48] that the Landauer formalism can be applied to heat transport in the ballistic regime and that in this regime thermal conductance depends upon the carrier subband structure without regard to the details of dispersion. In effect, at very low temperatures, phonon thermal conductance depends only on the cutoff modes. In addition, although each subband is a contributing channel of thermal conductance, the fundamental modes–those whose dispersions go to zero at zero k–contribute linearly in T. At the lowest temperatures, contributions other than those from the fundamental modes are exponentially suppressed. Therefore, as $T \to 0$, the overall conductance due to phonons in a mesoscopic "wire" reduces to a "quantum limit" at which conductance is computed as a constant coefficient multiplying the temperature. In 2000, Schwab, Henriksen, Worlock and Roukes published the results of an experiment [49] verifying that, at very low temperatures, thermal conductance by ballistic phonons in mesoscopic channels saturates at a value consistent with this "quantum of thermal conductance."

It is generally assumed that as the background temperature increases beyond the range of effective saturation associated with the quantum limit, that the temperature dependency of thermal conductance becomes cubic. I will show that for as much of the transition range as coupled phonons remain ballistic in

a rectangular wire, that the temperature dependency approaches a fractal power less than three.

Since the result is recent, but straightforward in its derivation, I will re derive as much of the Rego and Kirczenow result as I will need. Following the outline in reference [48], the Landauer heat flux in a wire due to ballistic phonons between reservoirs at temperatures T_L and T_R is given by

$$\dot{Q} = \frac{1}{2\pi} \sum_m \int_0^\infty dk \; \hbar\, \omega_m(k) \; [n_R(\omega_m(k)) - n_L(\omega_m(k))] \, \zeta_m(k) \, v_{g_m} \tag{7.1}$$

where $n_i(\omega_m(k))$ is the average mode occupation number at T_i given by the Planck distribution, $\zeta_m(k)$ is the transmission probability, and v_{g_m} is the energy transport velocity $(\partial\omega_m/\partial k)$. The subscript m indexes dispersion branches across all families of modes.

It is assumed that connections to the reservoir are efficient enough that I can set $\zeta_m(k) = 1$. Then equation (7.1) is converted to an integral over frequency. In converting to a frequency integral, a 1-D density of states will place $\partial\omega_m/\partial k$ in the denominator which cancels the group velocity removing any dependency on the dispersion. The result is

$$\dot{Q} = \frac{1}{2\pi} \sum_m \int_{\omega_{m_c}}^\infty d\omega \, \hbar\, \omega \; [n_R(\omega_m(k)) - n_L(\omega_m(k))] \tag{7.2}$$

where ω_{m_c} is the $k = 0$ value of ω in the mth phonon subband. Note that for each of the four fundamental modes, $\omega_{m_c} = 0$.

Since actual measurements involve inducing extremely small differences between T_L and T_R, considering only the net energy transport in the $\Delta T \to 0$ limit will suffice. In this limit, the conductance, defined as $\dot{Q}/\Delta T$, can be found by differentiating $n_i(\omega)$ with respect to T under the integral at the "average"

temperature. The result (expressed similarly in Ref. [47]) is

$$K_{\Delta T\to 0} = \frac{k_B^2\, T}{h} \sum_m \int_{x_m}^{\infty} \frac{x^2 e^x}{(e^x-1)^2}\, dx \tag{7.3}$$

where the lower bound of integration for the mth mode is given by $x_m = \frac{\hbar\omega_{m_c}}{k_B\, T}$. In this form, the conductance at a given temperature induced by very small temperature differences is determined entirely by the set of cutoff frequencies of the phonon subbands.

I have heretofore expressed cutoff frequencies in dimensionless units (see section 4.3) which absorb the physical scale of the wire and the shear velocity. In this context, however, it will be helpful to avoid confusing the symbol for half-width with Planck's constant and I will repeat the definition in terms of d as the full width of the smallest wire width.

$$\Omega_{m_c} = \frac{\omega_{m_c}}{\omega_o} \qquad \omega_o = \frac{2c_s}{d} \tag{7.4}$$

I will also rescale the temperature component of equation (7.3) to absorb physical size and material properties (other than $R = c_\ell/c_s$) by using

$$t = \frac{T}{T_o} \qquad T_o = \frac{2\hbar c_s}{k_B\, d} \tag{7.5}$$

The relationship between T_o and ω_o then has the mnemonic form

$$k_B\, T_o = \hbar\omega_o \tag{7.6}$$

It is to be noted that T_o will be close to 1K for d on the order of 200 Angstroms and c_s on the order of 10^3 m/sec. In anticipation of larger values of d, I will compute conductance over a range of 0 to 10 units of T/T_o. This does not imply

an assertion about the likelihood that phonons will remain ballistic over the same range for a given situation and I expect that the ballistic assumption breaks down primarily as a function of actual temperature. This range is chosen merely to demonstrate consistency of the result and preserve possible applicability to a wide range of physically realizable situations.

Finally, to fully isolate the universal T-linear contribution to K from those contributions which reflect the phonon subband structure, I focus on computing K/T rescaled to dimensionless units as

$$\Lambda = \frac{Kh}{k_B^2\, T} = \frac{K\pi d}{c_s\, k_B\, t} \tag{7.7}$$

so that the equation (7.3) is now rearranged into the completely dimensionless form:

$$\Lambda = \sum_m \int_{x_m}^{\infty} \frac{x^2 e^x}{(e^x - 1)^2} dx \qquad x_m = \frac{\Omega_{m_c}}{t} \tag{7.8}$$

The only material properties which survive in equation (7.8), though not expressed explicitly, are the assumption of isotropy and the ratio R of longitudinal to shear velocities which will be chosen for representative computations below to be $\sqrt{3}$.

The right hand side of equation (7.8) has all of its temperature dependence in the nonzero cutoff bounds within $\{x_m\}$. Moreover, the set of $\{\Omega_{m_c}\}$, and thus $\{x_m\}$, includes zero-valued elements corresponding to the beginning of each fundamental dispersion branch. For each such zero, there is a common-valued temperature-independent term in the sum on the right hand side of equation (7.8). For coupled modes in rectangular wires, I have shown that there are six such terms which correspond to the fundamental dispersion branches for dilatational (one), flexural modes (two), and torsional (three) mode families. Each such zero in $\{x_m\}$ contributes a "quantum of conductance" whose dimensionless value is $\pi^2/3$. [48,49].

7.2 Effective Dimension & Density of Propagating Modes as $k \to 0$

From equation (7.8), ballistic phonon thermal conductance is determined by the level spacing distribution of the full set of cutoff modes. Were it not for coupling, these levels could be enumerated as a collection of point sets each of which occupies a two-dimensional phase space with each of these characterized by two independent quantum numbers. The $k = 0$ uncoupled mode enumerations derived in section 6.2 are examples.

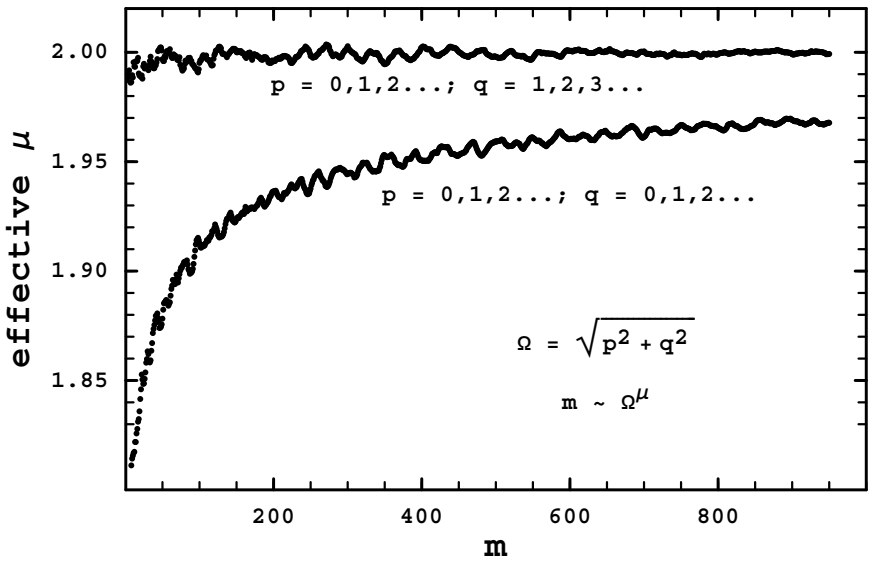

Figure 7.1: Effective capacity dimension for typical 2-D phase spaces. For a modest number of eigenvalues (m), the effective dimensionality of the phase space reveals the graininess of the discrete eigenvalues and the general trend is subject to small systematic distortion with minor changes in the details of mode enumeration. Shown is a moving average of 100 nondecreasing modes. The rule for Δm was to take the maximum of 15% of m or 15 modes.

The actual set of points defined by the $k \to 0$ limit of coupled propagating modes are not subject to a simple enumeration and good quantum numbers are not known. As prior chapters show, coupling results in a complex band structure and an unobvious level spacing distribution. Since phase space dimensionality and constancy of phase space density normally are manifest as integral power laws for

physical properties, it should be expected that the complexity of coupled modes will distort power-law behavior. I will show that dimensionality and phase space density is irregular in the lowest eigenvalue regions of coupled $k \rightarrow 0$ values and that as higher values of the sets are considered, the effective phase space dimensionality tends towards a fractal value distinctly less than two.

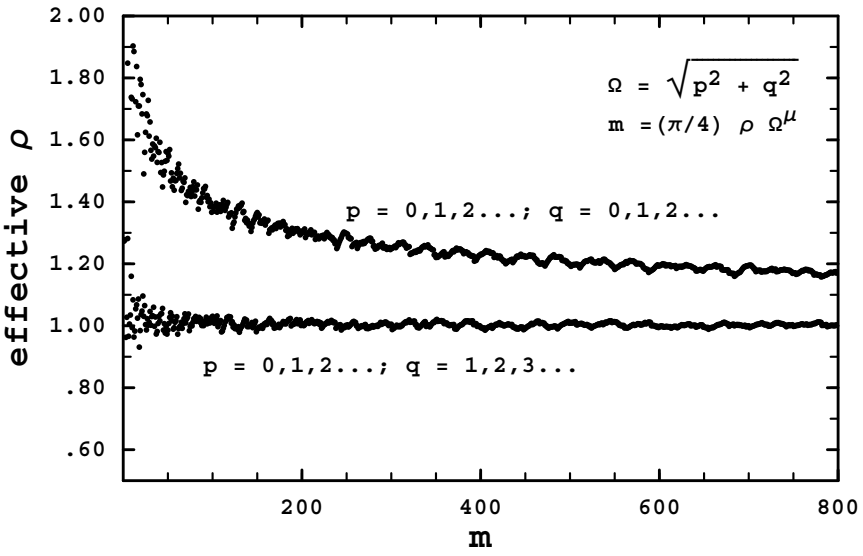

Figure 7.2: Effective density for 2-D phase spaces of Fig 7.1.

For definiteness I will focus on a square cross section. At the end of the section I will also show the effective dimension and density of states in frequency associated with an aspect ratio of 1:2 which, despite having some greater variability, retains the same general trends exhibited for the square cross section.

When the number of active modes is large and the density of modes in phase space is a constant, there is no danger in relating the number of modes m to the energy or frequency measure via a power-law relation such as $m = \rho\Omega^{\mu}$ where μ is the number of equally spaced quantum numbers that enumerate the modes. In these kinds of cases, we can simply define the dimensionality of the set of modes as the dimensionality of the phase space used to enumerate them. This is roughly

equivalent to what is also called the capacity dimension, or, sometimes, the box-counting dimension on account of a formal definition that involves counting the minimum number of boxes of a diminishing size needed to contain points in a region spanned by some measure [50]. In its physical effect, it is obviously the measure of how fast capacity increases with Ω.

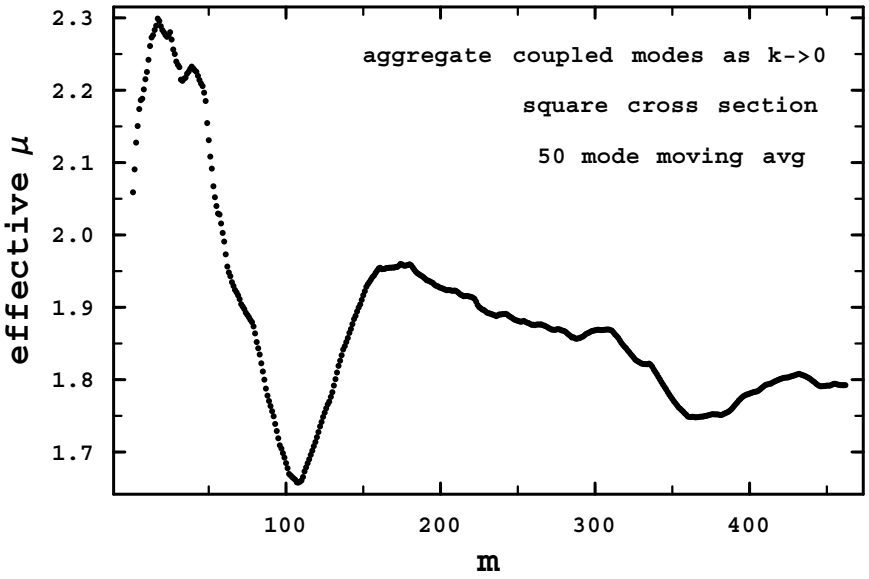

Figure 7.3: Effective capacity dimension for propagating modes at $k \to 0$. Δm was computed by taking the maximum of 50% of m or 15 modes. Compare with Fig. 7.1

If the phase space density changes or regularity is lost in the enumeration scheme, it should still be meaningful to consider the rate of change in capacity with respect to Ω. This flexibility is usually irrelevant in cases characterized by a large number of points and periodic boundary conditions that result in a constant phase space density of points. However, when the number of excited modes is small, I will demonstrate that this flexibility can easily become relevant even in phase spaces with constant density. To accommodate all situations where the effective dimensionality and density might change, I will define "effective"

dimensionality μ as a function of accumulated capacity by simply connecting the points of progressive capacity from some accumulated base m to some higher capacity level m_f such that $m = \rho_m \Omega_m^\mu$ and $m_f = \rho_{m_f} \Omega_{m_f}^\mu$ simultaneously. If I assume ρ to be effectively unchanged over Δm, and if I impose some rule for systematically computing Δm as a function of m, then μ at m can be precisely defined by

$$\mu = \frac{\log(m_f/m)}{\log(\Omega_f/\Omega)} \tag{7.9}$$

In addition, once I find μ I can define a precise $\rho(m)$ at each capacity level by

$$\rho(m) = \frac{m_f - m}{\Omega_{m_f}^\mu - \Omega_m^\mu} \tag{7.10}$$

To illustrate the meaningfulness of this flexibility and the fragility of dimension and density when the number of available modes above the ground state is small, consider the simple cases of

$$\Omega = \frac{\pi}{4}\rho\sqrt{p^2+q^2} \qquad \begin{cases} p = 0, 2, 3 \ldots; \; q = 1, 2, 3 \ldots \\ p, q = 0, 1, 2 \ldots \end{cases}$$

The phase spaces of the two cases is identical except for the removal of points along the q axis in one of them. Figure 7.1 shows a computation of the effective dimension for both cases shown as a moving average over 100 points. The modest capacities result in considerable graininess which averaging diminishes. The systematic fluctuations are a result of degeneracies over which Ω briefly stalls as m continues to increase. Figure 7.2 shows the effective ρ over the same capacity range for the two cases. I have factored the geometric factor applicable to a quarter-circle from ρ so that it has units which compare directly to a high-m value of 1.

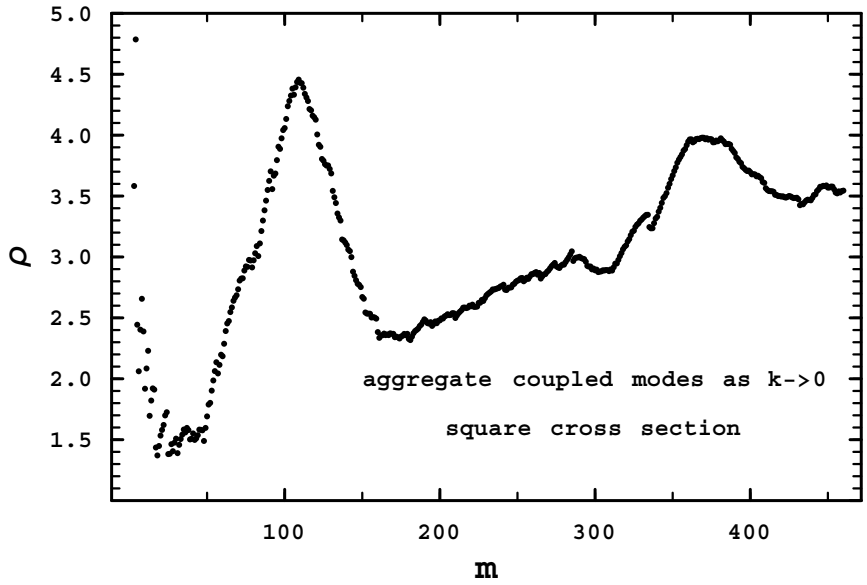

Figure 7.4: Effective phase space density of points for propagating modes at $k \rightarrow 0$. ρ is computed with respect to the effective μ's shown in Fig. 7.3. All geometric factors are absorbed into ρ. Compare with Fig. 7.2

The foregoing cases should now be compared qualitatively and quantitatively to similar computations with respect to the $k \rightarrow 0$ coupled modes for a square wire. To examine the effective dimension and density of the coupled modes in the $k \rightarrow 0$ limit, root intersections were plotted up to an Ω of $50\pi/2$ for all mode families and the results aggregated into a single set which was then sorted by eigenvalue. Since zero eigenvalues are given special treatment in the heat conductance calculations, only non-zero eigenvalues were included. Figure 7.3 shows a moving average progression of the effective capacity dimension (per equation (7.9)) of the resulting set.

In contrast with the effective μ for 2-D sets shown in Fig. 7.1, the effective μ for coupled modes in the $k \rightarrow 0$ limit shows a decreasing oscillation and the lack of systematic degeneracies results in a smoother curve albeit with larger overall

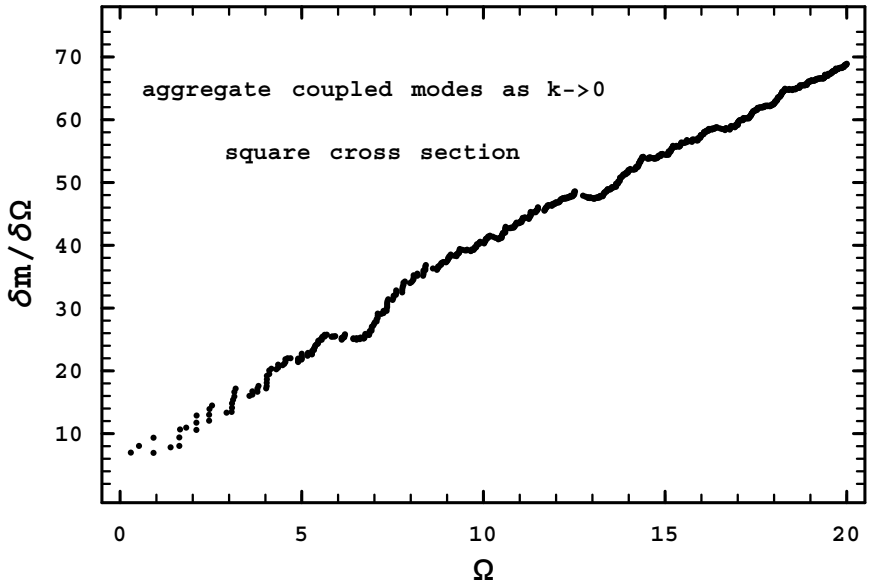

Figure 7.5: Effective density of states in frequency for propagating modes at $k \to 0$.

variation. Instead of trending up to a dimensionality of 2, the set of coupled modes appears to be trending down toward a dimensionality distinctly less than 2.

As the effective dimensionality undergoes oscillation for increasing m so too does the density. As Fig. 7.4 shows, there is, in general, a compensation for increased dimensionality in the form of decreasing phase space point density. As a result of this compensation the density of states in frequency–a dominant factor in the calculation of many physical properties–is spared the large variations evident for effective dimension and phase space density at small m. As Fig. 7.5 shows, the density of states in frequency is roughly linear albeit with a slight overall curvature. In the case of a perfect 2-D phase space with constant phase space density, the density of states in frequency would be perfectly linear.

Surprisingly, increasing the aspect ratio from 1:1 to 1:2 is not accompanied by an obvious dissipation of the dimensional reduction effect. The overall density of points rougly doubles, and the variability of effective μ increases, but as Fig. 7.6 shows, for the same general capacity level the trend is towards a dimensionality

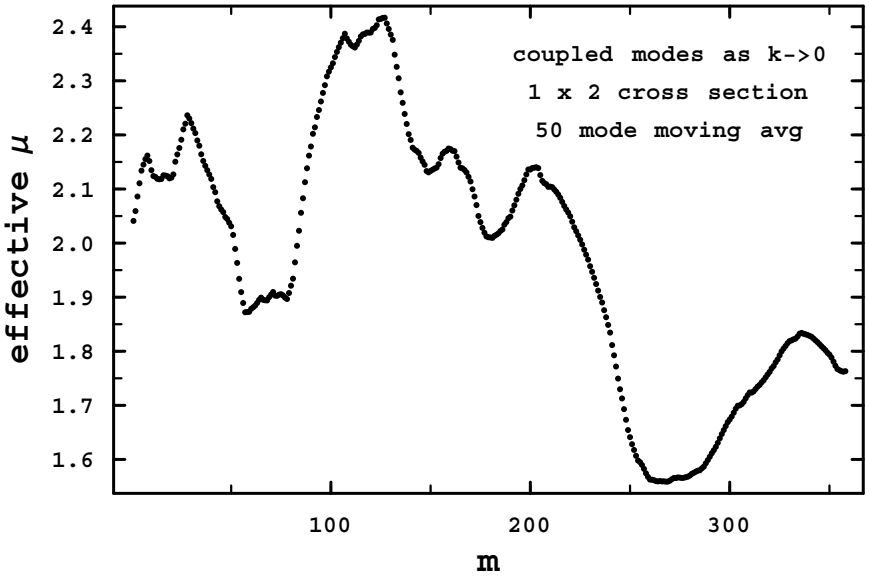

Figure 7.6: Effective phase space dimensionality for propagating modes at $k \rightarrow 0$– 1x2 cross section.

comparable to that for a square cross section. Similarly, as shown in Fig. 7.7, the density of states in frequency stilll averages towards a linear form.

7.3 Low Temperature Heat Conductance

Applying results of the prior section, I proceed now to compute the low temperature conductance due to coupled phonons in a square wire. Equation 7.8 involves an infinite sum over a function of the discrete set of modes at $k \rightarrow 0$ and in the prior section the effective capacity dimension and phase space density of the low-lying members of that set has been examined. I assume that the number of points which brings Ω to $20\pi/2$ (approximately 770 points) is sufficient to warrant the assumption that μ and ρ will not deviate significantly from their values at the highest capacities shown in Figs. 7.3 and 7.4. To facilitate a consistent method of computation, I therefore extrapolate the point set beyond the capacity

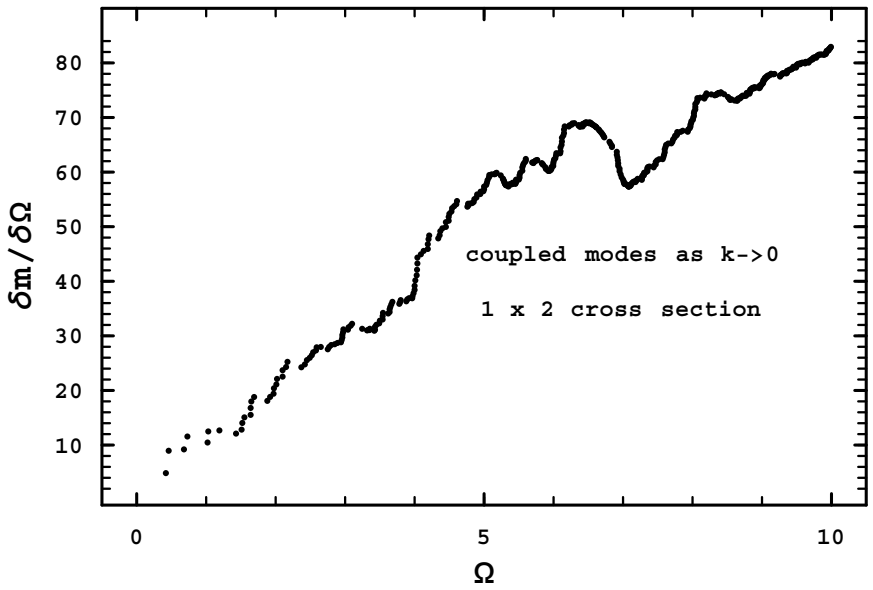

Figure 7.7: Effective density of states for propagating modes at $k \rightarrow 0$–1x2 cross section.

corresponding to $\Omega = 20\pi/2$ by adding points generated by

$$\Omega_m = (\frac{m}{\rho})^{1/\mu}$$

using $\rho = 1.593$ and $\mu = 1.793$–the density and effective dimensions of the set at $\Omega_m = 20\pi/2$. Extrapolating to $\Omega_m = 50\pi/2$ brings the total capacity of the set to approximately 3970 points.

It will be helpful to denote the dimensionless integral in equation (7.8) with the following notation:

$$F(y) = \int_y^\infty f(x)dx \qquad f(x) = \frac{x^2 e^x}{(e^x - 1)^2} \tag{7.11}$$

Λ, of course, is a sum of $F(y)$ over all possible discrete values of y contained in $\{x_m\}$. Since the extrapolation process sets an upper bound on $\{\Omega_m\}$ of $50\pi/2$, at each value of t there will be a maximum available $x_{\mathsf{max}} = (50\pi/2)/t$. Then, as a

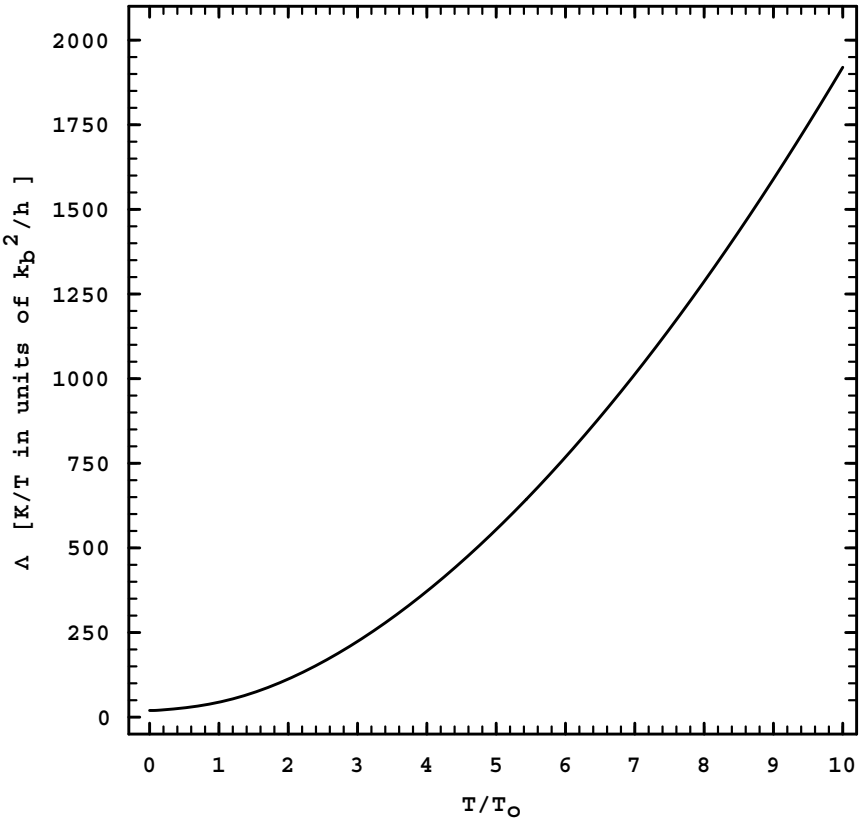

Figure 7.8: Thermal Conductance due to propagating modes at $k \to 0$.

precaution, I will contemplate computing Λ up to $t = 10$ and consider analytically estimating the tail of this sum when I reach this value for each $0 < t \leq 10$. To this end I first note that

$$\text{for } x > 5 \quad f(x) \to x^2 e^{-x} \tag{7.12}$$

so that for $y > 5$:

$$F(y) \to (2 + y(2 + y))e^{-y} \tag{7.13}$$

Assuming the x_m are in nondecreasing order, I can convert the upper range of the sum to an integral. Then, for $x_m \geq x_{\mathsf{lim}} > 5$:

$$\sum_{m=\mathsf{lim}}^{\infty} F(x_m) \to \rho\, \mu\, t^{\mu} \int_{x_{\mathsf{lim}}}^{\infty} y^{\mu-1} \left(2 + y(2 + y)\right) e^{-y}\, dy \tag{7.14}$$

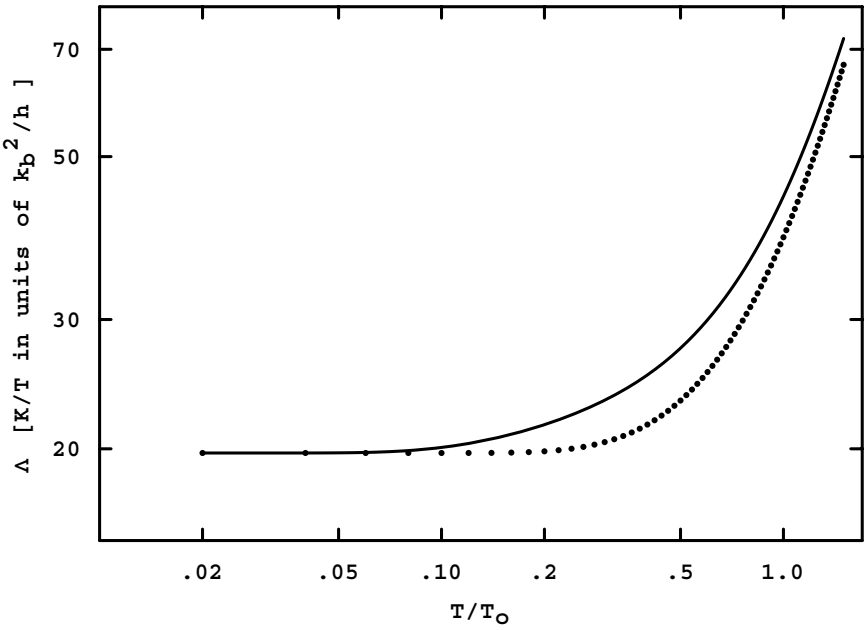

Figure 7.9: Low Temperature Region of Fig. 7.8 (Log-Log). The saturation to a constant value near the quantum limit is clear and conspicuous. The dotted line show the conductance recomputed with the first two nonzero eigenvalues removed. These values come from the discovered quadratic torsional modes. Removing them doubles the length of the quantum plateau.

For t at its largest considered value of 10, $x_{\mathsf{lim}} = (50\pi/2)/t$ is approximately 7.85 guaranteeing that approximations (7.12, 7.13) are appropriate at all lower temperatures. Evaluating equation (7.14) for $x_{\mathsf{lim}} = (50\pi/2)/t$ over the whole range $0 < t \leq 10$ shows that it is negligible below $t = 5$ and rises steeply to about 35 at $t = 10$. If I compute Λ by summing over the extrapolated set without making this correction, I discover that it is approximately 1900 as $t \to 10$ so that the tail correction is never greater than about 17 parts in 10^3. The tail correction is thus appropriately small.

Applying the tail correction while numerically computing Λ per equation (7.8) implements an algorithm characterized by the following equation:

$$\Lambda(t) = 6\frac{\pi^2}{3} + \sum_{m=1}^{\mathsf{lim}} F(x_m) \; + \rho\,\mu\,t^{\mu} \int_{x_{\mathsf{lim}}}^{\infty} y^{\mu-1}\,(2 + y(2+y))e^{-y}\,dy \qquad (7.15)$$

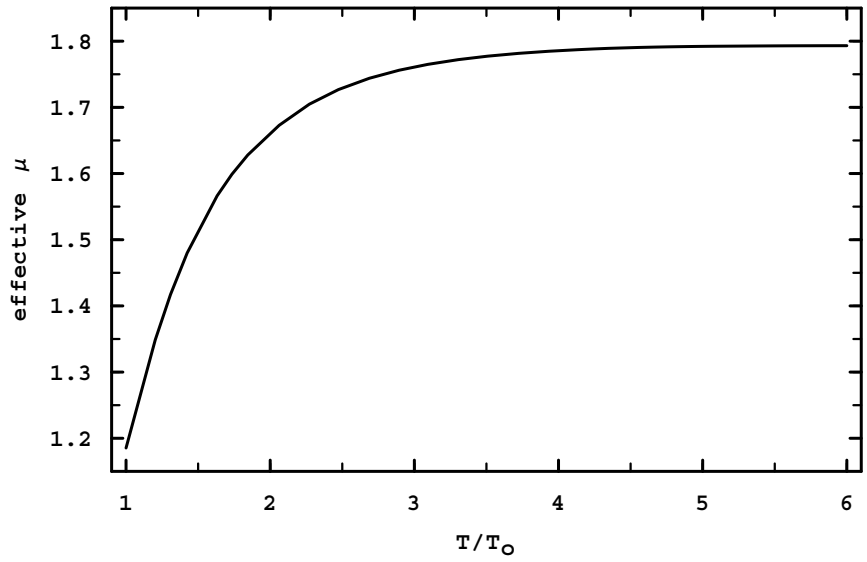

Figure 7.10: Effective Power-Law Exponent for Thermal Conductance. In the absence of scattering, the exponent for temperature dependency will be fractal and approach the dimensionality of the coupled modes in the $k \to 0$ limit.

$$\text{where} \quad x_1 = \frac{\text{lowest nonzero } \Omega_m}{t} \qquad x_{\text{lim}} = \frac{50\pi/2}{t}$$

The resulting thermal conductance, characterized by Λ, is shown in Fig. 7.8. The summation within equation (7.8) thermally smears discrete contributions to heat conductance and precludes manifestation of any stepwise progression.

Equation (7.14) incidentally shows that the effective dimension of the set of eigenvalues translates directly into the power-law temperature dependency of the thermal conductance once clear of the low temperature region. The progress of the power law for temperature dependence can be directly exhibited using the same technique used to measure effective dimensionality via equation (7.9). The result is shown in Fig. 7.10. It is, of course, based on an assumption that the propagating phonons remain ballistic over the entire range. If the wire is not uniform in its cross section or if impurity scattering is relevant, assumptions of the Landauer model will weaken and the effects of fractal dimensionality are likely to be lost.

CHAPTER 8
CONCLUSION

The main goal of this study has been to demonstrate that it is possible to analytically derive the mesoscopic phonon eigenspectrum of a rectangular isotropic elastic waveguide. The success of this enterprise should impact a range of applications and interests, but perhaps the strongest contemporary relevance comes from broad interest in very small scale structures in nano circuitry. Quantum wires are commonly rectangular and small enough to exhibit dimensional confinement of electrons. At that physical scale, however, they are also small enough that their lowest phonon frequencies are high enough to constitute a significant source of interaction with the electrons. In general, at mesoscopic scales, phonons are considered to be the most important source of electrical resistivity.

To put the importance of scale in perspective, consider equation (4.12). This simple definition of dimensionless rescaling used in the study also shows that actual frequency scales with the inverse of the smallest cross sectional half width. Typical dimensionless frequencies shown in this study are less than $10\pi/2$. Assuming, for the sake of discussion, a shear velocity on the order of 10^3 m/s (appropriate for metals), corresponding actual frequencies will be on the order of MHz for half widths of a centimeter, tens of GHz for half widths of a micrometer, and on the order of 10^{13} sec^{-1} for half widths of a nanometer. Since $k_B/\hbar$ is only approximately $1.3\text{x}10^{11}$, it is not hard to see that mesoscopic phonons in small-to-nanoscale objects have the potential for electronic interaction at quite reasonable temperatures.

Aside from the applicability to thermal and electronic properties on small scales, this study highlights generally that coupled phonons live in a fractal-dimensional

phase space. As one enters the mesoscopic regime, surface-coupled phonons eclipse bulk phonons in importance and one expectation should be that any property connected to the dimensionality of the phonon phase space will be affected. I have demonstrated this effect by calculating a fractal temperature dependency due to mesoscopic phonons in low-temperature rectangular environments. In general, power-law behaviors tied to phonon phase space dimension will experience a reduction in the applicable exponent to some non-integral number at least up to the point at which interactions erode the direct affects of the surface.

This study also demonstrates what may well be a generally helpful technique for dealing with nonseparable problems. It is the nonseparability of the boundary conditions which justified the widespread belief that the problem had no analytic solution. It is surprising, therefore, to see the problem yield to an extremely simple transform together with a carefully chosen, though ultimately quite simple, set of representations. The ability to characterize the spectrum without fully solving the superpositions is surprising–but not at all esoteric in its demonstration. In the end, it was careful attention to very elementary aspects of the problem which led to this result and that may well prove helpful in another nonseparability situation.

Finally, this study lays the groundwork for further research. It is reasonable to expect that the results of this study will be a starting point for several endeavors. These include building a thorough mathematical description of the superpositions, and extending the results to non isotropic materials.

REFERENCES

[1] N. Nishiguchi, Y. Ando, and M. N. Wybourne, Acoustic Phonon Modes of Rectangular Quantum Wires, Journal of Physics, Condensed Matter **9**, 5751 (1997).

[2] B. Auld, in *Acoustic Fields and Waves in Solids*, 1st ed. (Wiley, New York, 1973), Vol. II, Chap. 10, in which unavailability of solutions for free-standing rectangular waveguides is remarked upon specifically.The same observation does not appear in the 2d ed. because the entire subsection containing it was removed – ostensibly to make room for a new topic.

[3] R. W. Morse, The Velocity of Compressional Waves in Rods of Rectangular Cross Section, The Journal of the Acoustical Society of America **22**, 219 (1950).

[4] S. Yu, K. W. Kim, M. A. Stroscio, G. J. Iafrate, and A. Ballato, Electron–Acoustic-Phonon Scattering Rates in Rectangular Quantum Wires, Physical Review B **50**, 1733 (1994).

[5] A. I. Beltzer, *Acoustics of Solids* (Springer-Verlag, Berlin, 1988).

[6] L. Pochammer, Beitrag zur Theorie der Biegung des Kreiscylinders, Journal Für Reine Und Angewandte Mathematic (Crelle) **81**, 33 (1875).

[7] C. Chree, Longitudinal Vibrations of a Circular Bar, Quarterly Journal of Pure & Applied Mathematics **XXI**, 287 (1886).

[8] L. Rayleigh, On the Free Vibrations of an Infinite Plate, &c., Proc. London Math. Soc. Series 1 **XX**, 225 (1889).

[9] H. Lamb, On Waves in Elastic Plate, Proc. Royal Society Series A **93**, 114 (1917).

[10] J. Micklowitz, *Elastic Waves and Waveguides* (North Holland, New York, 1978).

[11] R. D. Mindlin, *The Collected Papers of Raymond D. Mindlin* (Springer-Verlag, New York, 1989), reprinted from *Structural Mechanics*, Pergamon Press (1960).

[12] A. N. Holden, Longitudinal Modes of Elastic Waves in Isotropic Cylinders and Slabs, Bell System Technical Journal **30**, 956 (1951).

[13] G. E. Hudson, Dispersion of Elastic Waves in Solid Circular Cylinders, Physical Review **63**, 46 (1943).

[14] D. Bancroft, The Velocity of Longitudinal Waves in Cylindrical Bars, Physical Review **59**, 588 (1941).

[15] S. Timoshenko, *Vibration Problems in Engineering*, 3d ed. (D. Van Nostrand, New York, 1955).

[16] M. A. Medick, One-Dimensional Theories of Wave Propagation and Vibrations in Elastic Bars of Rectangular Cross Section, Journal of Applied Mechanics **Sept 66**, 489 (1966).

[17] K. Washizu, On the Variational Principles of Elasticity and Plasticity, MIT Aeroelastic Structure Research Laboratory Technical Report 15.18, 1955.

[18] H. C. Hu, On Some Variational Vibrational Principles in the Theory of Elasticity and the Theory of Plasticity, Acta Physica Sinica **10**, 259 (1954).

[19] L. Rayleigh, in *The Theory of Sound* (Macmillan, London, 1877), Vol. i.

[20] E. T. Whittaker, in *A Treatise on the Analytical Dynamics of Particle and Rigid Bodies*, 4th ed. (University Press, Cambridge, 1937).

[21] R. Holland, Resonant Properties of Piezoelectric Ceramic Rectangular Parallelepipeds, The Journal of the Acoustical Society of America **43**, 988 (1968).

[22] R. Holland and E. P. EerNisse, *Design of Resonant Piezoelectric Devices* (MIT Press, Cambridge, Mass., 1969).

[23] H. H. Demarest, Jr., Cube-Resonance Method to Determine the Elastic Constants of Solids, The Journal of the Acoustical Society of America **49**, 768 (1971).

[24] G. Arfken, *Mathematical Methods for Physicists*, 3d ed. (Academic Press, San Diego, 1985).

[25] W. M. Visscher, A. Migliori, and T. M. Bell, On the Normal Modes of Free Vibrations of Inhomogeneous and Anisotropic Elastic Objects, Journal of the Acoustical Society of America **90**, 2154 (1991).

[26] A. Migliori, *Resonant Ultrasound Spectroscopy* (Wiley, New York, 1997).

[27] V. B. Campos, S. Das Sarma, and M. A. Stroscio, Phonon-Confinement Effect on Electron Energy Loss in One Dimensional Quantum Wires, Physical Review B **46**, 3849 (1992).

[28] R. Mickevičius, V. V. Mitin, K. W. Kim, M. A. Stroscio, and G. J. Iafrate, Electron Intersubband Scattering by Confined and Localized Phonons in Real Quantum Wires, Journal of Physics; Condensed Matter **4**, 4959 (1992).

[29] M. A. Stroscio and M. Dutta, *Phonons in Nanostructures* (Cambridge University Press, Cambridge UK, 2001).

[30] S. Das Sarma and V. B. Campos, Low-temperature Thermal Relaxation of Electrons in One-dimensional Nanometer-size Structures, Physical Review B **47**, 3728 (1993).

[31] S. Das Sarma and V. B. Campos, LO-phonon Emissions by Hot Electrons in One-Dimensional Semiconductor Quantum Wires, Physical Review B **49**, 1867 (1994).

[32] J. Seyler and M. N. Wybourne, Acoustic Waveguide Modes Observed in Electrically Heated Metal Wires, Physical Review Letters **69**, 1427 (1992).

[33] J. R. Senna and S. Das Sarma, Giant Many-Body Enhancement of Low Temperature Thermal-Electron-Acoustic-Phonon Coupling in Semiconductor Quantum Wires, Physical Review Letters **70**, 2593 (1993).

[34] R. Mickevičius, V. Mitin, and V. Kochelap, Radiation of Acoustic Phonons from Quantum Wires, Journal of Applied Physics **77**, 5095 (1995).

[35] G. D. Sanders, C. J. Stanton, and Y. C. Chang, Theory of Transport in Silicon Quantum Wires, Physical Review B **48**, 11067 (1993).

[36] R. W. Morse, The Dispersion of Compressional Waves in Isotropic Rods of Rectangular Cross Section, Ph.D. thesis, Brown University, Providence, R.I., 1949.

[37] B. Auld, *Acoustic Fields and Waves in Solids*, 1st ed. (Wiley, New York, 1973), Vol. I & II.

[38] B. Auld, *Acoustic Fields and Waves in Solids*, 2nd ed. (Krieger Publishing Company, Melbourne, Florida, 1990), Vol. I & II.

[39] K. W. Kim, S. Yu, M. U. Erdoğan, M. A. Stroscio, and G. Iafrate, Acoustic Phonons in Rectangular Quantum Wires: Approximate Compressional Modes and the Corresponding Deformation Potential Interactions, Proceedings of SPIE **2142**, 77 (1994).

[40] C. Herring and E. Vogt, Transport and Deformation-Potential Theory for Many-Valley Semiconductors with Anisotropic Scattering, Physical Review **101**, 944 (1956).

[41] G. D. Mahan, Temperature Dependence of the Band Gap in CdTe, Journal of Phys. Chem. Solids **26**, 751 (1964).

[42] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity*, Vol. 7 of *Course of Theoretical Physics* (Wiley, New York, 1970).

[43] E. H. Love, *A Treatise on the Mathematical Theory of Elasticity*, 4th ed. (Dover, New York, 1944), reprint of 1927 Cambridge publication *Mathematical Theory of Elasticity*.

[44] F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Dover, New York, 1990), first published by Frederick Ungar Publishing Company, New York, 1955.

[45] H. Sagan, *Boundary and Eigenvalue Problems in Mathematical Physics* (Dover, New York, 1989 reprint of 1961 Wiley, New York publication), see Theorem VII.8, section 3.3.

[46] R. Peierls, *Surprises in Theoretical Physics* (Princeton University Press, New Jersey, 1979), section 3.5.

[47] D. E. Angelescu, M. C. Cross, and M. L. Roukes, Heat Transport in Mesoscopic Systems, Superlattices and Microstructures **23**, 673 (1998).

[48] L. Rego and G. Kirczenow, Quantized Thermal Conductance of Dielectric Quantum Wires, Physical Review Letters **81**, 232 (1998).

[49] K. Schwab, E. Henriksen, J. Worlock, and M. Roukes, Measurement of the Quantum of Thermal Conductance, Nature **404**, 974 (2000).

[50] B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1982), revised ed. of 1977 Title.

BIOGRAPHICAL SKETCH

Steven E. Patamia was born in 1947 in New Orleans, Louisiana, but grew up as a military dependent. Consequently he has lived in England, the far East, and various "stateside" locations. He completed high school in Las Vegas, Nevada while his father was stationed at Nellis AFB outside of Las Vegas.

Steven has had strong scientific and technical interests since childhood. Among the manifestations of this interest, he was an active "amateur rocketeer" in the 1960's and had two rockets fired at the Army test range at Fort Sill, Oklahoma. Upon graduation from high school, Steven was the local recipient of the Bausch and Lomb Science Award given to recognize students with the most scientific promise.

Steven began undergraduate studies while still in Nevada, but transferred to UCLA at the same time his father was assigned to duty in Vietnam. Eventually, he rejoined his family in Tampa, Florida, where his father eventually retired from the Air Force. He completed his undergraduate degree in mathematics from the University of South Florida in 1972. Completion of that degree was interrupted by military service during which he was assigned to a Defense Intelligence Agency facility in Hawaii.

Prior to being drafted into military service in 1970, and while working on his BA, Steven was a director of systems programming at what was then known as the Computer Research Center at the University of South Florida. He remained professionally involved in commercial software development and support over a period of approximately 30 years.

Following military service, Steven pursued flight training and ultimately became an active general aviation instrument flight instructor.

Steven attended the University of Florida Law School and received a Juris Doctor degree in 1977. He practiced commercial law in the Miami area, but eventually concentrated on various consulting ventures. Out of a sense of frustration at not having fulfilled his early scientific ambitions, Steven eventually applied to and was admitted into the PhD program in physics at the University of Florida. Research on his PhD topic was completed at the Los Alamos National Laboratory in New Mexico and Steven hopes to be able to complete further development of the topic at that location.