Design and Modeling of a PZT Thin Film Based Piezoelectric Micromachined Ultrasonic Transducer (PMUT)

by

Katherine Marie Smyth

B.S., Massachusetts Institute of Technology (2010)

Submitted to the Department of Mechanical Engineering in partial fulfillment of the requirements for the degree of Master of Science in Mechanical Engineering at the MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2012

© Massachusetts Institute of Technology 2012. All rights reserved.

Author.................................................................
Department of Mechanical Engineering
May 11, 2012

Certified by ........................................ Sang-Gook Kim
Professor of Mechanical Engineering
Thesis Supervisor

Accepted by.................................................................
David E. Hardt
Chairman, Department Committee on Graduate Theses
Design and Modeling of a PZT Thin Film Based Piezoelectric Micromachined Ultrasonic Transducer (PMUT)

by

Katherine Marie Smyth

Submitted to the Department of Mechanical Engineering on May 11, 2012, in partial fulfillment of the requirements for the degree of Master of Science in Mechanical Engineering

Abstract

The design and modelling framework for a piezoelectric micromachined ultrasonic transducer (PMUT) based on the piezoelectric thin film deposition of lead zirconate titanate (PZT) is defined. Through high frequency vibration (1-16MHz) of a thin plate, the PMUT transmits and receives pressure pulses to construct medical ultrasound images as an alternative to bulk piezoelectric transducers currently in use. Existing transducers are difficult to fabricate and lack the small scale necessary for small form factor, high resolution 2D imaging arrays. From acoustic principles, the potential PMUT acoustic pressure output is determined and compared to a radiating rigid piston model. Acoustic pressure is shown to scale with the volumetric displacement rate, which is related to the plate deflection. A Green's function approach is then used to explicitly solve for the plate deflection of a bimorph and unimorph PMUT with an arbitrary number of circular or ring electrodes. The resulting solution is much simpler and more flexible than previously published solution techniques enabling the optimization of electrode configuration for large deflection and acoustic pressure. Additionally, the contribution of residual stress is examined; particularly its effects on bandwidth, sensitivity, and resonant frequency and an appropriate electrode coverage of the PMUT plate is suggested. Based on modelling, an initial PMUT design is proposed and is currently being fabricated.

Thesis Supervisor: Sang-Gook Kim
Title: Professor of Mechanical Engineering
Acknowledgments

There are many people without whom my master’s degree and this thesis would not have been possible. First, I would like to thank my advisor Professor Kim for his guidance throughout the past year. His big picture insight and industry background has given me a practical perspective in addition to his technical knowledge, which has taught me a great deal about design and research. I look forward to continuing to work with and learn from him in the future.

I would also like to thank my labmates Heonju Lee, Steve Bathurst and Ray Xu, and our Masdar collaborator, visiting professor Firas Sammoura. I truly appreciate all of the helpful suggestions and guidance that they have provided regarding fabrication and research. As a senior member of the lab, I am especially greatful for Steve’s help - he is always available to answer questions and offers to aid in fabrication and other challenges. I would not have been able to make it as far as I have on the project without his insight. Getting started on a project and finishing a Master’s thesis in one year is not an easy feat. The initial groundwork performed by Firas and his guidance early on and throughout this year on the PMUT project has been invaluable. I am very greatful for this positive collaboration and support. I would also like to thank Ray Hardin for his administrative work.

The support of the Microsystems Technology Laboratories staff has also been invaluable so far on this research effort. In particular, Dennis Ward has taught me almost all of what I know about practical fabrication and getting things to work using tricks in the MTL. I am also grateful for the training I have received from Dave Terry, Eric Lim, Bernard Alamariu, and Paudely Zamora. I would like to thank Vicky Diadiuk for giving me a chance to start fabrication work early in the MTL based on my sparse fabrication knowledge and faith. Without this added time, the current fabrication progress would not have been possible.

Finally, I would like to thank my mother, father, my brother Jason, and friends for their emotional support. To my mother and father, I cannot imagine growing up in a more supportive household and I am grateful for all of your love. When I am stressed, happy or unhappy, I always feel comfortable talking to both of you and you never manage to miss an
opportunity to make me feel better.
# Contents

## 1 Introduction

1.1 Research Objective .................................. 28

## 2 Background ........................................ 31

2.1 Medical Ultrasound Imaging .......................... 31
  2.1.1 History ......................................... 31
  2.1.2 Basic Ultrasound Physics and Imaging .......... 34
  2.1.3 Trends toward 3D Imaging ...................... 41
  2.1.4 Bulk PZT Transducer Fabrication .............. 43
  2.1.5 Transducer Developments ...................... 45

2.2 Capacitive Micromachined Ultrasonic Transducers (CMUTs) ........ 47
  2.2.1 History and Overview .......................... 47
  2.2.2 Fabrication ................................... 48
  2.2.3 CMUT Modelling ............................... 51
  2.2.4 Recent Innovations ........................... 53
  2.2.5 Limitations .................................. 54

2.3 Piezoelectric Micromachined Ultrasonic Transducers (PMUTs) .... 63
  2.3.1 Overview of Previous PMUTs .................. 63
  2.3.2 Fabrication ................................... 67
  2.3.3 Device Characterization ...................... 69
  2.3.4 PMUT Potential for Innovation ............... 73
### 3 Design and Modelling

3.1 Design Framework .......................................................... 77

- 3.1.1 Determination of Design Parameters ................................. 81

3.2 PMUT Plate Modeling ............................................................ 94

- 3.2.1 Classic Plate Theory [48] ............................................. 94
- 3.2.2 Homogeneous Solution .................................................... 100
- 3.2.3 PMUT Clamped Plate Resonance Example .......................... 103

3.3 Ultrasonic Acoustics ............................................................. 106

- 3.3.1 Wave Equation [43] ....................................................... 106
- 3.3.2 Acoustic Pressure Output of Vibrating Circular Plate [43] .... 109
- 3.3.3 Numerical Solution of Acoustic Pressure ............................ 110
- 3.3.4 Acoustic Pressure Comparison between Potential PMUT Designs 112

3.4 Green’s Function Solution ...................................................... 114

- 3.4.1 Derivation of Governing Plate Vibration Equation ................. 114
- 3.4.2 Bimorph Vibration Equation ........................................... 117
- 3.4.3 Unimorph Vibration Equation ........................................... 118
- 3.4.4 Arbitrary Electrode Configuration .................................... 120
- 3.4.5 Definition of Green’s Function ....................................... 122
- 3.4.6 Arbitrary Electrode Green’s Function Solution .................... 124

### 4 Results

4.1 Model Validation ............................................................... 129

- 4.1.1 Acoustic Pressure ......................................................... 129
- 4.1.2 Deflection Model Verification ....................................... 130

4.2 Electrode Configuration Optimization .................................... 130

### 5 Fabrication

5.1 Current Progress ............................................................... 135

- 5.1.1 Initial Processing Steps .................................................. 136
5.1.2 Bottom Electrode Lift-Off ........................................... 137
5.1.3 PZT Processing ......................................................... 139
5.2 Next steps ................................................................. 142

6 Summary ................................................................................. 145
6.1 Conclusions ................................................................. 145
6.2 Future Work ................................................................. 147
List of Figures

2-1 Array types used for imaging different parts of the body [30] ............... 36
2-2 Diagram of 1D transducer slicing. Elements have pitch $p$, width $w$ and length $l$ [56]. ................................................................. 45
2-3 Ultrasonic transducer equivalent circuit model based on lossless system with deflection in plate mode [32] .................................................. 55
2-4 Cross-section diagram of CMUT [11] .................................................. 56
2-5 Bandwidth as a function of plate radius for various membrane thicknesses. ... 57
2-6 Pull-in voltage related to gap height of a 500nm thick plate for different plate radii. Dielectric breakdown voltage plotted based on the dielectric strength of silicon nitride. ......................................................... 59
2-7 Transformation ratio as a function of gap height for various plate thicknesses. From Equation (2.13) and the definition of $n$ in 2.2, the transformation ratio is independent of radius near the pull-in voltage. ................................. 60
2-8 Relations between non-dimensionalized values of stiffness $k_{eff}/k$, restoring force $k_{eff}/k (1 - \zeta)$, and displacement $\zeta = 1 - x/t_a$. Negative values based on the sign convention of the deflection. Maximum values described above based on absolute value. ................................................................. 62
2-9 Cross-section of typical PMUT device design ........................................ 65
3-1 PMUT functional hierarchy ................................................................. 78
3-2 PMUT design parameter hierarchy ...................................................... 79
3-3 Unimorph stack with \( n \) with \( i^{th} \) layer having thickness \( h_i \) at a distance \( z_i \) from the reference \( z_0 = h_0 = 0 \). 

3-4 Values of \( \gamma^2 \) for various boundary conditions [10]. Boundary condition abbreviations specified in Table 3.1. 

3-5 Fundamental frequency based on loaded and unloaded conditions and resonant frequency. Values of \( \gamma_{01}^2 \) taken from [10]. 

3-6 Quality near resonance for \( f < f_{01,load} \) based on residual stress. Values shown are for a single layer silicon plate with \( a = 50\mu m \) and \( h_{Si} = 4\mu m \). 

3-7 Mechanical impedance based on residual stress with 3-dB bandwidth indicated by gray lines. Values shown are for a single layer silicon plate with \( a = 50\mu m \) and \( h_{Si} = 4\mu m \). 

3-8 Schematics showing coordinate system and geometry of plate displacement. 

3-9 Forces and moments acting on an infinitesimal volume element. 

3-10 Various vibration modes \( (m, n) \) of clamped, circular plate [7]. 

3-11 Non-dimensionalized first mode deflection profile. 

3-12 Optimal plate dimensions for various resonant frequencies. 

3-13 Point source geometry of acoustic radiation from a rigid plane boundary. 

3-14 Non-dimensional deflection profile (black) and circular arc fit of deflection profile (blue). 

3-15 Acoustic Pressure scaling for circular plate with radius \( a = 50\mu m \) resonating at 3MHz with plate thickness set by 3.51. 

3-16 Basic schematic of PMUT design configurations. 

3-17 Unimorph PMUT with \( N \) circular/ring electrodes in \( d_{31} \) mode. 

4-1 Schematics of representative unimorph PMUT test case with specified dimensions and materials.
4-2 Plots of PMUT deflection for geometry pictured in Figure 4-1 for 1st mode and the first 15 axisymmetric modes (n=15 in equation 3.132). Additional modes added to energy-based derivation by summation similar to that in equation 3.132.

4-3 Maximized deflection based on electrode geometry determined through Mat-
lab SQP optimization.

4-4 Deflection profile with one electrode for maximized deflection and acoustic pressure.

5-1 Microscope images of top side alignment marks.

5-2 Image reversal resist profile and lift-off.

5-3 Microscope pictures of spin coated PZT after pyrolysis.

5-4 Square root intensity scaling of XRD raw data with superimposed platinum and perovskite PZT peaks. Numbers on peaks indicate crystal orientation of PZT.

5-5 Peak matching data with measured XRD peaks on top and perovskite PZT and platinum peaks underneath. Perovskite PZT and platinum match all of the recorded peaks in the XRD data.

5-6 Cross-section views of PMUT device throughout fabrication process. Starting from top left and proceeding left to right down the page: oxide growth, bottom electrode lift-off, PZT after wet etch, top electrode lift-off, and back side DRIE144.
List of Tables

2.1 Acoustic impedance of water and in various parts of the body [30] ........... 38
2.2 Approximate equivalent circuit values for CMUT with material properties:
  dielectric permittivity of gap $\epsilon_0$ and membrane $\epsilon_r$, Youngs modulus $Y$, and
  poisson's ratio $\nu$ [32], [38], [11] .......................................................... 56
2.3 Resonant frequency equations and error between analytical and experimental solution [29], [44], [53], [62], [18], [60], [16]. Variables defined as follows:
  flexural rigidity $D$, tension $T$, plate radius/square side length $a$, first mode
  circular constant $\alpha$, first mode square constant $\beta$, membrane mode $m$, plate
  mode $p$, plate density $\rho$, plate thickness $h$ ........................................ 70
2.4 Coupling coefficients for various PMUT designs [44], [5], [16], [60] [53] ...... 71
3.1 Abbreviations for boundary conditions shown in Figure 3-4 ................. 86
5.1 Fabrication Process steps. Steps above the bold line have successfully been
  completed. ................................................................. 143
Nomenclature

Latin Letters

\( A \) constant related to average or maximum plate displacement
\( a \) plate radius
\( a' \) inner radius of electrode
\( a'' \) outer radius of electrode
\( B \) constant in plate deflection equation
\( c^E \) compliance
\( C_0 \) parasitic capacitance
\( C_m \) plate stiffness capacitance
\( c_m \) speed of sound in imaging medium
\( D \) flexural rigidity when referring to plate bending, electrical displacement in piezoelectric constitutive relations
\( d_{31} \) transverse piezoelectric constant
\( d_{33} \) longitudinal piezoelectric constant
\( E \) electric field
\( E_L \) electrical loss
\( F_e \) electrostatic force
\( f \) frequency
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_a$</td>
<td>anti-resonant frequency</td>
</tr>
<tr>
<td>$f_d$</td>
<td>doppler frequency shift</td>
</tr>
<tr>
<td>$f_r$</td>
<td>resonant frequency</td>
</tr>
<tr>
<td>$f_t$</td>
<td>transmit frequency</td>
</tr>
<tr>
<td>$h$</td>
<td>plate material thickness</td>
</tr>
<tr>
<td>$I$</td>
<td>acoustic intensity</td>
</tr>
<tr>
<td>$I_0$</td>
<td>moment of inertia</td>
</tr>
<tr>
<td>$I_2$</td>
<td>second moment of inertia</td>
</tr>
<tr>
<td>$IL$</td>
<td>insertion loss</td>
</tr>
<tr>
<td>$K$</td>
<td>kinetic energy</td>
</tr>
<tr>
<td>$k$</td>
<td>wave number</td>
</tr>
<tr>
<td>$k_{1,2}$</td>
<td>constants in stressed plate mechanical impedance equation</td>
</tr>
<tr>
<td>$k_c$</td>
<td>coupling coefficient</td>
</tr>
<tr>
<td>$k_m$</td>
<td>plate stiffness</td>
</tr>
<tr>
<td>$k_{eff}$</td>
<td>effective stiffness including spring softening</td>
</tr>
<tr>
<td>$L$</td>
<td>distance between source $R_0$ and target $R - R_0$</td>
</tr>
<tr>
<td>$l$</td>
<td>element length</td>
</tr>
<tr>
<td>$L'$</td>
<td>distance between source $R'_0$ and target $R - R'_0$</td>
</tr>
<tr>
<td>$M$</td>
<td>bending moment</td>
</tr>
<tr>
<td>$m_a$</td>
<td>added mass</td>
</tr>
<tr>
<td>$m_p$</td>
<td>plate mass</td>
</tr>
<tr>
<td>$N$</td>
<td>normal force in plate bending, number of circular/ring electrodes in Green’s function derivation</td>
</tr>
<tr>
<td>$p$</td>
<td>acoustic pressure, element pitch when referring to transducers</td>
</tr>
<tr>
<td>$p_f$</td>
<td>frequency parameter</td>
</tr>
</tbody>
</table>
$p_\omega$ magnitude of harmonic pressure independent of time

$Q$ quality factor in Chapter 2, axial stiffness coefficient in Chapter 3

$q$ multiplicative pressure scaling factor

$R$ distance from source used to calculate acoustic pressure

$R'_0$ point source location symmetric to $R_0$ about the z-axis

$R_0$ point source location

$r_0, \theta_0$ coordinate system along plate used in Green's function derivation

$R_a$ acoustic resistance

$R_m$ mechanical resistance

$s^E$ stiffness

$n$ transformer ratio in Chapter 2, number of material layer in Chapter 3

$T$ membrane tension

$t$ time

$t_a$ gap height

$t_n$ membrane thickness

$TL$ transmission loss

$(u_r, u_\theta, u_z)$ plate displacement field

$U_e$ parasitic capacitance energy

$u_z$ time dependent plate displacement velocity

$u_\omega$ harmonic plate displacement velocity independent of time

$U_{ac}$ acoustic energy
$V$  voltage
$v$  blood flow velocity
$V_{PI}$  pull-in voltage
$W$  acoustic power in Chapter 2, potential energy in Chapter 3
$w$  element width in Chapter 2, plate displacement in Chapter 3
$x$  CMUT displacement
$X_a$  acoustic reactance
$X_m$  mechanical reactance
$Y$  Young’s modulus
$z$  height of material layer with respect to bottom of stack
$z_a$  general acoustic impedance
$Z_m$  stressed plate mechanical impedance
$z_N$  neutral axis height measured from plate bottom
$z_{ml}$  acoustic impedance of matching layer
$z_{tis}$  acoustic impedance of tissue
$z_{trans}$  acoustic impedance of transducer

Abbreviations

1D, 2D, 3D  one-, two-, three-dimensional
AIUM  American Institute of Ultrasound in Medicine
AIN  aluminum nitride
B-mode  brightness mode imaging
BOE  buffered oxide etchant
CMOS  complementary metal oxide semiconductor
CMUT  capacitive micromachined ultrasonic transducer
CT    computed tomography
DP    design parameter
DRIE  deep reactive ion etch
FR    functional requirement
HF    hydrofluoric acid
KOH   potassium hydroxide
LPCVD low pressure chemical vapor deposition
LTO   low temperature oxide
MRI   magnetic resonance imaging
PECVD plasma enhanced chemical vapor deposition
PMUT  piezoelectric micromachined ultrasonic transducer
POSFET PVDF metal oxide semiconductor field effect transistor
PVDF  polyvinylidene difluoride
PZT   lead zirconate titanate
SOI   silicon-on-insulator
SQP   sequential quadratic programming
ZnO   zinc oxide

Greek Letters

\( \alpha \) frequency constant for first vibration mode of circular plate
\( \beta \) vibration constant: for first vibration mode of square plate in Chapter 2, clamped circular plate in Chapter 3

\( \chi(R) \) particular solution for green's function

\( \nu \) non-dimensionalized residual stress

\( \Delta \omega \) angular frequency bandwidth

\( \Delta f \) bandwidth

\( \varepsilon^e \) dielectric constant measured at constant strain

\( \varepsilon_0 \) dielectric permittivity of free space

\( \varepsilon_r \) relative dielectric permittivity of membrane

\( \eta \) transmit energy efficiency

\( (\varepsilon_{rr}, \varepsilon_{\theta\theta}, \varepsilon_{zz}) \) plate longitudinal strain field

\( \gamma \) constant from plate vibration equation

\( (\gamma_{r\theta}, \gamma_{z\theta}, \gamma_{rz}) \) plate transverse strain field

\( \kappa \) numerical constant for critical buckling stress

\( \lambda \) wavelength

\( \nu \) poisson's ratio

\( \Omega \) plate surface mid-plane

\( \Psi \) characteristic shape profile for vibrating clamped, circular plate

\( \rho \) plate material density

\( \rho_m \) density of imaging medium

\( \sigma_o \) residual stress

\( \varphi \) doppler imaging angle

\( \omega \) angular frequency

\( \zeta \) non-dimensionalized membrane displacement
Superscripts

' unimorph or arbitrary layer stack
(0) membrane strain contribution
(1) flexural strain or curvature contribution
* critical value
σ residual stress

Subscripts

01 first or fundamental vibration mode
e shape profile dependent on \( \cos (m\theta) \) or piezoelectric constant
i stack layer
k electrode
m membrane when referring to previous PMUT resonant frequencies otherwise radial vibration mode
N neutral axis
n vibration mode in \( \theta \) direction
o shape profile dependent on \( \sin (m\theta) \)
p plate when referring to previous PMUT resonant frequencies otherwise piezoelectric induced
σ related to \( \theta \) dependence of shape profile, stress when referring to plate bending

23
Chapter 1

Introduction

Considering its relatively brief history, medical ultrasound imaging has made staggering advances in the past 50 years. Starting as an idea inspired by radar and sonar from the military, the first interest in ultrasound’s use for therapeutic applications was reportedly in the early 1950’s [2]. Later in the decade, one-dimensional ultrasound diagnostic tools had already been developed and showed the ability to monitor heart valves and detect stomach cancer. It was not long into the next decade before the first two-dimensional bright mode (B mode) imaging, which forms the basis for many modern medical ultrasound imaging techniques, had been demonstrated and the diagnostic applications further broadened. Conventional ultrasound uses a short ultrasonic pulse (usually a few MHz) to transmit waves into the imaging medium. The incident pulse reflects off features in the human body and the timing between the incident pulse and the one received by the transducer determines the feature depth. In brightness mode imaging, the relative intensity of the received signal determines the image brightness and a two-dimensional image can be constructed.

Since the 1960’s, medical ultrasound sound has found uses in therapeutics through laser ablation and other surgical techniques and in diagnostics, with doppler imaging to measure blood flow, elastography commonly used for breast cancer detection, and three-dimensional fetal imaging [8], [30], [56]. Applications are ever expanding and the enumerated list is hardly exhaustive. In recent years, increasingly improved software capability has resulted
in advances in image construction through electronic beamforming - leading to drastic improvements in focus and increasing image acquisition speed.

When examined in parallel to ultrasound imaging advances, the developments in transducer technology has remained relatively stagnant over the past five decades. Initial technology, enabling one-dimensional imaging, used a single element of the bulk piezoceramic material lead zirconate titanate (PZT) for transduction. Manufacturing advances have enabled miniaturization of these elements, which can be assembled into one-dimensional arrays to produce higher resolution, two-dimensional images. Current fabrication begins when a backing, electrodes, PZT and matching layers are bonded and then manual diced into elements. The elements are then fitted with a lens and another backing layer for damping before final assembly into the ultrasound probe. All of these steps require an immense amount of manual labor, which leads to low yield and makes assembly expensive and inefficient. The latest beamforming techniques for 2D arrays to passively acquire 3D images (other 3D imaging techniques rely on controlled mechanical manipulation of a 1D array) require a large number of elements and even more intricate assembly.

Given this difficult fabrication process, producing such arrays is not practical or feasible on a large scale. In addition, assembly also limits element sizes and spacing between elements to slightly less than 1mm introducing interelement cross-coupling issues [25], [54], [34]. If assembly did not already limit the possibility of 2D array construction, size constraints would make the arrays bulky preventing use in intravenous surgical applications like angiography. Beyond fabrication, the physics of imaging prevent the bulk PZT transducer from achieving high sensitivity. The impedance mismatch between the elements and imaging medium introduce significant losses that can be improved with matching layers but still constrain bandwidth to ≤60% [56].

Smaller elements with high packing density and low mechanical impedance that can be easily and inexpensively manufactured are necessary for newly developed advanced imaging techniques to be fully realized. Such miniaturization needs lead naturally to a microfabrication. The latest, fundamental shift in transduction design is the capacitive micromachined ul-
trasonic transducer (CMUT). The CMUT uses an electrostatic actuation and sensing scheme with a plate geometry separated from the substrate by a vacuum gap. Changes in the electric field across the gap generate mechanical movement of the plate for actuation and the opposite effect is utilized for sensing. At small length scales, the electric field is large enabling high sensitivity not feasible with a macro-scale electrostatic device. Small plate mode operation significantly reduces mechanical impedance resulting in remarkably high bandwidth without the need for matching layers [35]. Conventional microfabrication is used construct CMUT element with sizes and spacing on the order of 50\(\mu\)m creating significantly smaller arrays. With the CMUT, 2D arrays can be more easily manufactured at a much smaller length scale making high resolution 3D imaging and small form factor probes possible. 2D and 3D images have already been demonstrated with CMUT technology and the latest developments are in the incorporation of CMUT devices onto catheter tips as a surgical tool [34].

With the CMUT’s success, there are still some drawbacks inherent to a capacitively driven design. A high bias voltage is required for operation stressing the material to the point of electrical breakdown and creating safety concerns. Higher sensitivity is achieved with reducing gap size to the limits of microfabrication capabilities, which constrain plate deflection and reducing the output acoustic pressure. The necessity of a non-linear drive voltage also introduces complexity into operation as conventional transducers respond to linear drive voltages. As such, existing electronic drive technology would have to be altered complicating the electrical system design. As solutions to these problems, piezoelectric micromachined transducers have be introduced as an alternative, microfabricated design solution. Such devices respond linearly to drive voltage, do not require high voltage, and their deflection and fabrication is not constrained by gap size. A similar plate mode actuation to that of the CMUT again lowers impedances making high bandwidth possible.

Despite its promise, most fabricated PMUTs fail to meet performance expectations. Coupling between the electrical and mechanical domain is generally much smaller than predicted [44], [5], [16], [53], [60], and sensitivity, especially in transmission, is so poor that researchers have turned to separate devices for transmission [9] or suggest that alternative
applications, not involving ultrasound transmission, would be best for implementation [44]. Many of these shortfalls are attributed to residual stress [44], [29], which has been alleviated through a partially suspended boundary condition [26], [44]. Up to this point, design optimization, including the realization of a partially suspended boundary condition, has largely been the result of trial and error since few analytical models exist to predict device performance. Attempts have been made to solve for device deflection [45], [53], [44] but the resulting solution is complex in most cases requiring numerical integration making it difficult to use for design optimization. Analytical models have yet been published for addressing the residual stress problems and defining an equivalent circuit creating further barriers to appropriate PMUT design based on modelling.

1.1 Research Objective

We report analytical modelling of acoustic pressure, deflection, and residual stress that can be used to improve the PMUT design process. The effect of residual stress on performance including bandwidth, sensitivity and resonant frequency is analyzed, and a means for alleviating the problem through optimized electrode coverage is proposed. Separately the potential acoustic pressure output of the PMUT is explored. High acoustic pressure is necessary for applications like intercranial pressure (ICP) monitoring of head injuries. Normally an invasive procedure that requires immense skill by the technician, a high acoustic pressure array could enable easier image acquisition that is non-invasive and high resolution. Potential acoustic pressure output of the PMUT is determined based on fundamental acoustic radiation equations and is compared to an equivalent rigid piston model based on average displacement previously used to describe CMUT pressure [37]. With the framework established to predict acoustic pressure, the design of the PMUT can be tailored for high acoustic pressure applications.

A novel solution technique for deflection is also introduced using the Green’s function approach. A Green’s function is defined for the fourth order plate vibration equation for a bimorph and unimorph PMUT based on [43]. With a stepwise defined forcing function
resulting from the applied piezoelectric moment, the overall plate deflection equation is explicitly solved for an arbitrary number of circular or ring electrode geometries. The resulting equation applies to a much larger variety of potential PMUT designs and can be more easily solved compared to previous solutions [44], [45]. Combined with the acoustic pressure analysis, the flexibility of the deflection solution is preliminarily demonstrated to optimize the plate deflection for high acoustic pressure.

From the modelling, a first generation PMUT design is selected and is currently being fabricated. With completion of the first generation device, we hope to prove the model validity and with performance knowledge and modelling expertise, move forward with more advanced designs. Our goal is to provide a more detailed modelling framework to PMUT design for application to the development of small form factor, high acoustic pressure devices. In the future, such a device could enable 3D real time imaging, and for particular applications like ICP, less invasive screening.
Chapter 2

Background

2.1 Medical Ultrasound Imaging

Ultrasound is an attractive medical imaging alternative to magnetic resonance imaging (MRI) and x-ray computed tomography (CT) scanning because of it is comparatively inexpensive and less harmful to the body [34]. However, the current 3D imaging capability of medical ultrasound devices lack the resolution and image quality of existing MRI and CT scan technology. With one-dimensional transducer arrays, 3D ultrasound is limited by slow image acquisition rates and dependence on operator expertise, which makes constructing a 3D image slow and lacking repeatability. The advent of a new transducer technology has the potential to eliminate these limitations making 3D real time imaging rapid and repeatable, and expanding the diagnostic applications of medical ultrasound.

2.1.1 History

In many regards, the use of ultrasound for medical applications is relatively young having only been commercially available for the past 50 years. Before its advent, ultrasound technology had first been developed in 1917 for sonar applications by the military [34]. The story goes that the observation of small fish killed in the wake of ultrasound beams inspired the application of ultrasound to medical therapeutics [61]. Even if this story does is not
entirely true, its suggestion that medical ultrasound was based on borrowed technology is valid. The first medical ultrasound was pieced together from existing sonar, metallurgical and radar systems at the time [30].

According to the American Institute of Ultrasound in Medicine (AIUM), a group of doctors attending a medical conference in 1951 shared a common interest in the use of ultrasound for therapeutics. In subsequent years, this same group formed the basis of the AIUM and medical ultrasound research efforts began to gain momentum. It was not long before the first images, although primitive, were captured. These A-mode and M-mode images were essentially one dimensional and displayed in the form of line plots. In M-mode, the echo amplitude is plotted as a function time and A-mode displays echo amplitude as a function of depth. These imaging techniques were applied to analyzing the human body with promising results. A-mode image scans showed a distinct difference between a normal stomach wall and one with cancer, and M-mode scans could display the motion of the anterior heart valve.

With these rapid results, it is not surprising that medical ultrasound technology became commercially available in the 1960s, and development of the medical technology continued at a similarly fast pace. The 1960s brought the first fully two-dimensional images; first bistable ones in black and white and then gray scale images not vastly different from those many are familiar with today. In 1964, the first commercially available machine produced by Physionic Engineering, Inc. produced bistable images gaining use in obstetrics, and the diagnosis of abdominal and pelvic diseases. With the availability of commercial machines, education and training programs became an important focus for groups like the AIUM.

More advanced imaging techniques and image quality improvements were gradually developed in subsequent decades. In the 1970s, doppler techniques including spectral, continuous wave, and color imaging, and equipment were developed for analyzing blood flow. Additionally, B-mode or brightness mode imaging, which forms the basis and most commercial medical ultrasound imaging today was implemented. An example system used B-mode imaging to detect defects in the human eye. By the end of the 1970s, most of the major
building blocks, particularly two-dimensional Doppler and B-mode imaging, that form the base of medical ultrasound imaging today had been established. Subsequent decades saw the expansion and modification of these techniques for new diagnostic and eventually therapeutic applications. Based on doppler techniques, power doppler imaging was developed in the 1980s, and B-mode imaging led to advances such as ultrasound mammography and renal cyst imaging in the 1990s. Stemming from ultrasound mammography, elastography, a technique to measure the effective stiffness of imaged masses, was applied to the detection of breast cancer becoming widely available in the 1990s.

The most logical next step from two-dimensional imaging was the realization of surface and volumetric images in three-dimensions. The first grayscale 3D images were demonstrated in the 1980s and implemented for fetal ultrasound in the 1990s. Until now, developments in the technology include improvements in resolution and scanning techniques [2]. As the development of 3D imaging continues, a final step toward 3D real time imaging (or 4D imaging) is the fundamental limit. For 4D imaging to be commercially realizable, scans must be consistent, high resolution, and occur at a fast acquisition rate. As we will show in the following sections, all of these factors are currently inhibited by transducer technology. Unless significant advances are made to ultrasonic transducers in the near future, 3D imaging resolution improvements will reach a final limit, and 4D imaging will not be possible.

Not just specific to 3D imaging, high resolution is becoming increasing important in newly defined applications. Improved image quality with aberrations present can be achieved through harmonic imaging, which requires high bandwidth to sense frequencies at twice the transmit frequency. New techniques involving the addition of contrast agents including color agents and micro-bubbles [30] for more detailed imaging similarly require high bandwidth. On top of the need for high bandwidth, superficial target imaging requires high frequencies, which again are not compatible with current transducer technology [8]. For now, we will sideline the topic of transducer technology for future discussion, but it is important to remember that significant current and future advances in medical ultrasound will hinge on its advance.
2.1.2 Basic Ultrasound Physics and Imaging

Imaging Types

From the previous discussion of ultrasound’s history, we have already introduced some information related to imaging types, but for further understanding, imaging principles will be explained in greater detail. In general, medical ultrasound imaging is based on the transmission of a short pulse, usually in a form similar to a sine wave, at ultrasonic frequencies on the order of 1-20MHz. Human anatomy is detected based on echo from features at various depth in the patient. As the pulse propagates through the human body, it is reflected and scattered from features in the anatomy, the echo, or the remaining reflected signal, propagates back to the imaging array where it is received. Based on the time between signal transmission and reception and the speed of sound in the media, the depth of the feature is determined. The first imaging techniques in A-mode and M-mode were based on this basic concept. A-mode also incorporated the amplitude of the received signal as a function of the depth.

In a sense, B mode or brightness mode is conceptually an extension of A-mode imaging to two-dimensions. The depth of the signal is determined using the same basic concept, and brightness reflects the relative strength of the received signal. For two dimensions, a longer one dimensional array is required to create a higher number of pixels for which the depth and brightness data can be recorded for construction into a 2D image.

Up to this point, it is assumed that the pulse is transmitted and received at the same frequency, but this is not necessary. As mentioned previously, receiving at harmonics of the transmitted frequency, usually the first harmonic, reduces imaging artefacts such as side lobes and the effects of non-linear wave propagation resulting in a cleaner signal [30]. For a higher signal to noise ratio, imaging at the first harmonic is preferential and is referred to as harmonic imaging.

Also based on B mode imaging, elastography is used to detect breast cancer. In elastography, a known force is applied to the tissue and a mass of interest, usually a lump suspected of being cancerous, is imaged during the process. The change in dimensions of the image are measured and the strain, stress, and effective elastic modulus of the mass can be determined.
Since cancerous masses differ in stiffness from conventional tissue, the elastogram can then be used to decide cancer diagnosis based on the elastic modulus of the lump [30].

Similar to elastography, doppler imaging techniques are used in the measurement of a physical property - in this case, blood flow. Velocity measurements are made by taking advantage of the doppler effect. When there is relative motion between a source and observer, the observed frequency relative to the constant frequency of the source changes with motion, commonly referred to as the doppler effect. For analyzing blood flow, the frequency shift $f_d$ related to the transmit frequency $f_t$ can be recorded to determine the blood flow velocity $v$ based on the imaging angle $\varphi$:

$$f_d = \frac{2f_tvcos\varphi}{c_m}$$

where $c_m$ is the speed of sound in the imaging medium [30]. Spectral doppler imaging displays the magnitude of the shifted frequency as a function of time with the intensity denoted by gray scale. When the spectral doppler image is superimposed with a B-mode image, color doppler images can be formed. In a color doppler image, the magnitude of the shifted frequency is denoted by the color scale on the two-dimensional image [30], [2]. More advanced forms of doppler imaging based on the same basic equation 2.1 exist, but their discussion will be excluded for simplicity.

**Array Types**

The first transducers used to construct the one dimensional A- and M-mode images consisted of a single bulk piezoelectric transducer element. For B-mode imaging, multiple, smaller transducer elements arranged in an array provide a two-dimensional image. The array type most commonly associated with fetal ultrasound is the trapezoidal or curvilinear for a wide viewing field. Resolution increases with the number of transducer elements, which is practically constrained by transducer element size and spacing between elements dictated by acoustic interactions. For digital control purposes, linear, trapezoidal, and curvi-linear
arrays can have as many as 256 elements but usually contain 128 elements because of element fabrication limitations. Each element is rectangular with typical width dimension of $1.3\lambda$, where $\lambda$ is the wavelength of the imaging frequency. This width is selected to create a narrow enough beam for focusing over a variety of angles while allowing for a wide field of view. The length of the element determines focusing ability with depth and is commonly chosen as $30\lambda$ to enable weak focusing [30]. For imaging in a water medium at 5MHz, the ideal transducer would then have a width of 0.4mm and length of 9mm.

Other array types include linear, sector and radial arrays, which are used for more specialized imaging. Sector arrays are most widely used for cardiac imaging where an entry point between the ribs requires a much smaller array. Additionally, the geometry of the sector array leads to a wide field of view at elevations consistent with the depth of the heart. When surgical applications require imaging internally, radial, linear and sector probes are used again because of their small form factor. In the radial format, the transducer can consist of one array element, which is rotates to construct a two-dimensional radial image [30].
Basic Imaging Physics

Understanding the basic fundamentals of acoustics is important in understanding the appropriate the design of ultrasonic transducer elements and arrays. As discussed in the previous section, acoustics governs the size and configuration of array elements for optimized focusing and field of view. In later sections, the governing acoustic equations and modelling for a particular element will be described in greater detail; however, for now, we will outline the basic acoustic concepts and associated metrics related to ultrasonic transduction. When referring to any type of acoustic wave propagation, the wave number $k$, angular frequency $\omega$, and wavelength $\lambda$ are basic algebraically defined values that are useful in subsequent, more detailed derivations.

$$\omega = 2\pi f$$  \hspace{1cm} (2.2)

$$k = \frac{2\pi}{\lambda}$$  \hspace{1cm} (2.3)

$$\lambda = \frac{f}{c_m}$$  \hspace{1cm} (2.4)

In the ultrasonic frequency range, the imaging wavelength in water is small ranging from 75$\mu$m for the high frequency of 20MHz to 1.5mm at 1MHz. Considering the density $\rho_m$ of the imaging media, another fundamental value, acoustic impedance $z_a$ can be introduced:

$$z_a = \rho_m c_m$$  \hspace{1cm} (2.5)

Similar to a mechanical and electrical impedance, the acoustic impedance relates an effort (acoustic pressure) to a flow variable (velocity). Acoustic impedance can thus be manipulated
in the same manner as a lumped circuit element when defining an equivalent transduction circuit. For a radiating boundary, the derived impedance is significantly more complex than the one shown above, but this simple definition still proves useful in understanding basic ultrasonic imaging.

In particular, the acoustic impedance mismatch, based on the definition in equation 2.5, is a significant loss mechanism in transmission. Since the acoustic impedance of the bulk piezoelectric transducer elements is much higher than tissue, approximately 80% of the transmitted signal is lost at the interface between the tissue and transducer. With such a high signal loss, sensitivity can be dramatically reduced. To overcome this loss, quarter wavelength matching layers with impedance $z_{ml} = \sqrt{z_{trans}z_{tis}}$ are added to the transducer face reduces signal loss. The quarter wavelength refers to the thickness of the matching layer and is chosen for optimal wave reflection and transmission from the matching layer. A transmitted pulse reverberates within the matching layer interfering destructively when reflected back to the transducer elements and constructively to form a larger amplitude pulse into the tissue. Multiple matching layers connected in series can reduce the transmitted signal loss to nearly zero [30]. During a fetal ultrasound, the gel applied to the probe tip in contact with the woman’s abdomen serves a similar purpose in impedance matching [34]. Similarly, the interface between bone and tissue suffers from high loss induced by acoustic impedance mismatch, so imaging is usually avoided in regions where transmission is required through bone. Hence, sector arrays are implemented for cardiac imaging through placement between the ribs, preventing any signal transmission through bone.

Table 2.1: Acoustic impedance of water and in various parts of the body [30]

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Acoustic Impedance (kg m$^{-2}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>$1.5 \cdot 10^6$</td>
</tr>
<tr>
<td>Liver</td>
<td>$1.66 \cdot 10^6$</td>
</tr>
<tr>
<td>Kidney</td>
<td>$1.64 \cdot 10^6$</td>
</tr>
<tr>
<td>Fat</td>
<td>$1.33 \cdot 10^6$</td>
</tr>
<tr>
<td>Bone</td>
<td>$6.47 \cdot 10^6$</td>
</tr>
</tbody>
</table>

Beyond impedance mismatch, signal attenuation within the tissue contributes to loss
and limits imaging depth. The transmitted beam signal intensity is attenuated exponentially with depth into the tissue, and the rate of this decay is related to frequency. Higher frequency pulses are attenuated more quickly at shallower depths making low frequencies more appropriate for deeper imaging applications. As such, ultrasound frequency selection is a compromise between resolution best achieved at high frequencies and attenuation, which is less pervasive at lower frequencies. Superficial imaging, at depths where attenuation is not significant, is commonly performed at higher frequencies (10-15MHz), and imaging of deep organs or tissue is performed at lower frequencies (3-5MHz) [30].

**Imaging Metrics**

Large signal loss can be avoided with the incorporation of matching layers and the selection of an appropriate transmission frequency resulting in increased sensitivity. However, inherent limitations in sensitivity exist based on the transducer design and other loss mechanisms that are difficult to circumvent. Array bandwidth provides a direct measure of sensitivity and allows for the comparison of a variety of transducer designs and array configurations. A high bandwidth is desirable because sensing can be achieved over a wider range of frequencies, and higher axial resolution is possible. With a high enough bandwidth, transmission and receiving can occur at different frequencies enabling harmonic imaging with a single device. Bandwidth is measured from the signal intensity waveform (whether transmitted or received) over range of frequencies. Commonly -3 dB bandwidth is reported and refers to the range of frequencies over which the signal intensity is greater than or equal to half the maximum intensity, which occurs at the center frequency [30]. With matching layers, commercial transducers are capable of -3dB bandwidths of over 60% of the center frequency [30], [56].

Other metrics including insertion, transducer, and electrical loss are commonly reported to describe device efficiency [56]. Insertion loss compares the equivalent circuit power with the resistive load of the transducer to the power in the same circumstances with no transducer present. Insertion loss is a measure of the overall efficiency as opposed to transducer loss,
which measures one-way efficiency. Overall power compared to the power arriving at the acoustic load denotes transducer loss. Through the definition of electrical loss $EL$, the transducer $TL$ and insertion loss $IL$ are directly related:

$$TL = EL \cdot IL$$ (2.6)

With relation between electrical, insertion and transmission loss, the effect of the individual loss factors on the whole can be analyzed [56].

Separate from loss and efficiency, output power, intensity and pressure are also critical to device performance. From acoustics, the radiated pressure from the surface is determined in later sections. Here we will assume that the pressure profile is known and determine the power and intensity accordingly. Formally, the definition of intensity $\bar{I}$ is related to the pressure $p$ and the complex conjugate of the propagating wave velocity profile $u^*$:

$$\langle \bar{I} \rangle = Re \{ p u^* \}$$ (2.7)

Physically, the intensity has direction and is acoustic power over unit area, and can be reduced to:

$$|I| \approx \left| \frac{p}{\rho m c_m} \right|^2$$ (2.8)

The above intensity equation holds true for plane waves and those emitted radially from a point source [1]. From this physical argument, the power can logically formally written:

$$W = \int \int \langle \bar{I} \rangle \cdot dS$$ (2.9)

where $S$ is the surface over which the power is to be computed [1]. For medical ultrasound transducers, the output can be measured in terms of power, intensity or pressure. Most commonly pressure values or sound pressure levels are reported; however, the conventions of intensity and power might also be useful. For example, the intensity can be used to determine the acoustic projection of an array, and since the power is not directionally important, it can
be a more universal metric for comparison.

### 2.1.3 Trends toward 3D Imaging

3D imaging has generated interest as an attractive application for medical ultrasound for over 20 years [54], [25] especially given the limitations of 2D imaging, which is highly dependent on the technician limiting image reproducibility. In comparison to CT or MRI technology, ultrasound equipment is less expensive and can acquire images at a faster rate; however, these capabilities do not come without caveats. Ultrasound is sensitive to noise sources inherent in acoustic physics such as aberrations, speckle, and shadowing, which must be overcome to achieve the necessary resolution for real-time 3D diagnostics. As a result, the first research tackling 3D imaging focused on the development of new techniques to reduce and alleviate the effects of acoustic imaging noise including phase correction, elevation focusing and synthetic aperture imaging approaches [54]. Other work examined the system level design laying applying existing technology for clinical testing of 3D imaging for radiology and echo-cardiology [25].

Image acquisition can be achieved via mechanical scanning with a 1D array or 2D array construction. Mechanical scanning is performed free hand with the aid of acoustic, magnetic or articulating arm positioner, or with a mechanical localizer using linear, fan or rotational scanning. With mechanical localization, the image third dimension is acquired through a mechanical mechanism that precisely sweeps the 1D array. This varies from the free hand positioning, where a technician moves the 1D array guided by a positioning feedback system. With either technique, third dimension acquisition adds significant mechanical and control complexity that could be eliminated through use of a 2D array. Image acquisition rate is also limited and is determined by the speed of the mechanical scanning motion.

Volumetric imaging possible with three dimensions can be useful for a variety of applications - many of which require accurate measurement of physical features. Volumetric imaging has been possible with 1D array scanning for intravascular ultrasound. The resulting volume images were acquired for measurement of vascular dimensions and atheroma volume.
Although measurement of these dimensions was possible, imaging lacked repeatability making the initial measurement useless without accurate further imaging for comparison. In addition, the motional nature of mechanical scanning introduced imaging distortions further limiting imaging accuracy [25].

The inaccuracy and complexity of 1D array scanning could be overcome with an imaging system utilizing a 2D array that requires no moving parts. The acquisition rate with a 2D array can also be much faster and is only limited by acoustic physics, particularly the speed of sound in the imaging medium. A faster acquisition rate means that patient examination time could be greatly reduced. With its potential, it seems surprising that 2D array imaging has not been the dominant approach for 3D imaging. Many authors argue that the largest challenge in the implementation of a 2D array is existing transducer technology [8], [25]. The labor intensive assembly and low yield of the manufacturing process makes large 2D array nearly impossible to fabricate, and the large number of elements create electrical interfacing problems [25], [34]. A larger number of elements require many electrical connections introducing lead bundling difficulties [25] and crosstalk problems [54]. Beyond manufacturing and array interfacing problems, the high acoustic impedance of the bulk PZT limits resolution reducing bandwidth and sensitivity and increasing insertion loss [54]. Given fabrication and electrical crosstalk challenges, 2D array imaging has best been accomplished by undersampling of a large number of array elements. However, undersampling limits dynamic range making small features difficult to detect and can excite parasitic, lateral vibration modes in the bulk PZT devices [54].

In addition to the previously mentioned advantages of the 2D array, noise reduction and imaging resolution improvements to existing transducer technology could broaden the range of 3D medical ultrasound applications. For ideal image resolution, 3D images would be constructed from infinitely thin 2D images along the elevation direction. In reality, slice thickness, the width of the beam in the elevation direction, constrains resolution. For a given height, the slice thickness is determined by beam focus. A larger number of elements electronically controlled in 2D would enable focusing over a wide variety of depths, which
is currently impossible with a 1D array even with the addition of a lens. Electronic beam-forming would also enable focusing in directions within the scan plane [54]. With improved focusing, it might be possible to circumvent imaging noise resulting from tissue inhomogeneities including aberrations.

Image acquisition, particularly related to the bulk PZT transducer technology, is the major bottleneck in the implementation of high speed, high resolution 3D imaging [25], [34]. With system level research, image processing and visualization is already possible [25] with existing visualization adapted from MRI and CT technology [30]. Construction of sparse 2D arrays in a research environment has also enabled the development of the necessary beam forming techniques [54]. The development of new applications based on these beamforming techniques and the potential of real time 3D ultrasound place further requirements on the design of a new transducer. In vivo elastography requires high resolution and repeatability for stiffness measurements in real time. High acquisition rates and high frequencies are necessary to achieve a high signal to noise ratio with a constant depth scan.

2.1.4 Bulk PZT Transducer Fabrication

As is clear by now, for 3D real time and more advanced imaging to be possible, the transducer technology, particularly manufacturing, needs to be altered. The fabrication technique to produce current transducers has already reached a fundamental limit that is a balance between yield, element size and assembly time [34]. Other packaging and integration factors including wire connections and bundling play a role in the final determined array size [54], but in general, the element size and spacing between elements is dictated by the manufacturing process [20].

Since the manufacture of ultrasound transducers is big business, nearly $1 billion annually [34], it is not surprising that many of the details of the manufacturing process is proprietary; however, a general flow of the process can be pieced together from research and ultrasound related publications. Manufacturing begins with producing a slab of piezoceramic material, usually a particular compound of lead zirconate titanate, that is later diced with a diced
with a diamond wire or wheel dicing saw [20], [54]. The piezoceramic is constructed from a powdered form through conventional powder processing and sintering [51]. Typically the material stack of the main array is then assembled and includes the bulk piezoceramic with top and bottom electrodes, impedance matching layers on the front, and a backing material that acts a support in the subsequent dicing step [56]. For appropriate bonding between layers, lapping and polishing are required before stack assembly with epoxy [58], [51] or the aid of a bonding jig [54].

A diamond wire or wheel dicing saw is used to the separate the elements in the material stack. The width of the saw, which can be less than 50μm [20], [58], determines the spacing between elements, referred to as kerf. Kerf is not necessarily the same as the cutter width - for a 45μm blade, a kerf of 150μm has been reported [58]. The periodicity of elements in the array is the pitch, which is simply the sum of the kerf and sliced transducer width [56]. Cross-coupling between elements is determined by kerf and thus width of the dicing blade. Crosstalk can be parasitic in imaging applications, and the kerf should be ≤ λ/2 to minimize coupling. For an imaging frequency of 5MHz, this corresponds to a kerf ≤ 150μm, which is just barely possible with a 45μm blade [58]. As a result, it is difficult to fabricate an array with this kerf specification, so cross-coupling is a problem. A more flexible manufacturing process could tailor the spacing between elements to reduce crosstalk without changing the manufacturing process - here again microfabrication proves to be a nice alternative.

Once diced, the one-dimensional array elements are fitted with a lens for focusing, and then additional bonding is required for electrical connections. Before electrical connections are made, the elements are taken off the backing layer support. In some cases, cleaning has been performed to remove unwanted particles remaining from the cutting process [51]. For research applications, elements are bonded with epoxy before electrical connections are made [51], [54], [58]. Presumably this step is added to make the assembly process easier since later processing does not require the manipulation of individual array elements. In industry, a flexible printed circuit is electrically connected to the top and bottom electrodes, and the final device is secured on a new backing layer that provides damping during operation. All of
the above processing steps describe fabrication of a one-dimensional array. Two-dimensional arrays are constructed in a similar way except additional dicing and bonding steps are required making the process even more time consuming and labor-intensive [51], [54]. After reviewing the process, the amount of manual labor required for assembly is very clear, and it is understandable that such a process could result in low yield, repeatability, and poor quality devices especially with a brittle piezoceramic base material. The constraints inherent in the fabrication process also directly limit the capabilities of the transducer. Specifically, the dicing saw width introduces cross-coupling that reduces sensitivity. An alternative transducer design involving a new and more innovative fabrication process will be crucial to increasing device sensitivity and improving manufacturing yield.

2.1.5 Transducer Developments

The shortfalls of traditional transducer fabrication has understandably spawned interest in alternatives. With the original development of 3D imaging in the 1980’s, 2D array fabrication has become increasingly necessary making manufacturing even more difficult. In the
early 1980's, transducer improvement focused largely on developing alternative materials to replace the piezoceramic. Plastics like polyvinylidene difluoride (PVDF) and its derivatives were particularly attractive because of their lower mechanical impedance enabling improved impedance matching and low lateral coupling for better sensitivity. High frequency applications were explored for implementation but low capacitance and high parasitic capacitance made the resulting elements difficult to use [61]. Other interest has been in composite or piezo-polymer materials, which do not require dicing. Despite improvements in fabrication, bandwidth is still limited to $\approx 60\%$ (similar to bulk PZT transducers with matching layers) and interelement crosstalk is reduced but not eliminated with contributions of $< 30dB$ [13].

More recent work has shifted to PZT thick film deposition techniques again for high frequency applications. High frequency imaging works best for superficial targets and has generated interest for eye, skin and intravascular imaging. Thick PZT films, still thinner than those produced by conventional manufacturing techniques, would naturally resonant at a high frequency while still functioning like a bulk piezoceramic. Large scale use of thick PZT films is limited by the low porosity which significantly reduces coupling, piezoelectric constant and dielectric permittivity. For improved film quality, densification can be performed at higher temperatures with unfavorable results. The large change in temperature necessary for densification increases residual stress and at high enough temperatures, lead oxide sublimes creating lead deficiencies in the film. New deposition techniques are currently under development to improve film porosity without requiring an elevated temperature step. Electrophoretic deposition, screen printing, and spin coating are plausible alternatives; however, advanced processing techniques still result in lower film quality compared to thin film deposition [36].

Perhaps the most promising step in next generation transducer design will come with micro-fabrication. At high volume, micro-fabrication is less expensive and naturally supports the needs for smaller transducers for high resolution imaging. At this scale, entirely new transducers designs must be considered and not just improvements to the piezoelectric material properties and deposition. With the right design, micro-fabrication could drasti-
cally improve the transducers and new, advanced imaging techniques could finally be commercialized. In the next section, we will discuss recent developments in microfabricated ultrasonic transducers (MUTs) currently under development, which show remarkable potential for larger scale implementation.

2.2 Capacitive Micromachined Ultrasonic Transducers (CMUTs)

2.2.1 History and Overview

Capacitive ultrasonic transducers are not a completely new idea. Some of the first capacitive transducers research for ultrasound applications occurred in the 1950s and 1960s [15]. These transducers were fabricated from a back plate with machined grooves for the individual transducers and a thin metallized, dielectric membrane, sometimes mylar, was used for actuation [24]. In the early 1990's, the capacitive ultrasonic transducer was revisited with the advent of microfabrication. Initially devices were only partially micro-fabricated with separately bonded dielectric films. One example used micro-machining to form grooves in an aluminum backplate and a 5\(\mu\)m thick polyethylene terephthalene membrane coated with gold was tensioned across the front [6]. The fully micro-machined transducer was introduced in 1994 and laid the groundwork for most of the CMUT work completed today [4], [22].

The earliest capacitive ultrasonic transducers and current CMUTs use the same electrostatic transducer mechanism for actuation and sensing. Each transducer consists of a fixed, conductive back plate separated from a structural membrane layer by an air or vacuum gap. By metallizing the membrane, charge can be carried across the gap generating an electric field. When the electric field is strong enough, the electrostatic force exceeds the mechanical stiffness force of the membrane causing deflection. For actuation, a voltage pulse is applied varying the electric field, which pulses the membrane sending an ultrasonic pressure wave into the medium. Once reflected from an object of interest, the pulse can be received by the same CMUT. This time the pressure applied to the membrane causes deflection changing
the gap height and thus the electric field generating a signal. When multiple CMUTs are operated in parallel making up an array, the received signal can be constructed into an image of the object of interest.

CMUTs are arranged into arrays in a slightly different configuration from traditional bulk piezoelectric transducers. An element consists of rows and columns of CMUT single devices or cells. Cells can be square, rectangular, hexagonal or circular and the spacing between cells is kept small, on the order of the cell size itself, and is determined by acoustic coupling. A CMUT element can be similar size to a bulk piezoelectric one, so the CMUT elements have a much higher number of cells in columns as opposed to rows. For a 128 element one-dimensional CMUT array, elements were 5x150 cells corresponding to a 200$\mu m$ width and 6000$\mu m$ length [22]. When imaging, all cells in the element are actuated in parallel and beamforming is performed by varying the firing time between elements.

2.2.2 Fabrication

CMUT fabrication can be roughly divided into two categories: sacrificial release and wafer bonding processes. The earliest CMUTs and those that followed for many years after were largely produced with sacrificial release based techniques. For increased control, yield and uniformity, newer devices have been produced via the wafer bonding approach [24]. Wafer bonding techniques also have the advantage of requiring fewer masks and processing steps thus reducing manufacturing time.

Most sacrificial processing begins with a silicon substrate that is highly doped or low resistance on the surface serving as both a support and bottom electrode. For a majority of devices, low pressure chemical vapor deposition (LPCVD) silicon-rich silicon nitride is used for a variety of functions: as the membrane structure, an electrically insulating layer and/or an etch-stop for the release process. Once doped, a thin layer of LPCVD silicon nitride is deposited as a sacrificial layer - typically thickness on the order of 1000Åare selected to avoid pin hold problems. The sacrificial release process is commonly performed with two subsequent sacrificial deposition and etching steps. LPCVD polysilicon is deposited
and patterned via photolithography to define an outer cavity ring, and a dry etch is used to remove the polysilicon. Another LPCVD polysilicon deposition, photolithography and etching step is performed to define the cavity. The thickness of polysilicon from the first two depositions determines the cavity gap height. A second layer of LPCVD silicon nitride is then deposited to form the structural membrane layer. Another lithography and dry etch step defines holes in the silicon nitride to access the polysilicon during release. A potassium hydroxide (KOH) etch clears the polysilicon and a final silicon nitride deposition plugs the access holes sealing the cavity.

The first steps of the sacrificial release process generally follow these steps; however, different etches and sacrificial materials can be selected. For faster sacrificial layer etching, a sacrificial layer combination of phosphorous doped polysilicon and PECVD silicon dioxide is used, and the release step is performed with an isotropic silicon etch [12].

Alternatively, sacrificial release fabrication can be performed with a variety of low temperature processes. Beginning with thermal oxide growth on a silicon wafer for electrical isolation, an aluminum bottom electrode is deposited via sputtering. A thin layer of plasma enhanced chemical vapor deposition (PECVD) silicon nitride is deposited as an etch stop. Instead of polysilicon, chromium serves as the sacrificial layer and is evaporated and patterned onto the surface. Chromium is chosen as the sacrificial layer because of its high etching selectivity compared to silicon nitride. After chromium patterning, a second silicon nitride layer is deposited at low temperature, and wet etching is used for the membrane release. A third and final silicon nitride deposition then plugs the holes used for the release effectively sealing the cavity [14].

In each fabrication process, it is critical that the silicon nitride is low stress during the membrane layer deposition to avoid buckling or cracking of the membrane during subsequent processing and device operation. Stress for this deposition is usually controlled via careful regulation of the reactant gas chemistry [14], \( O_2 \) annealing post-deposition [12], and/or deposition temperature [22]. In nearly all cases after release, the aluminum top electrode, is sputtered and patterned on top of the membrane to electrically complete the device. For
immersion applications, an additional passivation layer is required to electrically separate
the electrode from the imaging medium. A thin layer of low temperature oxide (LTO) can
provide the necessary passivation [22] or alternatively a parylene coating can be applied [24].

Careful considerations must be made during the sacrificial processing steps, which deter-
mine some of the selected device and processing parameters. Polysilicon is highly compres-
sive, which can cause cracking or buckling during release. Etch chemistry and conditions
must also be carefully regulated during release to increase uniformity and avoid stiction,
which can cause permanent device collapse. As such, alternative etching techniques includ-
ing dry etching, supercritical drying or freeze drying can be employed for release. Finally,
silicon nitride is not highly adhesive and therefore the access hole dimensions must be kept
small otherwise the nitride will fill the cavity [22].

Based on these fabrication considerations and structural requirements, gap height is con-
strained to slightly less than 1\( \mu \)m using this conventional processing approach. Alternative
sacrificial techniques have been developed to bypass these constraints yielding a gap as small
as 55nm with a single sacrificial layer etch. However, the processing parameters required
even more precise control of a time hydrofluoric acid (HF) etch limiting the practicality
and repeatability of the process [4]. A small gap height is also not necessarily ideal. A re-
duced gap height leads to higher sensitivity but also reduces deflection and output acoustic
pressure. As a result, the optimal gap height is a balance between sensitivity and transmit
power [4]. Advances in microfabrication are enabling increased residual stress and thickness
control enabling smaller gap heights to be more easily fabricated via sacrificial etching. With
these improvements, it is likely that the gap height thickness limitations will continue to be
reduced. Regardless of the stated limitations of the sacrificial release process, it is still highly
reliable and can achieve remarkable results. Oralkan et al fabricated a 128 element array
containing a total of 96000 CMUT cells with 100% yield using sacrificial release [22].

In recent years, a wafer bonding technique has been introduced to remove the fabrication
requirements imposed by sacrificial etching. Thermal oxide is grown on a highly conductive
silicon layer and patterned via photolithography and a buffered oxide etch (BOE) to form

50
cavities. An additional, thin thermal oxide layer is grown for electrical isolation completing fabrication on the bottom wafer. Wafer bonding is performed at vacuum between the cleaned and etched silicon wafer and a silicon-on-insulator (SOI) wafer. The SOI device layer was bonded to the top of the patterned silicon substrate and served as the membrane layer after later processing. Grinding and etching with KOH removed the handle layer of the SOI and a BOE etch removed the buried oxide leaving only the device layer. A post-anneal is used to secure the bonding between the device layer and oxide posts in the substrate. Patterning with photolithography and dry etch exposes a channel to the bottom silicon layer, which forms the bottom electrode. The process is completed with the sputtering and patterning of the top, aluminum electrode [31]. In this process, the cavity depth is controlled by the thermal oxide thickness, and the membrane thickness is determined by the as-purchased device layer thickness of the SOI wafer. Unlike the release process, the uniformity between cells is high leading to predictable and consistent measured properties like resonant frequency, collapse voltage and electromechanical coupling [31].

The wafer bonding technique has introduced increased uniformity to a process that was already highly reliable with sacrificial release. CMUT manufacturing is now high yield and highly uniform leading to easy construction of larger arrays for one- and two-dimensional imaging.

2.2.3 CMUT Modelling

With reliable fabrication, many CMUT researchers have focused on modelling to produce high performing arrays without extensive trial and error. Most CMUT models are based on plate/membrane vibration theory and the linearized Mason equivalent circuit model pictured in Figure 2-3 with plate geometry shown in Figure 2-4 and parameters in Table 2.2. The listed parameters are specifically for a plate. If residual stress is to be considered, the plate capacitance and inductance should be replaced by the lumped mechanical impedance for a stressed plate $Z_m$ in equation 3.13 multiplied by the plate area $\pi a^2$ [39].

Although Mason's model provides a strong basis for approximately CMUT performance,
non-linearities especially in driving the CMUT, electrical and mechanical losses, and complicated acoustic interactions between the CMUT and imaging medium are not fully realized with the basic Mason circuit. Hence modelling efforts have focused on modifying the Mason model or purposing new models to account for these influences. Loss mechanisms have been considered by adding parasitic elements to the transducer circuit. Electrode size was found to be related to parasitic capacitance and affects the transformer ratio for a coverage of less than 50% of the total plate area [11], [24]. Parasitic electrical and acoustic impedances have also been added in series to related circuit components to account for losses [4]. Despite these modelling efforts, the effect of parasitics can be difficult to predict and usually must be experimentally measured. A more easily predicted loss mechanism is determined by considering the effect of air loading by the cavity. From the governing plate vibration equation, the deflection is analytically determined and a finite element model examined the effects of tension, air cavity volume and stiffness on the mechanical impedance. It was found that the air cushion contributed to a 22% error between the membrane model and membrane-on-air model, which was validated experimentally [15].

Separately, a one-dimensional circuit model is employed to analytically predict the effects of non-linearities. One-dimensionality is achieved by decoupling the time and geometry dependent terms in the deflection equation. The equivalent circuit is derived by considering the time dependent displacement of an equivalent rigid piston, where the area and average velocity of the equivalent piston are defined based on finite element modelling of the coupled deflection profile. The influence of non-linearity is explored in the mechanical and acoustic domains assuming piston pressure radiation into a wave guide. The non-linear model is subjected to a voltage pulse input and the resulting sound pressure frequency response is determined [37].

Additional acoustic models have lumped the impedance of all cells electrically actuated in parallel forming an element. In the most basic model, acoustic loading is simply taken as the acoustic impedance $z_a$ of the imaging medium defined in equation 2.5 multiplied by the element area. For a low quality factor $Q$, the membrane impedance has been neglected
coupling the acoustic directly to the electrical domain [11]. In a different approach, the lumped element impedance is scaled down to the cell level by a factor related to the number of elements [12]. More detailed acoustic modelling has been performed to resolve artifacts experimentally observed when imaging with arrays. Device interactions across an array were found to create non-uniform fluid loading, which proves difficult to decouple from the mechanical domain. As a result, Caronti et al purpose a lumped mechanical-acoustic impedance in the Mason circuit for increased model accuracy [15].

Finally to validate newly derived models, more accurate device testing can be required. Pitch-catch measurements between a device and hydrophone are commonly used to determine acoustic responses [31], but impulse response is limited by transmitter and amplifier response. For this purpose, new equipment has been developed for improved electro-acoustical characterization [12].

2.2.4 Recent Innovations

Progress in modelling and fabrication has enabled the design of robust CMUTs with predictable and well-characterized performance. With these innovations, the logical next step for the CMUT is incorporation into larger arrays for advanced imaging and a push toward commercialization. In the past decade, CMUT arrays have demonstrated promising results with improved performance above the traditional bulk piezoelectric transducers. The first b-mode scanning images using 128 elements displayed 80% -6dB bandwidth at a frequency of 3MHz. As mentioned previously, these device were fabricated with a sacrificial release process producing cells with a remarkable 100% yield [22]. Other echographs performed with wafer-bonded CMUT arrays show further increases in sensitivity. A wafer-bonded array operating in the 3-13MHz range exhibited 130% -6dB bandwidth [39]. With the success of the first imaging arrays, CMUT arrays have been researched for other applications including an annular array for intravascular ultrasound imaging. Incorporated into a surgical probe tip, the resulting IVUS consisted of small form factor array with elements 43x140 $\mu m^2$ with high bandwidth (104% -10dB). All of these bandwidths are a staggering improvement to
the maximum ≈60% -3dB bandwidth [56] achievable with conventional PZT transducers and multiple matching layers.

As compared with conventional piezoelectric transducers, the CMUT has already demonstrated higher bandwidth without the need for matching layers, which can lead to additional imaging capabilities. Direct comparison with imaging probes of assembled CMUT and conventional PZT transducers arrays showed much higher resolution of the CMUT array in imaging a carotid artery and thyroid gland [39] at shallow depths. A further increase in the resolution of superficial targets can be achieved through high frequency imaging with CMUT cells. The small size of the CMUTs allows use at high resonant frequencies, which is not currently possible with existing transducer technology. Efforts to commercialize the CMUT are currently underway, but still face some limitations [34]. In particular, imaging at larger depths is still best with conventional transducers, which is likely related the acoustic pressure output of the CMUT arrays [39].

2.2.5 Limitations

Effective transducer design should consider a variety of efficiencies that are part of the overall system, electrical, and mechanical domains. For the pulse-echo imaging commonly used in medical ultrasound, the pulse length is shorter than the time between transmission and reception, so transmission and reception can be decoupled for transducer evaluation. Efficiency is a broad term when referring to the ultrasonic transducer and should be divided into a series of related metrics to create a clearer objective for the evaluation of device performance. The effectiveness of transmission, receiving and overall device performance will be based on the following set of goals, which are commonly used in ultrasound transducer design [56] and are outlined at the top level of the functional requirement hierarchy shown in Figure 3-1:

- High sensitivity
- High acoustic power
Figure 2-3: Ultrasonic transducer equivalent circuit model based on lossless system with deflection in plate mode [32]

- Low power consumption
- Resonant frequency appropriate for medical imaging

Although this might seem like a short list of metrics, each are dependent on parameters such as acoustic pressure, mechanical restoring force, and parasitics, which are critical to device performance.

Using the principles of electrostatic transduction and experimental data, the listed metrics can be used to determine the effectiveness of the capacitive micro-machined ultrasonic transducer (CMUT) and to determine its limitations. For the best understanding of the capabilities of the CMUT, it is useful to examine the equivalent circuit model. Although greatly simplified, the described metrics for transducer evaluation fall directly from the circuit model and can be compared with available experimental data for verification. A diagram of a typical CMUT is provided and the relevant equivalent circuit parameters are listed in the table below. These table values will be referenced in later description of the effective transducer metrics.

In traditional piezoelectric transducers, mechanical bandwidth is severely limited by acoustic impedance mismatch between the bulk piezoceramic material and the imaging medium. With the addition of a quarter wavelength matching layer, the -3 dB fractional bandwidth can be improved to 60% [56]. For ultrasonic devices operated based on plate or membrane mode, the mechanical impedance can be more closely matched to the imaging
medium resulting in greatly improved bandwidth. In literature, mechanical bandwidths of 100% and greater have been reported for such devices [35]. Thus, the CMUT does not suffer from the same bandwidth limitations.

In receiving mode, bandwidth should be maximized because a broader receiving band will exhibit less noise and will therefore reduce the signal-to-noise ratio of the incoming receiving signal. For the CMUT, the device bandwidth $\Delta f$ is inversely proportional to the RC time constant of the equivalent circuit model. Based on bandwidths dependence on mechanical impedance, plate geometry and acoustic medium density $\rho_m$ and sound speed $c_m$ are explicitly related to the bandwidth [11].

$$\Delta f = \frac{128Yt_n^3}{27\rho_mc_m(1-\nu^2)a^4}$$ (2.10)

Particularly, the device radius and plate thickness determine the bandwidth if the acoustic medium parameters are assumed constant. Bandwidth is maximized with a small plate radius and large membrane thickness. For an ultrasonic signal range of approximately 5MHz,
harmonic imaging is possible with sensing at 10MHz, requiring a minimum bandwidth of approximately 5MHz. From Figure 2-5 for a 5MHz bandwidth to be achieved, the plate radius is limited to less than 25\( \mu m \) for membrane thickness \( \geq 2 \mu m \).

High attenuation of the transmission pulse makes receiving sensitivity particularly important. Receiving sensitivity is defined as output voltage per unit of pressure and is directly related to the magnitude of the electrostatic force. A high electrostatic force translates to a greater sensitivity to receiving signal. Based on analysis from an energy perspective, the electrostatic force \( F_e \) is defined as follows [11]:

\[
F_e = \frac{C_0 V^2}{2(t_n + (\varepsilon_r/\varepsilon_0)(t_a - x))}
\]  \hspace{1cm} (2.11)

The electrostatic force is more sensitive to changes in voltages given the scaling with the quadratic scaling as opposed to changes in capacitance, which are linearly related to the electrostatic force. Thus, the voltage is more important than the electrical capacitance in determining the receiving sensitivity.

However, the effective gap height \( t_n + (\varepsilon_r/\varepsilon_0)t_a \) should also be considered in determining the magnitude of electrostatic force, and its relationship to the voltage makes the determination of the electrostatic force slightly more complex. By rewriting the equation (2.11) in
terms of only the gap height and related material parameters, the electrostatic force can be more simply defined:

\[
F_e = \frac{\varepsilon_0^2 k_m (t_n + (\varepsilon_r/\varepsilon_0) x)}{\varepsilon_0^2 (t_n + (\varepsilon_r/\varepsilon_0) (t_a - x))}
\]  
\tag{2.12}

For deflection to occur, a voltage must be applied across the electrostatic device. A small gap height resulting from a large deflection will create a high electrostatic force. Since the DC bias is much larger than the AC voltage required for the ultrasonic pulse, the DC voltage largely dictates the electrostatic force and thus the available electrostatic energy in the system. In receiving, changes in electrostatic energy create a signal and are determined by the acoustic pressure and the transformation ratio between the mechanical and electrical domains. Although the capacitance and transformation is important for the receiving signal, an initially high applied DC bias voltage increases the electrostatic force making small changes in voltage more easily sensed. Experimentally, the received signal has been shown to increase with applied bias voltage [12]. Logically, CMUTs are thus operated at high DC bias voltages for increased sensitivity, and it is not uncommon to see voltages on the order of 50V or higher [35].

However, the DC bias voltage is practically limited by pull-in and the dielectric strength of the plate material. The dielectric strength of silicon nitride, a material commonly used in the CMUT plate construction, is approximately 1000 V/m [3]. Although seemingly high, this dielectric strength can limit the thickness of the plate. For a DC bias voltage of 100V, the plate must be thicker than 100nm otherwise the dielectric strength of the plate will be compromised. Depending on the device design, the CMUT might already be limited by pull-in voltage before reaching the point of dielectric breakdown. Permanent collapse of the electrostatic device will occur when the pull-in voltage is reached. The pull-in voltage is determined by device geometry in addition to material parameters and gap height [11]:

\[
V_{Pl} = \sqrt{\frac{8k_m (t_n + (\varepsilon_r/\varepsilon_0) t_a)^3}{27\varepsilon_0 \pi a^2}}
\]  
\tag{2.13}
Figure 2-6: Pull-in voltage related to gap height of a 500nm thick plate for different plate radii. Dielectric breakdown voltage plotted based on the dielectric strength of silicon nitride.

As a general rule applicable to all electrostatic devices, the deflection cannot exceed 2/3 of the gap height otherwise collapse will occur [52]. For dimensions consistent with CMUT designs, the pull-in voltage is usually reached before dielectric breakdown as shown in Figure (2-6). However, for small plate radii and large gap height, it is possible for dielectric breakdown to occur before electrostatic pull-in.

Although slightly less important than the DC bias voltage, the transformer ratio $n$ should also be considered to maximize the device receiving sensitivity. As the connection between the electrical and mechanical domains, the transformation ratio can directly affect the received signal. The output current is proportional to the transformation ratio multiplied by the plate velocity generated from the receiving pulse. Thus, a high transformation ratio will create a better received signal. Near pull in, the transformation ratio is maximized with a thicker plate thickness and minimum gap height. It is important to recall that the bias voltage is dependent on gap height and it is therefore inherent that a large applied bias voltage is required to achieve a small gap height. Additionally, a large plate thickness requirement is important to maximize the transformation ratio and in some cases to prevent
Figure 2-7: Transformation ratio as a function of gap height for various plate thicknesses. From Equation (2.13) and the definition of $n$ in 2.2, the transformation ratio is independent of radius near the pull-in voltage.

the compromise of dielectric strength.

Discussion up to this point has focused on the CMUT operation in receiving mode where the sensitivity is key given the weak strength of the incoming signal because of attenuation. The strength of the output ultrasonic pulse in terms of acoustic power, energetic efficiency and sensitivity in transmission is an integral part of an effective transducer. Transmit sensitivity is defined as the pressure generated per applied voltage. It should be noted that this definition is the inverse of the receiving sensitivity in units only and not in actual expression.

As with receiving sensitivity, the transformer ratio is related to the acoustic pressure output but is not the most significant factor that affects transmit capability. The output force is proportional to the transformer ratio multiplied by the applied voltage, and as is the case in receiving, a high transformer ratio results in high transmit sensitivity. Since the transducer operates in a reciprocal system, the definition of the transformer ratio is the same in both directions and results in the same dependencies and optimization in transmit as in receiving mode.
The transformation ratio is an important metric to understand the functionality of the transducer in its ability to transfer signals between the electrical domain. Although the ratio magnitude does matter, its value only varies by a factor of less than 3 for an appropriately selected range of gap heights (50-1000nm) and device thickness on the order of 1µm. This small variation for a wide range of values leads to less significant variation in the transformation ratio for a wide range of CMUT designs.

The CMUT is likely operated near pull-in to achieve the best receiving signal at a high DC bias voltage. At pull-in, the restoring force dependent on the membrane stiffness \( k_m \) and the electrostatic spring softening \( C_0/n^2 \) is maximized providing an additional benefit for high DC bias operation. Membrane stiffness competes with electrostatic spring softening to determine the total restoring force. At the smallest gap height and therefore the largest applied DC bias voltage, the electrostatic spring softening is minimized and the restoring force is solely dependent on the membrane stiffness. Although maximized near pull-in, the restoring force is still constrained by geometric parameters. Membrane stiffness is increased with a smaller plate radius and thicker plate thickness. Larger restoring force has been experimentally linked to higher transmit sensitivity [28], so the plate radius should be kept small and the plate thickness should be maximized for optimal restoring force and higher transmit sensitivity.

In transmission, the acoustic pressure output and electrical energy efficiency are most important. For simplicity, the parasitic or shunt capacitance will be considered the most significant energy loss between the electrical and mechanical domains. Each time the transducer is actuated, the shunt capacitor must be charged in parallel to the acoustic output and energy is lost to this charging. The transmit energy efficiency \( \eta \) is related to the acoustic energy \( U_{ac} \) output and parasitic capacitance energy \( U_e \) as follows:

\[
\eta = \frac{U_{ac}}{U_{ac} + U_e} \quad (2.14)
\]

with

\[
U_e = \frac{1}{2} C_0 V^2 \quad (2.15)
\]
and

\[ U_{ac} = \frac{\omega^2 \pi a^4 A^2 \rho_m}{z} \exp \left( \frac{j \omega z}{c_m} \right) \]

where \( A \) is the average plate displacement, and \( z \) is the measurement distance from the plate. The acoustic energy equation is based on an equivalent rigid piston model, which has shown to be reasonably accurate for calculating CMUT acoustic pressure [37]. For subsequent calculations, the average displacement was determined based on the pull-in displacement and plate deflection equation, and the distance from the plate was kept small. Assuming a CMUT operates near pull-in voltage and the maximum deflection is near the pull-in deflection, the energy efficiency is on the order of 1% and decreases with higher pull-in voltages. This presents a challenge for CMUT operation. At high pull-in voltage, the CMUT is more sensitive in transmit and receiving mode, but actuation is incredibly energy inefficient.

Unfortunately the geometric factors that act to maximize the device bandwidth, sensitivity, restoring force, and transformation ratio create the CMUTs limitations. A CMUT with the ideal parameters described above including a small plate radius, large device thickness

\[ \text{discussion with Kailiang Chen, graduate student in electrical engineering advised by Professor Sodini} \]

62
operating near the pull-in voltage requires a significant DC bias voltage. In addition to requiring a large amount of power, a high DC bias voltage creates a large parasitic capacitance that lowers the energy efficiency when driving the system. Small gap height constrains the deflection of the CMUT membrane in transmit mode, and since deflection scales directly with acoustic pressure, the output pressure capability is thus limited. Fabrication challenges are also introduced for the small gap height requirement, evaporation after a wet etch creates stiction that can permanently adhere the membrane to the substrate surface [23]. For operation, the ideal CMUT should have a small plate radius, large thickness, and small gap height to maximize sensitivity, signal and output force. However, these parameters require a large amount of power, reduce efficiency, limit the acoustic pressure output and introduce fabrication challenges.

From the described effective transducer metrics, it is clear that the geometric constraints of the CMUT are imposed by factors related to electrostatic transduction. As such, the PMUT is not subjected to the same restraints and can likely be more energy efficient, require less power and be made larger. A larger size is advantageous for high acoustic pressure, which scales with the cube of the plate radius. With this scaling, a PMUT double the size of an existing CMUT would be capable of producing eight times the acoustic pressure. Additionally, the sensitivity and energy efficiency are independent of the bias voltage, so the PMUT could transmit and receive signals at a low bias without compromising performance.

2.3 Piezoelectric Micromachined Ultrasonic Transducers (PMUTs)

2.3.1 Overview of Previous PMUTs

The PMUT is a promising alternative to the existing CMUT technology since it does not suffer from the same limitations. Based on the previous discussion, the CMUT is limited to a small plate radius (≤50m), device thickness on the order of 1mm, a large bias voltage and a small gap height. Since a reduction in gap height results in high sensitivity, the smallest
possible gap height is ideal for CMUT design. In most cases, the gap height is determined by limitations related to fabrication. In PMUT designs, a gap is not necessary for transduction and as a result, a larger deflection is possible resulting in a higher potential acoustic pressure output. A large DC bias voltage is also not necessary for operation, and the plate radius is not constrained by factors related to electrostatic transduction. With few constraints on geometry and required voltage, the PMUT has the potential to improve the capabilities of the micro-machined ultrasonic transducer.

Some previous work has been completed based on PMUT design but leaves room for further improvements and design innovation. In contrast to the CMUT, the breadth of previous PMUT work is sparse with a limited number of publications over the past two decades. Unlike with the CMUT, it appears that PMUT research has yet to gain traction, which can likely be attributed to modeling difficulty and design sensitivity to residual stress, which results in significantly less than ideal performance. With a fresh approach to PMUT design and a better modeling capability, it is possible to circumvent these critical issues in future PMUT designs. First, it is important to learn from the previous designs and understand their limitations before proposing PMUT design improvements.

PMUT designs are largely based on a diaphragm configuration similar to that of the CMUT. Publication of one of the first PMUTs in 1990 determined that the diaphragm configuration resulted in less electrical capacitive loss to the surrounding structure. This initial work capitalized on the then recent popularity of polymer-based piezoelectrics using a thin sheet of PVDF as the main component of its diaphragm structure [41]. Subsequent publications focused on the PMUT array interface suggesting that significant crosstalk reduction was possible by placing PMUT devices on separately excited diaphragms [42]. Although this first PMUT was innovative for its time, it suffered from design and fabrication limitations that prevented larger scale adoption. The piezoelectric constant of the polymer thin film is much smaller than those of PZT, ZnO and AlN now commonly used for piezoelectric MEMS applications. Additionally, this film had to be separately processed and bonded to the front of the wafer in a manner largely incompatible with CMOS processing, and could
not be geometrically patterned. Despite these setbacks, the diaphragm design has largely been adopted in newer PMUT and CMUT designs possibly as a result of this previous work.

Although largely based on the diaphragm, subsequent PMUT designs have shown some improvements based on fabrication technology and clever design adaptations. Primary design changes have included relaxation of the clamped edge boundary condition [44], [26], and optimized electrode coverage [44], [16]. Boundary condition changes have focused on realizing a simply supported case, which is most closely achieved by removing material along the edge of the plate [44] or flexurally suspending the membrane on a few supports [26]. For the flexurally suspended case, improvements in coupling coefficient and acoustic pressure output nearly twice that of the clamped case have been reported [26]. Improvements to the coupling coefficient have also been proposed based on the electrode area coverage. Although electrode coverage is largely geometry dependent, in most cases, coverage of the entire device area leads to limited deflection, and alternative configurations are chosen. For the circular diaphragm, optimized coupling was observed with approximately 60% electrode coverage based on experiment and derived analytical and finite element models [44], [16]. Qualitatively, this is attributed to the change in sign of gradient of the deflection during the first mode of plate bending, which occurs at approximately 60% of the plate radius [44].

Inconsistent results related to bandwidth and device performance largely deviated from model predictions have been reported for most device designs. As previously discussed,
thin plate vibration devices have tunable acoustic impedance based on geometry and can 
be designed to impedance match the imaging medium. Impedance matching greatly reduces 
insertion loss and increases bandwidth, which can be larger than 100% as reported in CMUT 
literature [35]. Since the PMUT operation is based on a similar plate design, impedance 
matching and the resulting bandwidth improvement should be observed in PMUT device. 
Modeling of a rectangular plate PMUT with a smaller center electrode suggests that a 
bandwidth of over 100% should be possible [5], but experimental values have been much 
smaller [18], [46], even smaller than those of existing commercial bulk piezoelectrics. For a 
circular diaphragm design, Muralt et al do not report a bandwidth but strongly suggest that 
the PMUT should only be used in applications where large bandwidths are not required [44].

Non-optimal device performance and deviations from model predictions are most com-
monly attributed to problems with residual stress, but an appropriate means of alleviating 
this issue has yet to be implemented. Compressive residual stress can cause buckling, which 
cripples device performance. Since most devices are fabricated on SOI wafers with the buried 
oxide acting as an etch stop in the final membrane release step, a compressive oxide film is 
usually responsible for this buckling. Some authors have removed this oxide layer or added 
tensile thin film layers, such as silicon nitride, to balance the stress [41]. If residual stress 
does not cause buckling, it still hampers device performance, and membrane vibration is very 
sensitive to any residual stress. Although efforts have been undertaken to balance residual 
stress, high device sensitivity to any residual stress creates a reduced coupling coefficient and 
even potential bandwidth problems.

Since the effect of residual stress is felt largely along the edge of the plate, the simply 
supported boundary condition relaxes some of these residual stress issues, which can partially 
explain the improved device performance observed for the partially suspended devices. Other 
authors have added a bias voltage to overcome the effects of residual stress, but this requires 
the PMUT to operate at high voltages similar to those required for the CMUT [44], [16]. 
For significant acoustic pressure output improvement, the voltage has even been driven past 
the coercive voltage where the initial domain switching has enabled transmission sensitivity
as high as 20kPa/V at a distance of 20mm [19]. In the absence of a large applied DC bias voltage, acoustic power output is limited and transmit sensitivity is low [5]. An unbearably small transmit sensitivity has forced some authors to use PMUTs exclusively for signal reception operating a separate device for transmission [9].

2.3.2 Fabrication

Most current PMUT devices are micro-fabricated using largely similar techniques. The 1990 POSFET (PVDF-MOSFET structure) was one of the first attempts to apply the concepts of micro-fabrication, mainly CMOS processing techniques, to the production of small-scale piezoelectric ultrasound devices [41], [42]. Research and improved deposition of thin film piezoelectric materials like ZnO (zinc oxide), PZT, and AlN (Aluminum nitride) have changed PMUT device fabrication and capabilities. Despite these fabrication material improvements, most of the current PMUT processing steps are similar to those of the early POSFET.

Current PMUT fabrication generally follows two different fabrication approaches based on the starting substrate material. In either case, the process begins with an insulator addition to the silicon substrate via thermal oxidation, PECVD of oxide [26], [53], or LTO of sacrificial material [9], [16] and LPCVD silicon nitride [46], [41]. In the POSFET diaphragm design and earlier devices [5], [46], [16], [9], [29], the next step is boron doping of p-type or n-type silicon wafers that is later used as an etch stop for a final anisotropic wet etch with ethylenediamine-pyrocatechol-water-pyrazine [5], [46], [16], [9] or DRIE [29] to release the membrane structure. For the substrates coated with only oxide, photolithography on the front side oxide and etching forms the pattern used for the boron doping and drive-in anneal [16], [9]. After boron diffusion, the oxide layer is removed and LTO is again deposited on the substrate serving as an electrical insulator and etch stop during further processing. For a substrate initially coated with LTO and LPCVD nitride, the nitride has been patterned and etched to form the mask for boron doping and drive-in, and then cleared for further thermal oxidation and LPCVD of nitride and oxide [41]. Otherwise, the sacrificial layers
were kept until after electrode and/or piezoelectric material deposition to serve as an etch stop for the final membrane release [46], [29].

More current devices begin with an SOI wafer having a device thickness between 2 and 15m with a buried oxide layer ranging from 0.5 to 1m [18], [29], [62]. In this case, the final membrane release step is an anisotropic wet etch with ethylenediamine-pyrocathecol-water [62] or deep reactive ion etching (DRIE) [44], [18] using the buried oxide as an etch stop.

Intermediate processing steps are similar for both procedures and are largely dependent on the thin film piezoelectric material chosen for the device. After oxide addition, the bottom electrode is deposited onto the substrate using electron beam evaporation or sputtering. A thin adhesion layer (chromium, titanium or titanium tungsten) and electrode material (gold or platinum) are sequentially deposited and sometimes patterned via lift-off. In one device operating in $d_{33}$ mode, the bottom electrode was not necessary and fabrication immediately proceeded to spin coating prepared sol-gels of zirconium oxide and PZT with pyrolyzation and annealing steps occurring between each deposition [29]. For the $d_{31}$ devices after electrode deposition, PZT was spin coated, pyrolyzed and annealed on top of the platinum electrodes; patterning of the PZT was performed with photolithography and etching with a solution of H2O:HCl:HF [44], [5], [18], [16], [9], [62], [60]. AlN was sputtered onto platinum electrodes and later patterned and wet etched [26], [53]. For the piezoelectric ZnO, the gold electrode was patterned with a wet etching process and ZnO was sputtered onto the electrode surface and later patterned with a wet etch [46]. Electrode and piezoelectric material were paired according to crystallographic orientation compatibility.

Top electrodes were deposited in a similar manner with sputtering or e-beam evaporation and patterned with a lift-off technique or subsequent wet etching. In each case, the general final device is similar to the simplified schematic in Figure 1, and is completed with appropriate dicing and packaging based on experimental requirements. One and two-dimensional arrays were bonded onto chips and sealed for imaging experiments [18], [46], [9]. For devices incorporated into arrays, it is important to maintain the element pitch below the acoustic
wavelength in the imaging medium to prevent noticeable grating lobes [18], [60]. As such, DRIE is most convenient for membrane release because anisotropic wet etching can thin the material separating devices at such a small pitch [18].

Before packaging, some authors chose to pole their d31 mode PZT-based devices [5], [16], [60]. Although it is not entirely clear why poling is selected, poling has been outlined as a design parameter in one case [5]. As previously discussed, recent work links domain switching at applied voltages greater than the coercive voltage with improved acoustic pressure output [18], [19]. As such, poling orientation can influence device performance yet the potentially positive effect of initial poling is still unclear. For the $d_{33}$ mode device, poling was used to change membrane tension to alter displacement characteristics [14] and reduce the effects of residual stress, which will be discussed in the next section.

### 2.3.3 Device Characterization

Experimental data is largely used in PMUT characterization and is usually compared with some analytical modeling. However, with a wide variety of experimental characterization and a limited number of analytical models, it is difficult to compare many PMUT devices. Although experimental techniques are similar, direct comparison is not possible in most cases. For an array and sometimes individual devices, it is common to collect transmit and receiving sensitivity data, bandwidth, real and imaginary impedance plots, resonant frequency and deflection data. However, in most cases different imaging media are used in experiments including air, vacuum, water and various oils, and sensitivity is collected at a variety of distances using different units and reference pressures.

Since an analytical expression for the equivalent circuit of a unimorph PMUT has yet to be derived, impedance data cannot be matched with explicit expressions. Impedance or admittance data usually collected with an impedance analyzer is commonly reported [44], [5], [46], [53], [29], [60]. In some cases, the impedance data is substantiated with limited analytical modeling [44], [5], [53], [29], [60], or finite element modeling for a particular design case is used for comparison [45], [46], [16]. Since deflection, resonant frequency, and coupling
Coefficient are more conveniently computed, most PMUT literature relies on these metrics for reporting data.

For the thin plate PMUT design, classic plate theory (also referred to as Love-Kirchhoff analysis) has been used to determine equations for deflection and resonant frequency [26], [18], [46], [53], [29]. Explicit equations for resonant frequency based on circular or rectangular plate geometry are readily available in literature and can be extended to include loading effects from the acoustic medium. Thus, it is most common to see resonant frequency comparisons between analytical and experimental data, which show reasonable agreement.

Table 2.3: Resonant frequency equations and error between analytical and experimental solution [29], [44], [53], [62], [18], [60], [16]. Variables defined as follows: flexural rigidity $D$, tension $T$, plate radius/square side length $a$, first mode circular constant $\alpha$, first mode square constant $\beta$, membrane mode $m$, plate mode $p$, plate density $\rho$, plate thickness $h$.

<table>
<thead>
<tr>
<th>Design</th>
<th>Publication</th>
<th>Resonant Frequency</th>
<th>Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circular</td>
<td>Hong et al</td>
<td>$f_{m+p} = \frac{1}{2\pi a^{2}\rho h} \left( \frac{\alpha^2}{a} D + \alpha_m^2 T \right)$</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>Muralt et al</td>
<td>$f_p = \frac{\alpha_p}{2\pi a^2} \sqrt{\frac{E}{\rho h}}$</td>
<td>25*</td>
</tr>
<tr>
<td></td>
<td>Shelton et al</td>
<td>$f_{m+p} = f_p \left( 1 + \frac{a}{\alpha_p} \sqrt{\frac{T}{D}} \right)$</td>
<td>-</td>
</tr>
<tr>
<td>Square</td>
<td>Yamashita et al</td>
<td>$f_p = \frac{\beta_p}{2\pi a^2} \sqrt{\frac{E}{\rho h}}$</td>
<td>19.7</td>
</tr>
<tr>
<td></td>
<td>Dausch et al</td>
<td>$f_p \approx 5$ to $&gt; 100$ for large length</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Wang et al</td>
<td>Simulated</td>
<td>3.1</td>
</tr>
<tr>
<td>Rectangular</td>
<td>Cho et al</td>
<td>$f_{m+p} = \frac{1}{2\pi} \sqrt{\frac{k}{m}}$</td>
<td>-</td>
</tr>
</tbody>
</table>

*Comparison between clamped equation and suspended plate experimental data

**Based on lumped parameter model

The derivation of a deflection equation especially when considering the effects of a piezoelectric field is more complex and therefore less commonly reported; however, explicit solutions have been derived based on electrode area coverage using piece-wise [46], [45] and energy-based [44], [53] solution techniques. The deflection equation is particularly useful to determine the effect of residual stress on device performance. An applied tension term, which in the case of the PMUT is analogous to the residual stress, can be added to the governing plate vibration equation and incorporated into the deflection solution providing a means to estimate the residual stress in the structure. With an energy-based approach, Muralt et al.
extend their plate solution to define the coupling coefficient through minimization principles relating piezoelectric coupling, elastic and electrical energy. This derivation agrees well with experimental data, and is used to determine the optimum electrode area coverage based on maximum coupling. Muralt et al define this optimum area coverage as an electrode that starts or ends at 65% of the total radius [44].

Experimentally, the electromechanical coupling coefficient $k_c$ is determined from the resonant $f_r$ and anti-resonant $f_a$ peaks of the imaginary impedance by the following equation [44], [5], [16], [53], [60]:

$$\frac{k_c^2}{1-k_c^2} = \frac{f_r^2 - f_a^2}{f_a^2}$$

(2.17)

Since it is simply calculated from experimental data, the coupling coefficient is commonly used in device characterization and is most consistently defined. As a result, the coupling coefficient will be used to compare a variety of PMUT designs. In most cases, finite element modeling was also used to determine a maximum achievable coupling coefficient [44], [5], [16], [53] or for Muralt et al, this maximum value was computed with the analytical model. As previously mentioned, residual stress can have a grave impact on device performance and the large difference between the experimental and maximum achievable coupling can be attributed to this stress. A variety of voltage driving and biasing schemes have been used in an attempt to overcome the effects of residual stress and to improve device deflection and output acoustic pressure. For a $d_{31}$ mode device, driving with a high peak-to-peak voltage generated from an AC pulse has been used to improve acoustic pressure output, where an acoustic pressure of 340 Pa/V, was measured for a single device with 60V applied.
peak-to-peak [18]. Alternatively, the application of a DC bias voltage has been successfully implemented to increase the coupling coefficient [44], [16]. The DC bias voltage has a more significant effect on high stress as compared to low stress membranes where coupling improvements of 150% have been observed [16]. Despite this improvement, maximum coupling coefficients on the order of 10% are still predicted Table 2.4, which is low compared to the potentially > 100% coupling of the CMUT [35].

For further improvement in coupling, operation in $d_{33}$ mode has been suggested [5] and implemented [29]. For a poled $d_{33}$ mode PZT device, initial biasing and the application of a reverse bias larger than the coercive voltage created non-180° domain wall motion that increased membrane tensile stress and decreased residual stress. However, this change in residual stress was small even at high voltages, measuring 8MPa for a reverse biasing > 25V. Domain switching as a result of an applied voltage greater than the coercive voltage increased membrane displacement but was largely non-linear with increasing electric field [29].

Creative voltage operation techniques have improved PMUT device performance, but existing residual stress significantly limits device coupling even at high applied bias and peak-to-peak voltages. Alternate operation in $d_{33}$ mode still requires high voltages for small changes in residual stress, and non-linear behavior at voltages exceeding the coercive field makes device behavior more difficult to predict. Other factors have been attributed to this compromised coupling including piezoelectric film defects, membrane-electrode alignment error, and parasitic capacitance [5], but these effects have yet to be comprehensively investigated.

Finally, deflection and mode shapes measured via laser doppler [26], [53] and scanning [29] vibrometry and stroboscopic white light scanning interferometry [44] have been used to characterize PMUT devices. For the $d_{33}$ mode device, residual stress and Young's modulus of the membrane are calculated from the deflection profile via curve fitting, and vibration modes were compared to frequency predictions. In addition to curve fitting of the deflection profile, the residual stress can also be determined from the resonant frequency with a modified frequency equation that combines plate and membrane mode contributions. This equation
and associated data fitting is useful in both $d_{33}$ and $d_{31}$ mode devices to determine whether the plate deflection is governed by plate or membrane vibration.

Besides mode shape and frequency data, $d_{31}$ mode devices with partially released edge boundary conditions have been analyzed through the deflection profile. Added deflection near the edges [44] and a piston-like displacement [26] validate the relaxed edge boundary condition. In each case, improvement of acoustic pressure output [26] as previously described, and coupling coefficient [44] were observed. Of reported PMUT coupling data, the highest experimentally measured coupling occurred with a partially suspended device fabricated by Muralt et al Table 2.4.

Analysis is not limited to deflection, coupling coefficient and resonant frequency; however, the availability of analytical models and measurement ability of these metrics are convenient for reporting. Other parameters including receiving and transmission sensitivity, acoustic pressure, quality factor, phase, and transmission and response pulse forms are reported, but are difficult to compare because analytical models have yet to be determined and reporting is not consistent. Further analytical modeling will give deeper meaning to this experimental data and provide a larger basis for comparison.

### 2.3.4 PMUT Potential for Innovation

Previous PMUT work shows devices that often fall short of expectations in the mechanical and electrical domains whether measured via coupling coefficient, bandwidth, output acoustic pressure, etc. Although this can be discouraging, new design concepts and more detailed analytical modeling can improve device performance and provide a means to analyze data and optimize designs. Recent work with optimized electrode coverage and a partially suspended boundary condition [44], [26] has already made significant gains in device performance. Innovative device driving methods based on modifying pulse waveform and high peak-to-peak voltage have already been used to generate 3D images [19]. Although PMUT device performance has been limited, the PMUT has potential, that has largely gone unrecognized, to transform ultrasonic imaging.
Modeling advances especially pertaining to the equivalent circuit and deflection will be critical for optimized device performance. A deflection solution based on the Greens function is introduced for the case of an arbitrary number of circular/ring electrodes. The resulting equation explicitly defines deflection for a large number of potential geometries and for the case of one electrode, agrees with energy-based and piece-wise solutions [44], [26], [45], [46], [53]. Since it can be used for a wide variety of designs, the Greens function solution will be a powerful tool for deflection optimization. Further development of an equivalent circuit model based on the bimorph and unimorph design is also possible and underway. With analytical expressions for circuit parameters, quality, transformation, electrical and mechanical impedance can be analytically determined and compared to experimental values. With analytical predictive capability, the model development will also enable optimization before experiment leading to greatly improved device performance.

In comparison to the CMUT, the PMUT is less limited by geometric and design constraints. Based on design goals outlined in the previous section, optimal CMUT performance is achieved with a small plate radius, large thickness and small gap height. By controlling these geometric parameters, high sensitivity in transmission and receiving of the CMUT is achievable. However, these constraints in gap height and plate radius limit the volumetric displacement, which is directly proportional to acoustic pressure output. There is no limit on PMUT device deflection or radius and thus no limit on output acoustic pressure. As a result, an array could include large PMUT elements capable of high acoustic pressure output. Electrostatic actuation and geometry requirements also limit the breadth of potential CMUT designs. Since high sensitivity requires a sealed cavity with a high electric field and small gap, there are few possibilities for device functionality outside a plate or membrane mode CMUT with a gap for electrostatic operation.

A commonality between all PMUTs is a design largely based on a clamped diaphragm operating in either membrane or plate mode. However, the PMUT is not limited to a clamped plate configuration by its transducer scheme, which creates many design opportunities not available for the CMUT. Other designs have already been introduced including an
inter-digitated xylophone device and piezoelectric coated pillar structure, and experimental characterization is underway [40]. Creative thinking could be key for the best PMUT performance, moving away from a clamped diaphragm configuration can alleviate the challenges introduced by residual stress, and optimize overall device performance. Focus does not need be on improving the existing PMUT design but can be shifted to changing it for optimal functionality. With the proper analytical tools and a new design perspective, the PMUT can be optimized to reach its full potential.
Chapter 3

Design and Modelling

3.1 Design Framework

Based on the review of previous work on existing CMUT and PMUT designs, it is clear that great progress has been made, but challenges are still faced. Particularly, a high voltage requirement and gap-limited, small deflection prevent the CMUT from being used in high acoustic pressure applications. Although the PMUT is not deflection or voltage limited, high sensitivity to residual stress, especially when driving the device, has severely limited transmit bandwidth forcing some designers to turn to high applied voltages [18] or to abandon the use of the PMUT entirely in transmit mode [9].

In evaluating previous PMUTs and CMUTs, high acoustic power in transmission, low power consumption, high sensitivity, and a resonant frequency appropriate for general medical images (between 1 and 16MHz) are considered the most important factors based on medical ultrasound imaging requirements [56]. For appropriate design of a new small form factor, high resolution medical ultrasonic transducer, these must be the top level functional requirements for device performance. Although these requirements might seem simple, their coupling and the transduction and plate bending governing equations related to the PMUT device make the problem much more complicated. Beginning with the top level functional requirements, branching forming a functional hierarchy is constructed based on the more
detailed requirements of the PMUT design. This hierarchical approach makes the further design process less complex and naturally follows the hierarchical nature of the physical system [55].

From the functional hierarchy and the governing equations for the system, design parameters related to the functional requirements can be mapped in a similar functional hierarchy. Visual representation of these hierarchies are provided in Figures 3-1 and 3-2. Construction of the branching diagrams enables the functional requirements to be related to design parameters at each layer, breaking down the design problem into multiple, smaller forms. When segmented, the mapping between the function requirements to design parameters is a more tractable problem. For each layer, the functional requirements and design parameters are related via the design matrix. Before displaying the first design matrix, it is convenient for notation to define the top level function requirements:

**FR11** = Operate at resonant frequency appropriate for medical imaging ($1 \leq f_r \leq 16\text{MHz}$)
**FR12** = Low power consumption (Voltage $\leq 10\text{V}$ per transducer)
**FR13** = High sensitivity
**FR14** = High acoustic power in transmission
Figure 3-2: PMUT design parameter hierarchy

and design parameters:

DP11 = Applied moment caused by piezoelectricity
DP12 = Plate radius
DP13 = Mechanical impedance of plate
DP14 = Volumetric displacement of plate

A decoupled design matrix can be constructed from the top level requirements as:

\[
\begin{bmatrix}
  FR11 \\
  FR12 \\
  FR13 \\
  FR14
\end{bmatrix}
= 
\begin{bmatrix}
  x & 0 & 0 \\
  0 & 0 & 0 \\
  x & x & 0 \\
  x & x & x
\end{bmatrix}
\begin{bmatrix}
  DP11 \\
  DP12 \\
  DP13 \\
  DP14
\end{bmatrix}
\]

(3.1)

When branched further to the second layer, the more detailed functional requirements are listed:

FR21 = Minimize residual stress
FR22 = Optimize deflection profile
FR23 = Reduce interelement crosstalk
FR24 = Maximize center deflection
FR25 = Receive signal over a wide bandwidth
FR26 = High quality (sharp resonant peak) in transmission

corresponding to the design parameters:

DP21 = Plate thickness (overall thickness of material layer stack)
DP22 = Plate boundary conditions
DP23 = Element packing density
DP24 = Electrode configuration (size and radius)
DP25 = Receiving frequency (i.e. 2f_r for harmonic imaging, etc.)
DP26 = Transmission frequency (should be close to resonant frequency of plate)

which yields the design matrix relation:

\[
\begin{bmatrix}
\text{FR21} \\
\text{FR22} \\
\text{FR23} \\
\text{FR24} \\
\text{FR25} \\
\text{FR26}
\end{bmatrix} =
\begin{bmatrix}
\times & 0 & 0 & 0 & 0 \\
0 & \times & 0 & 0 & 0 \\
0 & 0 & \times & 0 & 0 \\
\times & \times & 0 & \times & 0 \\
\times & \times & \times & \times & 0 \\
\times & \times & \times & \times & \times
\end{bmatrix}
\begin{bmatrix}
\text{DP21} \\
\text{DP22} \\
\text{DP23} \\
\text{DP24} \\
\text{DP25} \\
\text{DP26}
\end{bmatrix}
\]

Again, this matrix is decoupled, which is useful in undertaking the design process. From the design matrices, some interesting observations can be made about the design. For example, it is convenient to separately define the receiving and transmission frequency, which means that different designs could be developed for transmission and receiving mode transducers. However, the design outline is still flexible allowing for a single transducer to be designed to both receive and transmit. If a transducer is designed with a high enough bandwidth,
harmonic frequencies can be sensed with a single device that has a resonant frequency appropriate for transmission at the fundamental mode.

Many of the functional requirements and design parameters are chosen based on predicted (outlined in the next sections) and previous device performance. In the future, experimental results and design iterations might slightly alter the appropriate functional requirements and design parameters. The design hierarchy should be iteratively updated and decoupled for continued design improvements.

Although it might not be initially clear, the design parameters are carefully chosen based on the system governing equations, which will be described later, and lumped parameters like the piezoelectric moment that appear in many of these equations. Their definition is left intentionally general to avoid unnecessary added constraints and coupling to the design process. For example, thickness refers to the overall plate thickness, which will consist of multiple layers of piezoelectric, electrode and supporting material. The design parameter thickness encompasses each layer thickness without specifying a number of layers or a particular material that must be used. In the following sections, the design levels and their functional requirements and design parameters will be considered for the design process.

3.1.1 Determination of Design Parameters

In the design matrices, the array elements have been rearranged to produce a multiplication matrix that is lower triangular, which corresponds to a decoupled system. Design parameters can be defined beginning with the first array element. For example, in the second design matrix (equation 3.2), the residual stress is only related to thickness and can therefore be determined from the residual stress requirement. Once the thickness is defined, the boundary condition can be set by the desired deflection profile. All design parameters are defined in a similar fashion by proceeding element-wise down the arrays in equations 3.1 and 3.2.
Residual Stress and Piezoelectric Moment

For the first functional requirement, the measured thin film residual stress $\sigma_o$ will be related to the processing conditions and is assumed to be constant in both the radial and theta-directions. Since it is unrealistic to deposit films with no residual stress, the effects of residual stress should be minimized by an appropriate force and moment balance along the thickness of the plate. If each layer $i$ has the same radius, the tension $T^o$ (force per unit length) caused by the residual stress $\sigma_o$ is:

$$T^o = \int_{z_{i-1}}^{z_{i}} \sigma_o, dz = \sum_{i=1}^{n} \sigma_o,i h_i$$

(3.3)

where $h_i$ is the thickness of layer $i$. For a membrane, the thickness in the $z$-direction is small making the bending stiffness negligible. In this case, the membrane equation of motion is related to the tension, a force per unit length similar to the derived residual stress tension. When the membrane is perturbed, the restoring force that acts to return the membrane back to its taut, equilibrium position is only related to the tension since the bending stiffness is small. For a plate subject to residual stress, the system is acting as both a plate and a membrane; thus, the governing plate and membrane equations can be superposed to determine plate deflection. The residual tension in equation (3.3) becomes important in the overall deflection equation of a stressed plate. From this qualitative explanation, it is clear that the tension should be set equal to zero for the plate equation to be valid and for optimal deflection behavior.

Based on the stressed plate equation, the critical stress can be derived. The critical stress will be negative or compressive and when reached, the plate will buckle losing all load bearing capacity. It is especially important that the residual stress is minimized to avoid buckling, so the critical stress should be quantitatively determined and avoided in PMUT design. Buckling occurs when the dimensionless residual stress $v$ is:

$$v = \frac{T^o}{T^*} = -1$$

(3.4)
where $T^*$ is the magnitude of the critical residual stress. In general for a circular plate with radius $a$ and flexural rigidity $D'$, 

$$T^* = \frac{\kappa D'}{a^2}$$

(3.5)

where $\kappa$ is a numerically determined constant based on the boundary conditions. For a circular plate with a clamped edge, $\kappa = 14.68$ [59]. For the PMUT transducer once the critical stress is exceeded, the plate will no longer vibrate, and ultrasonic waves cannot be generated in transmission or sensed in receiving mode. The device will fail to function.

When examining the effects of residual stress, it is easy to stop here since the stressed plate equation can be determined from equation (3.3) and the critical stress from equation (3.5). However, moment imbalance about the neutral axis $z_N$ might also cause buckling [27] and for lesser imbalances could adversely affect deflection. The residual moment $M^\sigma$ is calculated through the thickness of the plate for an arbitrary number of plate layers $n$:

$$M^\sigma = \int_{z_{i-1}}^{z_N} \sigma_{o,i} z dz = \sum_{i=1}^{n} \frac{\sigma_{o,i}}{2} \left[ (z_i - z_N)^2 + (z_{i-1} - z_N)^2 \right]$$

(3.6)

It should be noted that the distance $z_i$ of the $i^{th}$ layer from the reference axis of $z = h_0 = 0$ is different from the thickness of each individual layer. The individual layer height $h_i$ and the distance from the reference $z_i$ are directly related via summation $z_i = \sum h_i$, which is qualitatively shown in Figure 3-3.
Disregarding the effect of residual stress, the height of the mechanical neutral axis for the layer stack of isotropic material is given by [21]:

\[
z_N = \frac{\sum_{i=1}^{n} Y_i(z_i^2 - z_{i-1}^2)}{2(1-\nu_i^2)}
\]

(3.7)

By minimizing the residual moment and tension about the neutral axis, the negative effects of residual stress including buckling and compromised device performance can be averted. With expected residual stress values from [27], equations (3.3) and (3.6) can be minimized, and the appropriate thickness of each layer can be selected. In previous PMUTs, the clamped plate designs were particularly sensitive to residual stress, especially during transmission, resulting in severely compromised device performance. It is therefore crucial that one of the first design considerations is reducing the effects of this stress.

Similar to defining the thickness based on the residual stress, the voltage requirement is only dependent on piezoelectric moment and can thus the voltage requirement alone can be used to define the appropriate piezoelectric moment. As mentioned previously, the piezoelectric moment is a lumped parameter related to the layer thicknesses, neutral axis location, voltage and the piezoelectric and material properties of PZT. The piezoelectric moment is only present in transmit mode when a voltage is applied across the plate causing deformation. A radial tension is caused by the applied voltage that acts along the center of the PZT layer resulting in an applied moment about the neutral axis. The explicit derivation of the piezoelectric moment is provided in later sections. For now, the definition is introduced based on this physical interpretation. Assuming the piezoelectric material is located at layer \(i\), the piezoelectric moment \(M'_p\) is:

\[
M'_p = \frac{Y_{PZT}d_{31}V}{1-\nu_{PZT}} \left( h_{PZT}/2 + z_{i-1} - z_N \right)
\]

(3.8)

where \(d_{31}\) is the piezoelectric constant. Since the piezoelectric moment is linearly related to voltage, a larger applied moment and thus force occurs with increased voltage.
Boundary Conditions and Resonant Frequency

Again because of independence from other design parameters, the deflection profile requirement can be used to specify the necessary boundary conditions. This could be done initially since the profile is not coupled with any other parameters besides boundary condition; however, based on the convention of the design matrix, it is convenient to set the profile after the piezoelectric moment. Deflection profile is related to the output acoustic pressure because it directly defines the volumetric displacement rate. The direct relation between volumetric displacement and acoustic pressure output will be shown in later sections. For a given radius and fixed center displacement, volumetric displacement is maximized when the profile is similar to a rigid piston. For a clamped configuration, volumetric displacement is smaller because the fixed boundary condition reduces displacement near the edges of the plate. For circular plates, the deflection profile is related to a constant for each mode shape that is numerically determined from the boundary conditions. This mode shape constant is part of the deflection equation and determines the resonant frequency. For now, the constant will be specified as $\gamma_{01}^2$, where 01 refers to the first mode. The determination of this value from the boundary conditions for a clamped plate is shown in subsequent sections.

For a variety of boundary conditions, the $\gamma_{01}^2$ can be plotted and arranged by magnitude (Figure 3-4). The ideal case of a rigid piston displacement is approached for a plate with a free edge, which occurs when $\gamma_{01}^2 \approx 0$. As $\gamma_{01}^2$ increases, the deflection is assumed to be hindered by increasingly fixed boundary conditions. This assumption is consistent with the improved deflection that has been observed for a partially clamped plate [44], [26] and modeled for a simply supported as compared to a fixed plate [50]. The flexurally suspended boundary conditions realized in [44], [26] fall into the clamped along part of edge, simply supported along the rest case in Figure 3-4.

Although the free edge boundary condition is ideal, it is not physically realizable, as is the case with many other boundary conditions. When physically problematic boundary conditions are eliminated, some interesting boundary condition choices are revealed based only on $\gamma_{01}^2$. Simply supported at an arbitrary radius, a clamped edge with an added point
Figure 3-4: Values of $\gamma^2$ for various boundary conditions [10]. Boundary condition abbreviations specified in Table 3.1.

Table 3.1: Abbreviations for boundary conditions shown in Figure 3-4

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FE</td>
<td>Free Edge</td>
</tr>
<tr>
<td>CE</td>
<td>Clamped Edge</td>
</tr>
<tr>
<td>PMC</td>
<td>Point Mass at Center</td>
</tr>
<tr>
<td>UREL</td>
<td>Uniform Radial Edge Load</td>
</tr>
<tr>
<td>SSAR</td>
<td>Simply Supported at Arbitrary Radius</td>
</tr>
<tr>
<td>CC</td>
<td>Clamped Center</td>
</tr>
<tr>
<td>CPE</td>
<td>Clamped along Part of Edge</td>
</tr>
<tr>
<td>SSRE</td>
<td>Simply Supported along Remaining Edge</td>
</tr>
<tr>
<td>TSR</td>
<td>Torsion Spring Restraint</td>
</tr>
</tbody>
</table>

mass, and clamping at the center might result in improved deflection over the traditional fully clamped case and the flexurally supported devices in [44], [26]. $\gamma_{01}$ does not capture all of the factors related to deflection but can provide ideas for an optimum deflection profile.

The circular plate first mode resonant frequency $f_{01}$ is related to the plate radius $a$, the flexural rigidity $D'$ and rotary moment of inertia $I'$, which are related to the thickness, neutral axis height and material properties as will be derived in later sections.

$$f_{01} = \frac{\gamma_{01}^2}{\pi a^2} \sqrt{\frac{D'}{I'}}$$  \hspace{1cm} (3.9)

The above equation is also valid for higher order modes $(m, n)$ when the appropriate numerical value of $\gamma_{mn}^2$ are selected. At the first mode or fundamental frequency, vibrations
are axisymmetric, and the amplitude of oscillations is largest, making fundamental frequency mode actuation ideal for ultrasonic transduction applications. For transmission in typical medical ultrasound applications, the fundamental frequency range is shown in 3-1 [30]. If the system will be used in harmonic imaging, the receiving fundamental frequency should be ≈ 2f_{01} or the bandwidth requirement should be specified as sufficiently high for harmonic imaging using the same transmission device. Based on these guidelines, the plate radius can be determined either for a single device or two separate devices for transmission and receiving modes.

Unfortunately, loading from the fluid medium will shift the fundamental frequency, which will affect the ultrasonic transducer design. If the transducer is operated in air, this loading is small since the density of air is small. However, for an imaging medium like tissue, which has a density similar to water, density is larger and loading can cause a significant shift in the resonant frequency. An approximate equation for fluid loading with added mass \( m_a \) on one side of a plate with mass \( m_p \) is given by [10]:

\[
f_{01,\text{load}} = \frac{1}{\sqrt{1 + m_a/m_p}} f_{01}
\]  

(3.10)

Like \( \gamma_{01}^2 \), the added mass is numerically determined based on the geometry and boundary conditions. For a circular, clamped plate vibrating at its fundamental mode into a medium with density \( \rho_m \), \( m_a = 0.6689 \rho_m \pi a^3 \). Similar to loading, residual stress will also affect resonant frequency modifying the numerical constant \( \gamma_{mn}^2 \). The new values for \( \gamma_{mn}^2 \) based on the boundary condition and magnitude of the residual stress can be found in [10], and are again numerically determined.

Quality and Bandwidth

At the loaded resonance, the bandwidth and quality are both significant. Before continuing, it is important to specify the difference between resonant frequency and the operating frequencies in transmission and receiving. Vibration at the fundamental loaded frequency has an undamped system response that results in a quality that approaches \( \infty \) and band-
Figure 3-5: Fundamental frequency based on loaded and unloaded conditions and resonant frequency. Values of $\gamma_{01}^2$ taken from [10]

width close to zero. For a system transmitting at the fundamental frequency, an input signal will be transmitted indefinitely. Ultrasonic imaging is a series of short pulses separated by pauses. For short pulses to be possible, the quality should be high but not infinite; otherwise, ringing will significantly affect receiving signal quality. Similarly in receiving, high signal sensitivity occurs at large bandwidths. Too much excitation near the fundamental frequency will significantly reduce bandwidth and therefore sensitivity. As a result, it is common for the transmission frequency to be near the resonant frequency for a high quality and some ringing, but not at the resonant frequency where the quality is infinite and the ringing occurs indefinitely. Similar operation in receiving mode is also optimal.

Transmission quality can be approximated from existing literature related to plate vibration, transduction, and acoustics [17], [32], [43]. Quality can be defined in a variety of ways particular to the ultrasonic transducer such as a simply mechanical system and a mechanical system with acoustic radiation. Since loading is significant and the loaded definition is more thorough, the mechanical system loaded quality (henceforth referred to as quality) will be considered when optimizing transmission. For the ultrasonic transducer, the quality refers to the sharpness of the signal peak at resonance. Narrow peaks are of higher quality and
conversely, wide peaks are low quality. Quality is inversely proportional to bandwidth and as such, the bandwidth is low at high quality and vice versa. Considering the ultrasonic transducer in transmission mode, a narrow peak is ideal because it reduces the noise level of the receiving signal; however, the magnitude of the quality (i.e. narrowness of the peak) is limited by the damping required to reduce transmission ringing. From impedance curves, the quality $Q$ is derived based on the mechanical and acoustic reactances $X_m$, $X_a$ and resistance components $R_m$ and $R_a$ respectively.

$$Q = \frac{1}{2p_f} \frac{X_m + X_a}{R_m + R_a} \tag{3.11}$$

In the above equation, $p_f$ is the frequency parameter, a dimensionless number that relates the operating angular frequency $\omega = 2\pi f$ to the angular resonant frequency $\omega_{01}$ [32].

$$p_f = \frac{1}{2} \left( \frac{\omega}{\omega_{01}} - \frac{\omega_{01}}{\omega} \right) \tag{3.12}$$

The reactance and resistance values in equation 3.11 are the imaginary and real part of the impedance respectively. Mechanical impedance is defined analogously to electrical impedance with the flow term of velocity $u_w$ and force $F$ as effort [52]. For a harmonic excitation as frequency $\omega$, the pressure and velocity are related to the deflection as will be shown in later sections. To avoid the necessity of additional derivations, an existing model for the mechanical impedance $Z_m$ for a stressed plate subjected to uniformly distributed pressure load [38] will be used to estimated quality.

$$Z_m = \frac{F}{u_w} = j\omega I_0 \frac{ak_1k_2 [k_2J_0 (k_1a) I_1 (k_2a) + k_1J_1 (k_1a) I_0 (k_2a)]}{ak_1k_2 [k_2J_0 (k_1a) I_1 (k_2a) + k_1J_1 (k_1a) I_0 (k_2a)] - 2 (k_1^2 + k_2^2) J_1 (k_1a) I_1 (k_2a)} \tag{3.13}$$

$I_0$ and $I_1$ are hyperbolic Bessel functions of the first kind and $J_0$ and $J_1$ are Bessel functions of the first kind. $k_1$ and $k_2$ are constants related to material properties and previously defined parameters:
The mechanical impedance caused by acoustic radiation is separately calculated, again using a model found in literature \cite{43,17}. The deflection of the plate is explicitly described by a combination of Bessel functions, which complicates the determination of the radiated acoustic impedance in transmission. For simplicity, the device will be considered as a rigid piston that uniformly deflects at an average velocity determined by the more complicated deflection profile. This modelling approach is not uncommon \cite{17,37}, and in some cases includes the adjustment of the equivalent piston area for higher accuracy \cite{17}. As will be shown later, the equivalent piston model for a clamped, circular plate deflection proves to estimate the acoustic pressure with reasonable accuracy when the equivalent piston area equals the device area. Therefore, the actual device plate radius will be used in the equivalent piston impedance determination. From \cite{43},

\[ Z_a = R_a + jX_a \]  \hspace{1cm} (3.16)

with \( c = D'/I_0' \) and \( d = T^\sigma/I_0' \). For real inputs, the Bessel functions will always be real; however, complex inputs will create complex outputs for the first order Bessel functions. Since the radius input will always be real, \( k_1 \) or \( k_2 \) would have to be complex to yield a complex output. For \( k_1 \) or \( k_2 \) to be complex, \( c < 0 \), which is not physically possible as flexural rigidity and the moment of inertia are always positive, real values. Since \( k_1 \) and \( k_2 \) are always real, the Bessel functions will be real and the mechanical impedance is purely imaginary. With this model, the mechanical impedance will only introduce a reactance \( X_m = Z_m/j \) and the resistance \( R_m = 0 \). In reality, dissipative effects are part of the system, which could introduce a resistive component; however, for modeling purposes, this determination of reactance is sufficient.
\[ R_a = \pi a^2 \rho_m c_m \left[ 1 - \frac{c_m}{\omega a} J_1 \left( \frac{2\omega a}{c_m} \right) \right] \]  

(3.17)

\[ X_a = 4a^2 \rho_m c_m \int_{0}^{\pi/2} \sin \left( \frac{2\omega a}{c_m} \cos \alpha \right) \sin^2 \alpha d\alpha \]  

(3.18)

where \( c_m \) is the speed of sound in the imaging medium. In the approach to PMUT design, the quality in equation 3.11 should be considered specifically in transmission. When the PMUT is actuated, the applied voltage creates a piezoelectric moment (equation (3.8)) about the neutral axis, which modifies the governing plate equation and the derivations that follow. Mechanical impedance will also be affected by an electrically induced spring softening. For a bimorph PMUT, this spring softening is dependent on the thickness of the piezoelectric material and related, fixed piezoelectric and material properties [50]. However, for a multimorph configuration with more than one electrode, it is likely that the spring softening will depend on electrode radii, layer thickness and neutral axis location, but the exact equivalent circuit parameters have yet been derived. It is important to remember that the modelled mechanical impedance does not consider the effect of this piezoelectric moment and the spring softening, when in reality, it will affect the impedance and therefore the quality. However, to understand generally how the system behaves, the mechanical and acoustic load impedances from literature should be sufficient and it is assumed that the piezoelectric moment and spring softening will not greatly modify the system.

The transmission quality is estimated by substituting equations (3.18), (3.17) for the radiating acoustic impedance and (3.13) for the mechanical impedance into the quality equation (3.11). The frequency parameter in equation (3.12) is determined by the loaded fundamental frequency in equation (3.10). With all substitutions and simplifications complete, the quality near resonance can be approximated for compressive residual stress \( T^\nu < 0 \), tensile stress \( T^\nu > 0 \), and no stress \( T^\nu = 0 \) conditions.

As mentioned previously, the quality spikes to infinity at the resonant frequency, which is not ideal for actuation. In transmission, an operating frequency should be chosen near
the resonant frequency, where the quality is high, so it is more important to understand the behavior of the quality factor near resonance as opposed to at resonance. Approaching resonance, the quality is highly sensitive to operating frequency, so it is important that the proper operating frequency is selected. For a fixed quality factor, the appropriate frequency can be chosen based on the residual stress. A plate with tensile/compressive residual stress will transmit differently from a zero stress plate at the same operating frequency, and it is important to understand that similar functionality can be obtained by varying the transmit frequency based on stress conditions.

Alternatively, quality factor can be introduced in terms of bandwidth $\Delta \omega$:

$$Q = \frac{\omega_{01}}{\Delta \omega}$$

Thus, the bandwidth can be simply determined from the quality and resonant frequency. Another, more physical approach to understanding bandwidth is through comparing the mechanical and acoustic load impedance as outlined in [35]. For frequencies where the me-
Figure 3-7: Mechanical impedance based on residual stress with 3-dB bandwidth indicated by gray lines. Values shown are for a single layer silicon plate with \( a = 50\mu m \) and \( h_{Si} = 4\mu m \).

Mechanical impedance is small, nearly all of the load is absorbed by the acoustic mechanical impedance. At these frequencies, which determine the bandwidth, the circuit is most sensitive to acoustic loading. High bandwidth is ideal for receiving because the circuit is sensitive to a wide range of frequencies at which the mechanical impedance is small. With a large enough bandwidth \((\Delta f > 200\%)\), it is even possible to sense harmonic frequencies. Again, for the most accurate determination of bandwidth, spring softening should be added to the mechanical impedance for analysis. Unlike in transmission mode, there is no applied voltage across the piezoelectric layer in receiving mode, so the piezoelectric moment is identically zero and does not affect receiving bandwidth.

From the plot of mechanical impedance of the plate, the array receiving bandwidth can be roughly estimated neglecting spring softening. For a system operating in water, the acoustic impedance is \(1.5 \times 10^6 \text{N} \cdot \text{s/m}^3\), so the bandwidth will be determined by the frequency range for a mechanical plate impedance less than the water impedance denoted by the horizontal lines in Figure ??

The bandwidth is approximately 100\% for tensile stress, 125\% for no stress, and 150\% for compressive stress. Interestingly, the addition of compressive stress acts to increase bandwidth while tensile stress adversely affects receiving sensitivity. For
the CMUT, this bandwidth estimation is reasonably accurate. When low stress nitride is used as the membrane material \((T^* \approx 0)\), the measured fractional bandwidth is 123\% [35] corresponding well with the zero stress prediction.

By setting the layer thickness, piezoelectric moment, plate radius, and boundary conditions, the design parameter, center deflection can be set. Center deflection, especially in transmission, should be maximized to produce the highest output acoustic pressure, which is proven in later sections. For now, discussion of the center deflection and deflection profile will be neglected as subsequent sessions will derive explicit equations in great detail.

Despite the coupled nature of the PMUT design process, the design matrix can systematically guide PMUT design. It is easy to get lost in the details of the recent discussion of each design parameter and functional requirement, yet the top level design goals are not disregarded. The listed functional requirements and design parameters enable a high sensitivity, low power consumption, and high acoustic power device to be designed at the medical imaging frequencies based on fundamental equations either derived here or in previous transduction, acoustics or plate vibration literature. Although these governing equations do not cover the detailed transduction of the PMUT in great depth, they can provide rough guidelines and geometric parameters for operation that can be refined with more accurate modelling and experiment. Future PMUT design does not have to follow this exact approach and may not given discoveries of important factors that have not been considered in the above modelling. Design will be an iterative process, and this approach should be the starting point on which new PMUT designs can be based.

### 3.2 PMUT Plate Modeling

#### 3.2.1 Classic Plate Theory [48]

The design of a piezoelectric micro-machined ultrasonic transducer (PMUT) will utilize the vibration of a thin, circular plate coated with the PZT to transmit and receive acoustic pressure waves. A model PMUT geometry includes PZT sandwiched between top and bottom
platinum electrodes fixed to a silicon plate. Electrical insulation between the silicon plate and electrode is provided by an additional thin layer of thermal oxide. In receiving mode, the acoustic pressure mechanical loading on an array of micro-scale PMUT transducers will generate a voltage signal in the PZT to create an ultrasound image. In transmit mode, a high frequency voltage will be applied across the PZT layer to displace the silicon plate and generate an ultrasonic, acoustic wave.

The lateral dimensions of the proposed PMUT will be much larger than the thickness, classic plate theory is appropriate to describe the shape profile and vibration modes. In classic plate theory, the following assumptions are made:

1. Lines perpendicular to the mid-plane of a surface (transverse normals) are straight after deformation.

2. The transverse normals do not change in length.

3. Transverse normals remain perpendicular to the mid-plane after deformation.

Under the proposed assumptions, the displacement field $u=(u_r,u_\theta,u_z)$ can be derived for a circular plate with initially zero displacement along the mid-plane:

$$u_r(r,\theta,z,t) = -z \frac{\partial w}{\partial r}$$

$$u_\theta(r,\theta,z,t) = -z \left( \frac{1}{r} \frac{\partial w}{\partial \theta} \right)$$

$$u_z(r,\theta,z,t) = 0$$

where $w$ is the displacement along the $z$-axis. Plate deformation can be determined from the displacement field with the Green-Lagrange strain tensor. For moderate deformations, non-linear effects must be included in the equations for normal ($\varepsilon_{rr},\varepsilon_{\theta\theta},\varepsilon_{zz}$) and transverse ($\gamma_{r\theta},\gamma_{z\theta},\gamma_{rz}$) strain. With classic plate theory, these non-linearities are mostly negligible except for those involving $u_z$. The resulting strain equations are referred to as the von Kármán strains:

$$\varepsilon_{rr} = \frac{\partial u_r}{\partial r} + \frac{1}{2} \left( \frac{\partial u_z}{\partial r} \right)^2$$

95
Based on the displacements defined in 3.20, the normal $\varepsilon_{zz}$ and shear strains $\gamma_{z\theta}$ and $\gamma_{rz}$ are identically zero. The strain equations can be further divided into strains experienced along the surface mid-plane (membrane strains: $(\varepsilon_{rr}^{(0)}, \varepsilon_{\theta \theta}^{(0)}, \gamma_{r\theta}^{(0)})$) and bending contributions (flexural strains or curvatures: $(\varepsilon_{rr}^{(1)}, \varepsilon_{\theta \theta}^{(1)}, \gamma_{r\theta}^{(1)})$). In simplified form, the strain equations are written

$$
\varepsilon_{rr} = \varepsilon_{rr}^{(0)} + z\varepsilon_{rr}^{(1)}, \quad \varepsilon_{\theta \theta} = \varepsilon_{\theta \theta}^{(0)} + z\varepsilon_{\theta \theta}^{(1)}, \quad \gamma_{r\theta} = \gamma_{r\theta}^{(0)} + z\gamma_{r\theta}^{(1)}
$$

(3.24)

where:

$$
\varepsilon_{rr}^{(0)} = \frac{1}{2} \left( \frac{\partial w}{\partial r} \right)^2, \quad \varepsilon_{rr}^{(1)} = -\frac{\partial^2 w}{\partial r^2},
$$

(3.25)

$$
\varepsilon_{\theta \theta}^{(0)} = \frac{1}{2r^2} \left( \frac{\partial w}{\partial \theta} \right)^2, \quad \varepsilon_{\theta \theta}^{(1)} = -\frac{1}{r} \left( \frac{\partial w}{\partial r} + \frac{1}{r} \frac{\partial^2 w}{\partial \theta^2} \right),
$$

(3.26)

$$
\gamma_{r\theta}^{(0)} = \frac{1}{r} \frac{\partial w}{\partial r} \frac{\partial w}{\partial \theta}, \quad \gamma_{r\theta}^{(1)} = -\frac{2}{r} \left( \frac{\partial^2 w}{\partial r \partial \theta} - \frac{1}{r} \frac{\partial w}{\partial \theta} \right),
$$

(3.27)

PMUTs dynamically receive and transmit pressure waves during operation, so the above static plate structure analysis is limited. Instead, the potential and kinetic energy of displacement field should be considered using Hamilton’s Principle.

With Hamilton’s principle, the plate body is considered as a collection of material par-
articles subject to kinematic boundary conditions. During deformation, each particle travels along a path that varies with the virtual displacement \( \delta u \). Hamilton’s principle states that the optimal path of the material particles between time \( t_1 \) and \( t_2 \) occurs when the difference between the kinetic \( K \) and potential \( W \) energy over time is at an extremum.

\[
0 = \int_{t_1}^{t_2} \delta K - \delta W \, dt \tag{3.28}
\]

The integrand in the above equation is commonly referred to as the differential form of the Lagrangian function \( \delta \Pi \), and the evaluated integral expression forms the equations of motion for the system. Specifically for a circular plate, the Lagrangian energy function is determined by strain energy dependent on stress \( \sigma \), work done by external applied forces including the distributed load \( q \) and the foundation reaction force \( F_s \), and kinetic energy associated with plate displacement velocity \( \dot{u} = (\dot{u}_r, \dot{u}_\theta, \dot{u}_z) \). Hamilton’s equation thus takes the form:

\[
0 = \int_0^T \int_\Omega \int_{-h/2}^{h/2} \left( \sigma_{rr} \delta \varepsilon_{rr} + \sigma_{r\theta} \delta \varepsilon_{r\theta} + \sigma_{\theta\theta} \delta \varepsilon_{\theta\theta} + \sigma_{rr} \delta \gamma_{r\theta} \right) dz \, dr \, d\theta \, dt
- \int_0^T \int_\Omega \int_{-h/2}^{h/2} \rho \left( \dot{u}_r \delta \dot{u}_r + \dot{u}_\theta \delta \dot{u}_\theta + \dot{u}_z \delta \dot{u}_z \right) dz \, dr \, d\theta \, dt
- \int_0^T \int_\Omega (q + F_s) \delta w \, dz \, dr \, d\theta \, dt \tag{3.29}
\]

where \( \Omega \) denotes the mid-plane, and \( \rho \) is the plate density. It is convenient to define strain energy based on the normal forces \( N \) and moments \( M \) applied along the thickness \( h \) of an infinitesimal volume element.

\[
\begin{bmatrix}
N_{rr} \\
N_{r\theta} \\
N_{\theta\theta}
\end{bmatrix} = \int_{-h/2}^{h/2} \begin{bmatrix}
\sigma_{rr} \\
\sigma_{r\theta} \\
\sigma_{\theta\theta}
\end{bmatrix} \, dz \tag{3.30}
\]

\[
\begin{bmatrix}
M_{rr} \\
M_{r\theta} \\
M_{\theta\theta}
\end{bmatrix} = \int_{-h/2}^{h/2} z \begin{bmatrix}
\sigma_{rr} \\
\sigma_{r\theta} \\
\sigma_{\theta\theta}
\end{bmatrix} \, dz \tag{3.31}
\]

97
Since the plate experiences no transverse displacements, the kinetic energy can be simplified and written in terms of moments of inertia:

\[
\begin{bmatrix}
I_0 \\
I_2
\end{bmatrix} = \int_{-h/2}^{h/2} \rho \left[ \frac{1}{2} z^2 \right] dz = \rho \left[ \frac{h^3}{12} \right]
\]  

(3.32)

Euler-Lagrange equations are deduced from Hamilton’s principle by setting the \((\delta u, \delta v, \delta w)\) terms equal to zero separately. Since the only substantial displacement occurs along the thickness of the plate, the transverse virtual displacements can be neglected. The general form of Hamilton’s equation in the thickness direction becomes:

\[
0 = -\frac{1}{r} \left[ \frac{\partial^2}{\partial r^2} (rM_{rr}) - \frac{\partial M_{\theta\theta}}{\partial r} + \frac{1}{r} \frac{\partial^2 M_{\theta\theta}}{\partial \theta^2} + 2 \frac{\partial}{\partial r} \left( r \frac{\partial M_{r\theta}}{\partial \theta} \right) \right] \\
- \frac{1}{r} \left[ \frac{\partial}{\partial r} \left( rN_{rr} \frac{\partial w}{\partial r} + N_{r\theta} \frac{\partial w}{\partial \theta} \right) + \frac{1}{r} \frac{\partial}{\partial \theta} \left( N_{\theta\theta} \frac{\partial w}{\partial \theta} + rN_{r\theta} \frac{\partial w}{\partial r} \right) \right] \\
- q + F_s + I_0 \frac{\partial^2 w}{\partial t^2} - I_2 \frac{\partial^2}{\partial t^2} \left[ \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial w}{\partial \theta} \right) + \frac{1}{r^2} \frac{\partial^2 w}{\partial \theta^2} \right]
\]  

(3.33)

Given the material constitutive relations, the z-direction equation of motion can be writ-
ten in terms of materials properties and displacements for an isotropic material. From Hooke's law and with classic plate theory assumptions, the constitutive relations for a transversely isotropic material are:

\[
\begin{bmatrix}
\sigma_{rr} \\
\sigma_{\theta\theta} \\
\sigma_{r\theta}
\end{bmatrix} =
\begin{bmatrix}
Q & \nu Q & 0 \\
\nu Q & Q & 0 \\
0 & 0 & \frac{(1-\nu)}{2} Q
\end{bmatrix}
\begin{bmatrix}
\varepsilon_{rr} \\
\varepsilon_{\theta\theta} \\
\gamma_{r\theta}
\end{bmatrix}
\] (3.34)

where \( Q \) is the axial elastic stiffness coefficient, and \( \nu \) is poisson's ratio. For an isotropic material, the axial stiffness coefficient is:

\[
Q = \frac{Y}{1-\nu^2}
\] (3.35)

where \( Y \) is the Young's modulus of the material. PMUT will vibrate when electrical current flows through electrodes attached at the top and bottom of a PZT thin film fixed to a silicon plate. The thin plates of PZT and silicon that make up the PMUT design are both isotropic materials, and therefore the above constitutive relations apply directly to the design. The resulting bending stiffness \( D \) of an isotropic thin plate is constant along the plate thickness:

\[
D = \frac{Qh^3}{12}
\] (3.36)

For an adiabatic, isothermal process, the moment-deflection relationships for the isotropic circular plate can be written as:

\[
M_{rr} = -D \left[ \frac{\partial^2 w}{\partial r^2} + \frac{\nu}{r} \left( \frac{\partial w}{\partial r} + \frac{1}{r} \frac{\partial^2 w}{\partial \theta^2} \right) \right]
\] (3.37)

\[
M_{\theta\theta} = -D \left[ \nu \frac{\partial^2 w}{\partial r^2} + \frac{1}{r} \left( \frac{\partial w}{\partial r} + \frac{1}{r} \frac{\partial^2 w}{\partial \theta^2} \right) \right]
\] (3.38)

\[
M_{r\theta} = -(1-\nu) D \frac{1}{r} \left( \frac{\partial^2 w}{\partial r \partial \theta} - \frac{1}{r} \frac{\partial w}{\partial \theta} \right)
\] (3.39)

Using the Laplace operator and the Marcus moment simplification, the general equation
of the deflection can be derived from 3.33. Since \( I_2 = O(h^3) \) and the thickness \( h < 0.1a \) for a plate with radius \( a \), the rotary inertia \( I_2 \) is very small and can be neglected. The dynamic equation of motion can be written in the simplest, most general form as:

\[
D \nabla^2 \nabla^2 w + I_0 \frac{\partial^2 w}{\partial t^2} = q - F_s
\]

where the Laplacian operator is:

\[
\nabla^2 = \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2}
\]

This equation of motion considers an isotropic circular plate of a single, uniform material. With the applied boundary conditions (usually taken at the edge of the plate), the second order differential equation can be analytically solved for specified plate configurations. For the case of a piezoelectric plate, the constitutive relations for piezoelectricity must be considered in the derivation, and the resulting equation of motion has an added piezoelectric term that is dependent on the applied electric field and piezoelectric constant. For composite materials, the plate theory should be modified to consider the material properties of each layer and boundary conditions between layers.

### 3.2.2 Homogeneous Solution

From the derived equation of motion, plate vibration can be solved for the simple, homogeneous case of harmonic excitation. For the homogeneous solution, it is assumed that the plate is under no initial external load and there is a negligible external foundation reaction force. Equation 3.40 becomes:

\[
D \nabla^2 \nabla^2 w + I_0 \frac{\partial^2 w}{\partial t^2} = 0
\]

Under simple harmonic excitation with angular frequency \( \omega = 2\pi f \), the deflection \( w \) is assumed to take the form \( w = W(r, \theta) e^{i\omega t} \). The shape function \( W(r, \theta) \) describes the contour
of the plate during vibration, and can be inserted into the differential equation:

$$\left(\nabla^2 - \gamma^2\right) \left(\nabla^2 + \gamma^2\right) W = 0$$ \hspace{1cm} (3.43)

where \(\gamma\) is a constant defined as:

$$\gamma^4 = \frac{\omega^2 I_0}{D}$$ \hspace{1cm} (3.44)

In cylindrical coordinates, the plate equation of motion solution takes the form of a combination of Bessel functions and hyperbolic Bessel functions of the first kind. The Bessel function describes the power series expansion solution to an equation of the form:

$$\frac{d^2y}{dx^2} + \frac{1}{x} \frac{dy}{dx} + \left(1 - \frac{n}{x^2}\right)y = 0$$ \hspace{1cm} (3.45)

where the \(m^{th}\) order Bessel function is:

$$J_m(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{2^{k-m} k! (m+k)!} x^{2k+m}$$ \hspace{1cm} (3.46)

The hyperbolic Bessel function is related to the Bessel function as \(I_m(x) = i^{-m} J_m(ix)\), and provides the imaginary component of the shape function solution.

With the described Bessel functions, the general form of the natural vibration solution is:

$$W(r, \theta) = \frac{\cos(m\theta)}{\sin(m\theta)} \left(A J_m(\gamma r) + B I_m(\gamma r)\right)$$ \hspace{1cm} (3.47)

For a clamped plate of radius \(a\), the boundary conditions \(W(a, \theta) = 0\) and \(\left(\frac{\partial W}{\partial r}\right)_{r=a} = 0\) along the edges must be satisfied. The constants \(A\) and \(B\) are set by the applied boundary conditions:

$$B = -A \frac{J_m(\gamma a)}{I_m(\gamma a)}$$ \hspace{1cm} (3.48)

$$I_m(\gamma a) \frac{d}{dr} J_m(\gamma r) \bigg|_{r=a} - J_m(\gamma a) \frac{d}{dr} I_m(\gamma r) \bigg|_{r=a} = 0$$ \hspace{1cm} (3.49)

For the second boundary condition to be satisfied, \(\gamma\) must be a special set of real number
values that are specified by vibration mode shape \((m, n)\).

\[
\gamma_{mn} = \frac{\pi}{a} \beta_{mn}
\]  

(3.50)

where \(\beta_{mn}\) is a numerically determined constant also dependent on the vibration mode. Since \(\gamma\) depends on angular frequency, the above equation can be used to determine the fundamental frequencies of oscillatory modes in the system. It is common to uniquely define vibration modes by the number of radial modal lines \(m\) and lines \(n\) in the theta direction. From this definition, the vibration modes can be qualitatively described. In this way, \(\gamma\) can be written as \(\gamma_{mn} = (\pi/a) \beta_{mn}\), where \(\beta_{mn}\) is empirically tabulated and dependent on \(m\) and \(n\). The frequency \(f\) is dependent on \(\beta\) and from both definitions of \(\gamma\) provided, the frequency for each vibration mode can be directly determined. From the angular frequency \(\omega = 2\pi f\), the resonant frequency \(f_{mn}\) can then be determined for a given mode shape as:

\[
f_{mn} = \frac{\pi}{a^2} \sqrt{\frac{D}{I_0}} \left(\beta_{mn}\right)^2
\]  

(3.51)

which is the same as equation (3.9) except it applies to all mode shapes. With the known mode shape constants \(\beta_{mn}\) and the form of the deflection equation, the overall solution to the
The homogeneous vibration equation can be expressed as a series of the characteristic functions:

\[ \Psi_{emm} = \cos (m\theta) \left( J_m \left( \frac{\pi \beta_{mn} r}{a} \right) - \frac{J_m (\pi \beta_{mn})}{I_m (\pi \beta_{mn})} I_m \left( \frac{\pi \beta_{mn} r}{a} \right) \right) \]

with coefficients:

\[ \Lambda_{mn} = \frac{1}{\pi a^2} \int_0^{2\pi} d\theta \int_0^a \Psi_{\sigma mn}^2 r dr \]  

(3.53)

where \( \sigma = o \) or \( e \).

### 3.2.3 PMUT Clamped Plate Resonance Example

In the PMUT design, the electrode, PZT and silicon constitute connected layers of a thin plate. In initial analysis, the silicon plate is considered to be much thicker than the PZT and electrode layers. Preliminary calculations thus consider vibration modes and acoustic wave radiation from a circular, silicon plate source. At the first fundamental frequency mode \((m, n) = (0, 1)\), the characteristic deflection shape is ideal for acoustic wave radiation and sensing. The deflection shape at higher order modes can create destructive interference that reduces acoustic power output. At the first fundamental frequency, the deflection is axisymmetric about the z-axis:

\[ W (r) = A \left[ J_0 (\gamma_{01} r) - \frac{J_0 (\gamma_{01} a)}{I_0 (\gamma_{01} a)} I_0 (\gamma_{01} r) \right] \]

(3.54)

For a single material plate, the first mode frequency \( f_{01} \) is deduced from 3.51 for \( \beta_{01} = 1.015 \):

\[ f_{01} = 0.4671 \frac{h}{a^2} \sqrt{\frac{Y}{\rho (1 - \nu^2)}} \]

(3.55)

Depending on the medical ultrasound application, ultrasonic imaging frequencies typically range between 1 and 15 MHz. Optimal imaging is a balance between high resolution and acoustic beam attenuation [30]. At higher frequencies (10-15 MHz), high resolution imaging of superficial targets is possible with the short signal wavelength, but the high in-
tensity signal is also quickly attenuated in tissue. As a result, lower frequencies (< 5 MHz) are used to image large or deep organs. Lower frequency imaging encompasses a large variety of ultrasound applications, so first mode resonant frequencies between 1 and 5 MHz were used to determine the optimal dimensions for a vibrating silicon plate.

PMUTs are fabricated on silicon on insulator (SOI) wafers. The electrodes and PZT are patterned on the device layer, and the back side is etched up to the buried oxide layer to form the final PMUT device. Thin electrode, PZT and thermal oxide layers lie on top of silicon, which is the thickest layer of the device. After the back side etch process, the remaining silicon is the thickness of the SOI device layer ($O(1\mu m)$). With estimates of the imaging frequency, silicon plate thickness and material properties, the appropriate device plate radius can be estimated. It is important to note the assumptions inherent in the estimation, which have been listed previously but will be enumerated again for clarity:

1. The thickness $h$ of the silicon is much larger than the combined thickness of the electrode, PZT, and thermal oxide layers, and therefore material properties of the thinner layers can thus be neglected.

2. There is no initial applied load on the system ($q = 0$).
3. The rotary inertia $I_2$ is negligibly small.

4. The plate is clamped along its edges.

For a silicon plate excited at the first fundamental frequency mode, the optimal plate radius can be slightly larger than the practical sizes of the CMUT. Pull-in voltage limitations constrain the size of the CMUT; however, PMUT size does not have the same physical restrictions. The final design of the PMUT will have to consider acoustic pressure output, array interfacing requirements and coupling with the piezoelectric thin film as size constraints. For the case of an unloaded, clamped silicon circular plate, the initial fundamental frequency estimation provides a foundation for further study involving these added design considerations. Later sections will consider the influence of additional layers and piezoelectricity on the deflection.
3.3 Ultrasonic Acoustics

3.3.1 Wave Equation [43]

Plate vibration at a high frequency in a fluid media acts an acoustic wave source. Acoustic waves of pressure $p$ propagate three-dimensionally according to the fundamental wave equation:

$$\frac{\partial^2 p}{\partial t^2} = c_m^2 \nabla^2 p$$  \hspace{1cm} (3.56)

where $c_m$ is the speed of sound in the fluid media. Please note that the Laplacian $\nabla^2$ above is for a three-dimensional coordinate system, which varies from the plate specific definition in 3.41. The three-dimensional wave equation is derived from a force balance between acoustic pressure over an infinitesimal area balanced by mass times acceleration. The continuity equation for the infinitesimal volume element allows the acceleration to be written in terms of pressure, which is then substituted into the force balance to yield the one-dimensional wave equation. Through the application of common vector calculus operators, the three-dimensional wave equation is derived as shown above. As with the plate equations of motion, the second order partial differential wave equation can only be explicitly solved in specific circumstances for given boundary conditions.

The example of a simple source is considered as a solvable application of the wave equation. From a source, acoustic waves propagate spherically outward and are only radially dependent. Thus, the wave equation reduces to:

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial p}{\partial r} \right) = \frac{1}{c_m^2} \frac{\partial^2 p}{\partial t^2}$$  \hspace{1cm} (3.57)

The simple source solution is finite everywhere except for $r = 0$, so the general solution takes the form:

$$p = \frac{1}{r} F(r - c_m t) + \frac{1}{r} f(r + c_m t)$$  \hspace{1cm} (3.58)

which is similar to the plane wave equation.

This general form of the solution can be applied to the case of a simple harmonic source.
Assume the source is a simple harmonic function of the form $S_\omega e^{-j\omega t}$ fixed at the origin. At an arbitrary distance $R = (x, y, z)$ from the source, the general solution of the wave equation becomes:

$$p = \frac{jk \rho c}{4\pi R} S_\omega e^{jk(R-ct)} \tag{3.59}$$

where the wave number $k = \omega/c$ and $\rho$ is the density of the fluid media. The pressure from a simple harmonic source can be generalized for any source point $R_0 = (x_0, y_0, z_0)$:

$$p_\omega (R|R_0) e^{j\omega t} = jk \rho m c m S_\omega g_\omega (R|R_0) e^{j\omega t} \tag{3.60}$$

where $g_\omega (R|R_0)$ is the Green's function solution of the equation:

$$\nabla^2 g_\omega (R|R_0) + k^2 g_\omega (R|R_0) = -\delta (R - R_0) \tag{3.61}$$

and $\delta (R - R_0)$ is the three-dimensional delta function.

The Green's function $g_\omega (R|R_0)$ solves the above equation everywhere except for $r = r_0$. From the simple harmonic source example, the homogenous solution of 3.61 clearly takes the form:

$$g_\omega (R|R_0) = \frac{1}{4\pi L} e^{ikL} \tag{3.62}$$

where $L = R - R_0$. The Green's function is significant in that it enables the wave equation solution to be written in a generalized form for various initial and boundary conditions. For a specific case, the solution can be written in terms of green's functions, which are then evaluated with the provided conditions. An alternative form of the Green's function solves the most general case and is composed of the homogeneous $g_\omega (R|R_0)$ and particular $\chi (R)$ solutions:

$$G_\omega (R|R_0) = g_\omega (R|R_0) + \chi (R) \tag{3.63}$$

The particular solution is introduced to represent boundaries at a finite distance from the source or surface where the acoustic waves originate. For a simple source, there are no boundaries present and $\chi = 0$. However, for the case of acoustic wave radiation from a
Figure 3-13: Point source geometry of acoustic radiation from a rigid plane boundary

surface, boundaries are present and the particular solution no longer vanishes.

The simplest model of acoustic radiation and reflection from a surface considers a rigid, plane surface. In this model, the plane surface is assumed to be infinitely long and the material acoustic impedance of the plane is neglected. Along the plane of the surface, the fluid particles are at zero velocity, and the normal component (along \( \hat{n} \)) of the pressure gradient is also zero. With the definition of a simple source provided above, it is possible to model the acoustic pressure reflection and radiation from this rigid surface via symmetry and a combination of simple sources. Assume the rigid surface lies along the plane \( z = 0 \), and a point source exists above the plane at \( R_0 = (x_0, y_0, z_0) \). For acoustic pressure waves to satisfy the conditions of zero velocity and normal pressure gradient along the surface, there must be an equal point source \( R'_0 = (x_0, y_0, -z_0) \) symmetric to \( R_0 \) about the \( z = 0 \) plane. The resulting acoustic pressure equation includes this boundary, and thus, the Green’s function \( G_\omega (R|R_0) \) includes the particular solution \( g_\omega (R|R'_0) \) for the reflected source. For the rigid boundary, the Green’s function is explicitly defined as:

\[
G_\omega (R|R_0) = \frac{1}{4\pi L} e^{ikL} + \frac{1}{4\pi L'} e^{ikL'}
\]

(3.64)
where

\[
L = [(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2] \\
L' = [(x - x_0)^2 + (y - y_0)^2 + (z + z_0)^2]
\]

The geometry and relations in the above equations are schematically shown in Figure 3-13.

3.3.2 Acoustic Pressure Output of Vibrating Circular Plate [43]

In transmit mode, the circular plate PMUT will act as a boundary surface for acoustic wave radiation. For a simple harmonic motion of the surface with the velocity profile \(u_z = u_\omega e^{j\omega t}\), the normal pressure gradient is no longer zero and the boundary condition along the surface is redefined as:

\[
\frac{\partial p}{\partial z} = jk\rho c u_\omega (x, y) e^{j\omega t}
\]

(3.66)

The resulting acoustic pressure at an arbitrary point \((x, y, z)\) away from the surface is then dependent on the Green's function and velocity profile along the surface \((x_0, y_0, 0)\):

\[
p_\omega (x, y, z) = jk\rho c m \int u_\omega (x_0, y_0) G_\omega (x, y, z|x_0, y_0, 0) \, dx_0 dy_0
\]

(3.67)

For waves to propagate, the PMUT surface deflects with a high frequency harmonic excitation according to the deflection equation described in Section 3.2.2:

\[
w (r, t) = A \left[ J_0 (\gamma r) - \frac{J_0 (\gamma a)}{I_0 (\gamma a)} I_0 (\gamma r) \right] e^{j\omega t}
\]

(3.68)

Since the first vibration mode deflection is axisymmetric about the z-axis, the deflection can be written in terms of the shape function as \(w (r, t) = W (r) e^{j\omega t}\). The simple harmonic velocity profile along the surface should be equivalent to the derivative of the axial deflection with respect to time.

\[
u_z = \frac{dw}{dt} = j\omega W (r) e^{j\omega t}
\]

(3.69)
Thus, the amplitude of the velocity profile $u_\omega$ becomes:

$$u_\omega = j\omega W(r) = j\omega A \left[ J_0(\gamma r) - \frac{J_0(\gamma a)}{I_0(\gamma a)} I_0(\gamma r) \right] \tag{3.70}$$

and can be substituted into the acoustic pressure equation 3.67. The above equation is in polar coordinates and can be translated into cartesian coordinates by setting $r = \sqrt{x_0^2 + y_0^2}$. This final substitution yields an explicit integral equation for the acoustic pressure at any specified distance from the vibrating plate surface.

### 3.3.3 Numerical Solution of Acoustic Pressure

With the derived acoustic pressure equation, the influence of various plate dimensions and boundary conditions on the acoustic pressure can be investigated. The equation 3.67 was numerically integrated for specified dimensions to compare acoustic pressure output for specific plate configurations. A simple analysis where the plate thickness was fixed by the frequency according to equation 3.55 with a changing radius was performed to determine the relationship between plate radius and acoustic pressure. When the plate radius is doubled and the operating frequency is held constant, the acoustic pressure increases by a factor of eight. A similar relation was observed when the plate radius was tripled, and the acoustic pressure increased by a factor of 27. For various other multiplicative factors, this cubic scaling also held true. Thus, when the radius is multiplied by a factor $q$, the pressure increases according to the scaling:

$$p(qa) = q^3 p(a) \tag{3.71}$$

The acoustic pressure equation involves the surface integral of the velocity profile, which yields a volumetric displacement rate. Qualitatively, this means that the acoustic pressure should scale with the displaced volume between the deflection profile and the $z = 0$ plane, and explains the cubic term in the pressure scaling equation.
The pressure scaling with volume can be more clearly visualized by approximating the deflection profile as an circular arc in two dimensions and a spherical cap in three dimensions. The equation for an arc can be fit to three points that lie along a curve. For the first mode deflection, it is convenient to define these points at the clamped boundary of the plate ±a and at the center where the deflection is maximized with amplitude A. Evaluation of the matrix determinant:

\[
\begin{vmatrix}
  r^2 + z^2 & r & z & 1 \\
  a^2 & -a & 0 & 1 \\
  A^2 & 0 & A & 1 \\
  a^2 & -a & 0 & 1 \\
\end{vmatrix} = 0
\]  

(3.72)

yields the necessary parameters to derive the arc equation:

\[(z - z_0)^2 + r^2 = R^2\]  

(3.73)

where

\[
R^2 = a^2 + \frac{a(A^2 - a^2)^2}{4A^2a^2}
\]

\[z_0 = \frac{a(A^2 - a^2)}{2Aa}\]  

(3.74)

The acoustic pressure was numerically determined with the circular arc equation using 3.67, and the spherical cap volumetric scaling was compared to the numerically determined acoustic pressure. Since the first mode vibration deflection is axisymmetric, the two-dimensional displacement between the arc and \(z = 0\) plane can be extended into three dimensions as the volume of a spherical cap. The volume \(V\) of a spherical cap is given as:

\[
V = \frac{1}{6} \pi A (3a^2 + A^2)
\]  

(3.75)

For the PMUT circular plate design, the deflection is much smaller than the plate radius, and the second order maximum deflection term \(A^2\) can be neglected in the overall volume.
The numerically determined acoustic pressure for the circular arc deflection scales according to 3.71. Since the volume and pressure show the same radial dependence, the acoustic pressure must scale with the displaced volume of the first mode vibration.

### 3.3.4 Acoustic Pressure Comparison between Potential PMUT Designs

With the numerically determined acoustic pressure, the acoustic pressure can be determined for an arbitrarily defined velocity profile in x- and y- coordinates. For the highest acoustic pressure output, the optimal deflection profile can be determined by comparing various plate configurations. Acoustic pressure output is commonly compared to that of an equivalent rigid piston [47] [37]. For the circular plate first mode vibration, equivalent rigid pistons are defined at the maximum and average displacements. The average displacement...
of the deflection $\langle W \rangle$ is determined by the integral:

$$\langle W \rangle = \frac{\int 2\pi r W (r) \, dr}{\pi a^2}$$  \hspace{1cm} (3.77)

Both the average and maximum displacement can be substituted into the acoustic pressure equation using the same approach as with the deflection for the first vibration mode in Section 3.3.2. It is found that the equivalent piston model acoustic pressures at the average and maximum displacement were larger than the first mode vibration deflection. Since the acoustic pressure is larger in the equivalent piston model, a PMUT design based on a rigid piston should produce a larger acoustic pressure output with less overall deflection. The reduced deflection would be ideal for powering the PMUT in transmit mode. A lower voltage could likely be applied across the PZT to produce the appropriate acoustic pressure output.

The proposed numerical model is a useful design tool for the PMUT. For different plate geometries, the acoustic pressure can be determined and compared to evaluate potential designs.

Figure 3-15: Acoustic Pressure scaling for circular plate with radius $a = 50\mu m$ resonating at 3MHz with plate thickness set by 3.51
3.4 Green’s Function Solution

3.4.1 Derivation of Governing Plate Vibration Equation

Since the lateral dimensions of the piezoelectric disk and silicon support material are much larger than the thickness, classic plate theory is appropriate for analysis. Recalling the displacement field in equation 3.20, the bending strain-displacement relations are determined by plate curvatures \( (\varepsilon_{rr}, \varepsilon_{\theta\theta}, \gamma_{r\theta}) \) during deformation [48].

\[
\varepsilon_{rr} = 2\varepsilon_{rr}^{(1)} = \frac{\partial u_r}{\partial r} = -z \frac{\partial^2 w}{\partial r^2}
\]

\[
\varepsilon_{\theta\theta} = 2\varepsilon_{\theta\theta}^{(1)} = \frac{u_r}{r} + \frac{1}{r} \frac{\partial u_\theta}{\partial \theta} = -\frac{z}{r} \left( \frac{\partial w}{\partial r} + \frac{1}{r} \frac{\partial^2 w}{\partial \theta^2} \right)
\]

\[
\gamma_{r\theta} = z\gamma_{r\theta}^{(1)} = \frac{1}{r} \frac{\partial u_r}{\partial \theta} + \frac{\partial u_\theta}{\partial r} = \frac{2z}{r} \left( \frac{\partial^2 w}{\partial r \partial \theta} - \frac{1}{r} \frac{\partial w}{\partial \theta} \right)
\]

Contributions from bending dominate the strain field and the tension (membrane strain) experienced along the surface mid-plane can be neglected. From the displacement field, it is clear that the transverse strain components \( (\varepsilon_{zz}, \varepsilon_{rz}, \varepsilon_{\theta z}) \) are identically zero, and transverse normal stresses can also be neglected.

The aforementioned PMUT design uses lead zirconate titanate (PZT) as the piezoelectric material, and silicon as a support layer. Initially, the e-form of the linear, piezoelectric constitutive relations are chosen for further development of the stress \( (\sigma) \) - strain \( (\varepsilon) \) relations in the PZT thin film:

\[
\sigma_{ij} = c_{ijkl}^{E} \varepsilon_{kl} - \varepsilon_{ki} E_{k}
\]

\[
D_i = e_{ikl} \sigma_{kl} - \varepsilon_{ik} E_{k}
\]

where \( D_i \) is the electric displacement, \( e_{ki} \) is a piezoelectric constant, \( E_k \) is the electric field, and \( c_{ijkl}^{E} \) is the compliance measured at constant electric field, and \( \varepsilon_{ik} \) is the dielectric constant measured at constant strain [57], [33]. This high order tensor form denoted with integer subscripts \( i, j, k, l \) is the most general form of the piezoelectric relations. Based on the physical problem, major simplifications are made to the constitutive relations. Since
the PZT is actuated through an applied voltage between a top and bottom electrode, the $e_{31,f}$ mode is excited and all other directional piezoelectric contributions are neglected. For simplicity, the effective piezoelectric constant $e_{31,f}$ is written in term of the bulk properties: the piezoelectric constant $d_{31}$ and stiffnesses $s_{11}^E, s_{12}^E$.

$$e_{31,f} = \frac{d_{31}}{s_{11}^E + s_{12}^E} \quad (3.83)$$

Since PZT and silicon are isotropic, the compliance and stiffness terms can be written in terms of axisymmetric material properties (i.e. the Young’s modulus $Y = (Y_{PZT}, Y_{Si})$, and poisson’s ratio $\nu = (\nu_{PZT}, \nu_{Si}$)). Combining Hooke’s Law and the piezoelectric constitutive relations, the resulting plane-stress reduced constitutive relations for PZT become:

$$\sigma_{rr} = \frac{Y_{PZT}}{1 - \nu^2} [\varepsilon_{rr} + \nu \varepsilon_{\theta\theta}] + \frac{Y_{PZT} d_{31} E_z}{1 - \nu_{PZT}} \quad (3.84)$$

$$\sigma_{\theta\theta} = \frac{Y_{PZT}}{1 - \nu_{PZT}^2} [\nu \varepsilon_{rr} + \varepsilon_{\theta\theta}] + \frac{Y_{PZT} d_{31} E_z}{1 - \nu_{PZT}} \quad (3.85)$$

$$\sigma_{r\theta} = \frac{Y_{PZT}}{2(1 + \nu_{PZT})} \gamma_{r\theta} \quad (3.86)$$

Similarly, the constitutive relations for the silicon layer are obtained:

$$\sigma_{rr} = \frac{Y_{Si}}{1 - \nu_{Si}^2} [\varepsilon_{rr} + \nu \varepsilon_{\theta\theta}] \quad (3.87)$$

$$\sigma_{\theta\theta} = \frac{Y_{Si}}{1 - \nu_{Si}^2} [\nu \varepsilon_{rr} + \varepsilon_{\theta\theta}] \quad (3.88)$$

$$\sigma_{r\theta} = \frac{Y_{Si}}{2(1 + \nu_{Si})} \gamma_{r\theta} \quad (3.89)$$

For a circular plate of PZT operating in $d_{31}$ mode, the electrical charge flows between the top and bottom electrodes and along the transverse normals perpendicular to the surface mid-plane. An applied voltage creates a strain mismatch between the top and bottom of the piezoelectric material that causes the plane to deflect in transmit mode. In receiving mode, the incoming acoustic pressure induces an analogous strain mismatch generating an
electrical signal. Again using classic plate theory and the piezoelectric constitutive relations, the transverse electrical displacements are negligible and the reduced electrical displacement $D_z$ of the PZT layer becomes:

$$D_z = \frac{Y_{PZT}d_{31}}{1 - \nu_{PZT}} (\varepsilon_{rr} + \varepsilon_{r\theta}) + \varepsilon_z E_z$$ (3.90)

With the enumerated constitutive relations, the equations of motion for a vibrating plate can be determined. In classic plate analysis literature [48], the principle of virtual displacements or Hamilton’s principle is commonly used to derive the general plate equations of motion. A common form of the equation of motion can be obtained as a function of bending moments, displacements and normal forces. For infinitesimal strains, the normal force becomes negligibly small and the motion equation 3.29 simplifies to:

$$-\frac{1}{r} \left[ \frac{\partial^2}{\partial r^2} (r M_{rr}) - \frac{\partial M_{r\theta}}{\partial r} + \frac{1}{r} \frac{\partial^2 M_{r\theta}}{\partial r \partial \theta} + \frac{2}{r} \frac{M_{r\theta}}{\partial \theta} \right] = -I_0 \frac{\partial^2 w}{\partial t^2}$$ (3.91)

In the above equation, applied forces including the foundation reaction is neglected and higher order moment of inertia terms are considered negligibly small. For the bimorph and the more realistic unimorph model, the equations of the motion are derived separately but take analogous forms that will later be grouped into a single general equation.
3.4.2 Bimorph Vibration Equation

Assuming applied voltage $V$ is linearly proportional to the electric field, Maxwell's law for the piezoelectric plate reduces to:

$$E_z = V/h_{PZT} \quad (3.92)$$

where $h_{PZT}$ is the thickness of the PZT layer. The applied electric field is constant with the PZT plate thickness.

Substituting equation (3.92) into the equations (3.84) and (3.85) yields a $z$-dependent form of the stress equations. With the simplified stress equations, the stress-induced moments experienced along the center ($z = 0$) of the piezoelectric plate are expressed as:

$$M_{rr} = \int_{-h_{PZT}/2}^{h_{PZT}/2} \sigma_{rr} zdz = -D \left[ \frac{\partial^2 w}{\partial r^2} + \frac{\nu_{PZT}}{r} \left( \frac{\partial w}{\partial r} + \frac{1}{r} \frac{\partial^2 w}{\partial \theta^2} \right) \right] + \frac{Y_{PZT}d_{31} h_{PZT} V}{2(1-\nu_{PZT})} \quad (3.93)$$

$$M_{\theta\theta} = \int_{-h_{PZT}/2}^{h_{PZT}/2} \sigma_{\theta\theta} zdz = -D \left[ \nu_{PZT} \frac{\partial^2 w}{\partial r^2} + \frac{1}{r} \left( \frac{\partial w}{\partial r} + \frac{1}{r} \frac{\partial^2 w}{\partial \theta^2} \right) \right] + \frac{Y_{PZT}d_{31} h_{PZT} V}{2(1-\nu_{PZT})} \quad (3.94)$$

$$M_{r\theta} = \int_{-h_{PZT}/2}^{h_{PZT}/2} \sigma_{r\theta} zdz = -(1-\nu_{PZT}) D \left( \frac{\partial^2 w}{\partial r \partial \theta} - \frac{1}{r} \frac{\partial w}{\partial \theta} \right) \quad (3.95)$$

where $D$ is the isotropic bending stiffness:

$$D = \frac{h_{PZT}^3 Y_{PZT}}{12(1-\nu^2)} \quad (3.96)$$

The final equation of motion is deduced through substitution of the moment equations (3.93), (3.94), and (3.95) into equation (3.91). With further simplification, the final form of the plate vibration equation becomes:

$$D \nabla^2 \nabla^2 w + I_0 \frac{\partial^2 w}{\partial t^2} = \nabla^2 M_p \quad (3.97)$$

which is similar to the general plate vibration equation (3.40). For convenience, the piezo-
electric moment term \( M_p \) is introduced:

\[
M_p = \frac{Y_{PZT} d_31 h_{PZT} V}{2(1 - \nu_{PZT})}
\]  

(3.98)

which can be determined directly from equations (3.93) and (3.94). Since the thickness of the electrodes is assumed to be much smaller than the thickness of the PZT, the electrode material properties can be neglected. Therefore the bending stiffness and moment of inertia are only dependent on the material properties and geometry of the PZT layer. By integration, the moment of inertia for the bimorph is:

\[
I_0 = \int_{-h_{PZT}/2}^{h_{PZT}/2} \rho_{PZT} dz = \rho_{PZT} h_{PZT}
\]  

(3.99)

where \( \rho_{PZT} \) is the density of PZT. With the provided constants, equation (3.97) becomes the fully defined plate vibration equation for a bimorph plate. This final equation is a piezoelectric modified form of the plate vibration equations derived in acoustics and classic plate theory literature [48], [43]. It will be shown that the bimorph equation is a special case of a more general vibration equation of motion that will now be determined for the unimorph geometry.

3.4.3 Unimorph Vibration Equation

Since the unimorph model lacks symmetry about a central axis, the derivation of the equation of motion is slightly more involved. For equation (3.91) to apply, the bending moments must be determined along the neutral axis, pictured at a height \( z_N \) from the bottom of the silicon support in Figure 3-16b. For the circular plate geometry, the neutral axis lies on the surface where the net longitudinal stresses and strains are zero and is determined from the general equation (3.81):

\[
z_N = \frac{\frac{Y_{Si} h_{Si}^2}{1-\nu_{Si}^2} + \frac{Y_{PZT} [(h_{Si} + h_{PZT})^2 - h_{Si}^2]}{1-\nu_{PZT}^2}}{\frac{Y_{Si} h_{Si}}{1-\nu_{Si}^2} + \frac{Y_{PZT} h_{PZT}}{1-\nu_{PZT}^2}}
\]  

(3.100)
where $h_{Si}$ is the thickness of the silicon layer. Along the neutral axis, the bending moments can be determined in an analogous manner to those in equations (3.93), (3.94), and (3.95):

\[
M_{rr} = \int_{-h_N}^{h_{Si} - h_N} \sigma_{rr, Si} z dz + \int_{h_{Si} - h_N}^{(h_{Si} + h_{PZT}) - h_N} \sigma_{rr, PZT} z dz = D' \varepsilon_{rr}^{(1)} + D'' \varepsilon_{\theta \theta}^{(1)} + M'_p \tag{3.101}
\]

\[
M_{\theta \theta} = \int_{-h_N}^{h_{Si} - h_N} \sigma_{\theta \theta, Si} z dz + \int_{h_{Si} - h_N}^{(h_{Si} + h_{PZT}) - h_N} \sigma_{\theta \theta, PZT} z dz = D'' \varepsilon_{rr}^{(1)} + D' \varepsilon_{\theta \theta}^{(1)} + M'_p \tag{3.102}
\]

\[
M_{r \theta} = \int_{-h_N}^{h_{Si} - h_N} \sigma_{r \theta, Si} z dz + \int_{h_{Si} - h_N}^{(h_{Si} + h_{PZT}) - h_N} \sigma_{r \theta, PZT} z dz = \frac{D' - D''}{2} \gamma_{r \theta}^{(1)} \tag{3.103}
\]

Based on the new limits of integration, the thickness dependence within the applied piezoelectric moment term is altered and the unimorph piezoelectric moment $M'_p$ becomes:

\[
M'_p = \frac{Y_{PZT} V d_{31}}{1 - \nu_{PZT}} \left( h_{Si} + h_{PZT}/2 - z_N \right) \tag{3.104}
\]

Since the strain fields experienced by the Silicon and PZT plates are similar, the strain terms can be grouped into bending stiffness constants $D'$ and $D''$ related to the bending stiffness in the PZT $D_{PZT}$ and Silicon $D_{Si}$ layers.

\[
D_{PZT} = \frac{Y_{PZT}}{3(1 - \nu_{PZT}^2)} \left[ ((h_{Si} + h_{PZT}) - h_N)^3 - (h_{Si} - h_N)^3 \right] \tag{3.105}
\]

\[
D_{Si} = \frac{Y_{Si}}{3(1 - \nu_{Si}^2)} \left[ (h_{Si} - h_N)^3 + h_N^3 \right] \tag{3.106}
\]

\[
D' = D_{Si} + D_{PZT} \tag{3.107}
\]

\[
D'' = \nu_{Si} D_{Si} + \nu_{PZT} D_{PZT} \tag{3.108}
\]

Equations (3.101), (3.102), and (3.103) can then be substituted into the bending moment equation of motion (3.91). With further manipulation, the $D''$ terms cancel out and the
resulting vibration equation for the unimorph can be written as:

$$D' \nabla^2 \nabla^2 w + I_0 \frac{\partial^2 w}{\partial t^2} = \nabla^2 M_p'$$

(3.109)

The moment of inertia is also adapted to consider the bending contribution and density $\rho_{Si}$ of the silicon layer:

$$I_0' = \int_{-h_N}^{h_{Si}-h_N} \rho_{Si} dz + \int_{h_{Si}-h_N}^{(h_{Si}+h_{PZT})-h_N} \rho_{PZT} dz = \rho_{Si} h_{Si} + \rho_{PZT} h_{PZT}$$

(3.110)

Thus, the plate vibration equations (3.97) and (3.109) are of the same form. The above derivation can be extended to apply to a stack with $n$ layers. In this case, the moment of inertia becomes:

$$I_0' = \sum_{i=1}^{n} \rho_i h_i$$

(3.111)

and the flexural rigidity is:

$$D' = \sum_{i=1}^{n} \frac{Y_i}{3(1-\nu^2)} [(z_i - z_N)^3 - (z_{i-1} - z_N)^3]$$

(3.112)

which agrees with the equation for flexural rigidity derived in [21]. Recall that $z_i$ and $h_i$ are qualitatively defined in Figure 3-3. For simplicity, the unimorph equation of motion (3.109) will be considered the general form of the plate vibration equation, and it should be noted that for the special case of a bimorph, $D' = D$, $M_p' = M_p$ and $I_0' = I_0$. In subsequent sections, the simplified model considering only the thickness of the PZT and silicon will be used; however, the equation of motion can be applied to additional material layers with equations (3.81),(3.111), and (3.112).

### 3.4.4 Arbitrary Electrode Configuration

With the general form of the vibration equation, the equation can be extended to the case of an arbitrary electrode configuration. For completeness, the unimorph case will be analyzed in the derivation; however, the final solution can be easily solved for the bimorph
Figure 3-17: Unimorph PMUT with $N$ circular/ring electrodes in $d_{31}$ mode

through substitution of the appropriate bending stiffness and moment of inertia terms.

Consider a circular disk of PZT sandwiched between $N$ ring electrodes, each excited by an identical AC voltage $V e^{j \omega t}$ of angular frequency $\omega$. The PZT/electrode composite is supported by a circular plate of silicon as shown in Figure 3-17. Under simple harmonic excitation, the deflection $w$ is assumed to take the form $w = W(r, \theta) e^{j \omega t}$. The shape function $W(r, \theta)$ describes the contour of the plate during vibration, and can be inserted into the differential equation (3.109):

$$\left(\nabla^4 - \gamma^4\right) W = f(r) \quad (3.113)$$

where $\gamma$ is defined in equation 3.44.

For convenience in the later Green’s function derivation, the right hand side of equation (3.113) is grouped into the function:

$$f(r) = \frac{1}{D'} \nabla^2 M'_{p'}(r) \quad (3.114)$$

In the arbitrary ring electrode configuration, the PZT layer will be excited in areas beneath the electrode, so the equation of motion is only valid in the regions covered by an
electrode; otherwise, the applied piezoelectric moment is zero. For the above equation to be universally valid for the arbitrary electrode case, the piezoelectric applied moment is defined by a series of step functions:

\[
M'_p (r) = M'_p \sum_{k=1}^{N} (H(r - a'_k) - H(r - a''_k))
\] (3.115)

which satisfies the condition that:

\[
M'_p (r) = \begin{cases} 
M'_p & \text{under electrode} \\
0 & \text{no electrode} 
\end{cases}
\] (3.116)

This piezoelectric moment definition is particularly useful given the differentiation properties of the step and delta functions. With the boundary value problem defined in equation (3.113), it is most common to apply boundary conditions and solve for deflection piecewise. However, this limits the applicability of the solution to plates with only a small number of electrodes; otherwise, the solution process must be iterated to produce a solution that can become impossibly complex. Given these limitations, an alternative Green’s function solution technique is used to define a more tractable form of the deflection equation with wider applicability.

### 3.4.5 Definition of Green’s Function

In section 3.2.2, the case with no applied load was solved. To broaden the applicability of this homogeneous solution, we will now consider a force applied at point \((r_0, \theta_0)\) driving the steady-state plate motion. The influence of electrodes will initially be neglected. It is assumed that the resulting heterogeneous equation can be solved with Green’s function \(G(r\theta | r_0\theta_0)\):

\[
(\nabla^4 - \gamma^4) G = \frac{1}{r} \delta(r - r_0) \delta(\theta - \theta_0)
\] (3.117)
In mathematics, the Green’s function method is convenient for solving specific boundary value problems. The technique employs an auxiliary function known as Green’s function that is defined independently from the forcing function $f(r)$. If the form of Green’s function is known, the boundary value problem can be simplified into an integral form that relates the Green’s function and forcing function to the final solution. Since our arbitrary electrode case involves a boundary value problem with a forcing function $f(r) = \nabla^2 M_p(r)$, the Green’s function method is an convenient choice to solve for the deflection [49].

Because the homogeneous solution to the deflection equation is a series of characteristic functions, it is assumed that the Green’s function solution to equation (3.117) is also of the same form:

$$G(r, \theta|r_0, \theta_0) = \sum A_{\sigma mn} \Psi_{\sigma mn}$$  \hspace{1cm} (3.118)

Again from the homogeneous solution, we know that the constants denoted by $A_{\sigma mn}$ are related to the characteristic function coefficients, which will later be used for simplification. By substituting equation (3.118) into (3.117) and multiplying through by $\Psi_{\sigma mn} r dr d\theta$, equation (3.117) becomes:

$$\int_0^a \int_0^{2\pi} \Psi_{\sigma mn} (r, \theta) \left[ \sum_{s kl} A_{skl} (\gamma_k^s - \gamma_l^k) \Psi_{skl} (r, \theta) \right] r dr d\theta = \int_0^a \int_0^{2\pi} \delta(r - r_0) \delta(\theta - \theta_0) \Psi_{\sigma mn} (r, \theta) dr d\theta$$  \hspace{1cm} (3.119)

which is then integrated over the area of the plate. The integers $s$, $k$, and $l$ are dummy variables in the summation. In the resulting integral, the product of the characteristic functions $\Psi_{\sigma mn} \Psi_{skl}$ is only non-zero when $s = \sigma$, $k = m$ and $l = n$. Given the properties of delta functions and the definition of the characteristic function coefficients provided in equation (3.53), the constants $A_{\sigma mn}$ can be determined. By substituting $A_{\sigma mn}$ back into equation (3.118), the Green’s function becomes:

$$G(r, \theta|r_0, \theta_0) = \frac{1}{\pi a^2} \sum_{\sigma, m, n} \frac{\Psi_{\sigma mn} (r, \theta) \Psi_{\sigma mn} (r_0, \theta_0)}{\Lambda_{mn} (\gamma_m^\sigma - \gamma_n^\sigma)}$$  \hspace{1cm} (3.120)

and is now in terms of known values and functions.
3.4.6 Arbitrary Electrode Green’s Function Solution

For simplicity, only axisymmetric deflection will be analyzed and thus, the Laplacian operator becomes:

\[
\nabla^2 \equiv \Delta = \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial}{\partial r^2} \right) \tag{3.121}
\]

Since the pMUT is operated near the first resonant frequency (mode (0, 1)), the deflection mode of interest is axisymmetric. Near the first resonance mode, the magnitude of the (0, 1) deflection will be much larger than any other modes, so it is appropriate to model the problem as axisymmetric for the operating frequencies of interest. It is likely that neglecting the non-axisymmetric modes will result in a small but negligible error in the deflection profile calculation. For comprehensiveness, all axisymmetric modes are considered in the following analysis.

The deflection equations can now be solved by using the properties of Green’s function. First, equation (3.113) is multiplied by \( r G(r| r_0) \) and equation (3.117) is multiplied by \( r W(r) \). The second term is subtracted from the first and the result is integrated over the plate area:

\[
\int_{0}^{a-\epsilon} \left[ (\Delta^2 G(r| r_0) - \gamma^4 G(r| r_0)) W(r) - (\Delta^2 W(r) - \gamma^4 W(r)) G(r| r_0) \right] r dr \\
= \int_{0}^{a-\epsilon} [W(r) \delta(r - r_0) - rf(r) G(r| r_0)] dr 
\tag{3.122}
\]

The infinitesimal value \( \epsilon \) is added to the upper limit of integration so the integral properties of the delta function can be properly realized but does not limit the application of the clamped plate boundary conditions. On the left hand side of equation (3.122), the \( \gamma^4 \) terms cancel.
out and integration by parts eliminates the terms involving the bilaplacian $\Delta^2$:

$$\int_0^{a-\epsilon} \left[ (\Delta^2 G(r|r_0) - \gamma^4 G(r|r_0)) W(r) - (\Delta^2 W(r) - \gamma^4 W(r)) G(r|r_0) \right] r dr$$

$$= \left[ W(r) \left( r \frac{\partial}{\partial r} \Delta G(r|r_0) \right) \right]_0^{a-\epsilon} - \left[ \left( r \frac{\partial W(r)}{\partial r} \right) \Delta G(r|r_0) \right]_0^{a-\epsilon}$$

$$- \left[ \Delta W(r) \left( r \frac{\partial G(r|r_0)}{\partial r} \right) \right]_0^{a-\epsilon} - \left[ G(r|r_0) \left( r \frac{\partial}{\partial r} \Delta W(r) \right) \right]_0^{a-\epsilon}$$

(3.123)

In axisymmetric plate vibration, maximum deflection occurs at the center of the plate; therefore, it is assumed that $\frac{\partial W}{\partial r} |_{r=0} = 0$. With this assumption and the clamped plate boundary conditions, each term on the right hand side is identically zero for the integration limits. With further use of the delta function properties of integration, equation (3.122) becomes:

$$W(r_0) = -\int_0^{a-\epsilon} r f(r) G(r|\rho_0) dr$$

(3.124)

The right hand side is analytically determined with the definition of $f(r)$ in equation (3.114). Again, the differential properties of the delta and step functions prove useful.

$$r f(r) = \frac{M''}{D'} \sum_{k=1}^{N} \left[ a''_k \frac{\partial}{\partial r} \delta (r - a''_k) - a'_k \frac{\partial}{\partial r} \delta (r - a'_k) \right]$$

(3.125)

The integral in equation (3.124) thus becomes:

$$W(r_0) = \frac{M''}{D'} \int_0^{a-\epsilon} \sum_{k=1}^{N} \left[ a''_k \frac{\partial}{\partial r} \delta (r - a''_k) G(r|r_0) - a'_k \frac{\partial}{\partial r} \delta (r - a'_k) G(r|r_0) \right] dr$$

(3.126)

which reduces to:

$$W(r_0) = \frac{M''}{D'} \sum_{k=1}^{N} \left[ a''_k \left| \frac{\partial G(r|r_0)}{\partial r} \right|_{r=a''_k} - a'_k \left| \frac{\partial G(r|r_0)}{\partial r} \right|_{r=a'_k} \right]$$

(3.127)

Since the previous definition of the Green's function included all modes, the definition must be altered for the axisymmetric case to solve equation (3.127). The theta-dependent
term in the governing differential equation drops out reducing equation (3.117) to:

\[(\nabla^4 - \gamma^4) G(r|r_0) = \frac{1}{r} \delta(r - r_0)\]  

(3.128)

Axisymmetric modes occur when \(m = 0\) modifying the Green's function definition in equation (3.120) to become:

\[G(r|r_0) = \frac{1}{\pi a^2} \sum_n \frac{\Psi_{e0n}(r) \Psi_{e0n}(r_0)}{\Lambda_{0n}(\gamma_{0n}^4 - \gamma^4)}\]

(3.129)

The coefficient of the characteristic function is also altered in the axisymmetric case:

\[\Lambda_{0n} = \frac{1}{\pi a^2} \int_0^a \Psi_{e0n}^2 r dr\]

(3.130)

When differentiating the Green's function with respect to \(r\), all terms are constant except for a characteristic function. Given the Bessel function properties and the definition of the characteristic function in equation (3.52), the final form of the Green's function derivative is derived:

\[\frac{\partial G(r|r_0)}{\partial r} = -\frac{1}{a^3} \sum_n \frac{\Psi_{e0n}(r_0) \beta_{0n}}{\Lambda_{0n}(\gamma_{0n}^4 - \gamma^4)} \left[ J_1 \left( \frac{\pi \beta_{0n} r}{a} \right) + \frac{J_0(\pi \beta_{0n})I_1(\pi \beta_{0n} r/a)}{I_0(\pi \beta_{0n})I_1(\pi \beta_{0n} r/a)} \right]\]

(3.131)

Substitution into equation (3.127) yields the final form of the deflection shape profile.

\[W(r_0) = \frac{M_p}{a^3 D'} \sum_n \sum_{k=1}^N \beta_{0n} a_k \left[ J_1 \left( \frac{\pi \beta_{0n} a_k'}{a} \right) + \frac{J_0(\pi \beta_{0n})I_1(\pi \beta_{0n} a_k' a)}{I_0(\pi \beta_{0n})I_1(\pi \beta_{0n} a_k' a)} \right] - a_k'' \left[ J_1 \left( \frac{\pi \beta_{0n} a_k''}{a} \right) + \frac{J_0(\pi \beta_{0n})I_1(\pi \beta_{0n} a_k'' a)}{I_0(\pi \beta_{0n})I_1(\pi \beta_{0n} a_k'' a)} \right]\]

(3.132)

This solution final form is valid for an arbitrary number of circular, ring electrodes. Although the final equation is rather long, all terms except for \(\Lambda_{0n}\) are explicitly defined and not in integral form. A numerical integration can be used to determine the normal-
izing coefficient $\Lambda_{bn}$, but is simple compared to the numerical integration required in the approaches taken by [44], [53]. $\Lambda_{bn}$ is also independent of all design parameters except for plate radius, so for most design optimization (except for plate radius determination), the numerical integration only needs to be preformed once and can then be used for multiple design iterations.

The Green's function derivation is carried out specifically for the plate bending case, but a similar approach can be a useful tool to analyze other scenarios - particularly to determine the effects of residual stress. By modifying the initial differential equation 3.109, the effect of residual stress can be added through a $\nabla^2 T^\sigma$ term on the right hand side where $T^\sigma$ is defined in equation 3.3. Again, this would be for complete top and bottom electrode coverage, for incomplete electrode coverage and additional applied moment term, effecting the coefficient of $\nabla^4$ in equation 3.109 would have to be added as well. In this case, the Green's function must be redefined following the same approach outlined above. For brevity, this derivation will not be included. However, it is important to note that the Green's function approach can be applied to a variety of deflection derivations including changes in boundary conditions and governing differential equation.
Chapter 4

Results

4.1 Model Validation

For the derived models to be useful, it is necessary to compare the results with those obtained experimentally or with previously published solutions.

4.1.1 Acoustic Pressure

Validation is simpler for the acoustic pressure model. The plate vibration can be modelled as an equivalent rigid piston moving at the average velocity of the plate vibrations. This approach is not uncommon when analyzing similar plate or membrane vibration based systems and it is common to add a correction term, usually an equivalent area, to make up for discrepancies in the calculated pressure output [17]. For the derivation of a CMUT equivalent circuit model, an equivalent piston area with the average displacement has been defined with respect to the actual piston area and displacement profile [37], so the equivalent area approach has already been undertaken to define pressure for a MUT design. Since the numerically calculated and equivalent piston pressures are similar, a logical next step would be to define an equivalent area, which for the case of the PMUT, can be approximated as the actual plate area. From comparison between the pressure model and an equivalent piston, the modelling results make sense in the context of the problem and validate the solution.
4.1.2 Deflection Model Verification

The deflection model determination is decidedly more rigorous and given the novel approach, it can be difficult to check. Fortunately, existing energy-based solution techniques have been derived [53], [44] and can be used for comparison. Since the energy-based approaches are most easily solved for one electrode, the case of a single electrode with 60% coverage, previously mentioned as the coverage for optimal displacement, is used to calculate deflection in both models. The energy-based derivation is taken exactly from [21] and [44]; and numerical integrations were performed in Matlab. Multiple geometry cases were selected, and a representative unimorph test case based on the realizable geometries is presented in Figure 4-2. The maximum deflection in the energy-based and Green's function derivation for 1 and 15 modes axisymmetric modes were compared showing an error of 0.1% between the test cases. Error from other test cases was similarly small. With such a small error, the difference between the models can be considered negligible. The Green’s function derivation therefore accurately calculates the deflection profile, but requires substantially less computation and can easily be modified to represent alternative design geometries.

4.2 Electrode Configuration Optimization

Preliminary optimization of electrode configurations was carried out in Matlab using the non-linear optimization method of sequential quadratic programming (SQP). The built-
in optimization function 'fmincon' was used to find the minima of the objective function iteratively with the SQP approach. Initial values and constraints were specified separately from the objective function. The constraints were specified as follows:

**Voltage:** \(-1 \leq V \leq 1\)

**Electrode k:** \(a'_k < a''_k, \ a''_{k-1} < a'_k\)

with \(k\) defined in Figure 3-17. With these constraints, the first optimizations were completed for the case of one and two electrodes (\(N=1\) and 2 in equation 3.132). Two objective functions were selected; one to maximize deflection, and the other to maximize volumetric displacement for high acoustic pressure. For maximized deflection, the objective function aimed to minimize the center deflection. Since the deflection profile yields negative displacements by sign convention, the minimum of the deflection profile is the maximum deflection that occurs at the center. For maximized acoustic pressure, the volumetric displacement was calculated by integrating the deflection profile with respect to \(\theta\) and \(r\). Again, given the sign convention, the objective function minimized the volumetric displacement, which is negative, effectively leading to maximized acoustic pressure. Each optimization was completed multiple times with varying initial conditions to evaluate the robustness of the generated solutions.
A bimorph PMUT geometry was initially selected for simplicity with dimensions outlined in Figure 4-3. For one electrode, maximum deflection was achieved through 60% electrode coverage and the application of 1V, which is consistent with the prediction by [44]. Since this result agrees with previous values, the optimization method should be valid for more complex configurations. With two electrodes, deflection was maximized with 1V applied at a center electrode with 60% coverage and -1V applied to an outer ring electrode covering the remainder of the device. The optimized two electrode configuration nearly doubled deflection compared to the single electrode case.

Finally, it is interesting to compare the maximum deflection to the maximum volumetric displacement case. Higher volumetric displacement is not necessarily achieved through maximizing deflection. Based on the optimization script, slightly greater than 60% electrode coverage is suggested for high acoustic pressure resulting in a reduced center deflection. The preliminary optimization results further confirm the validity of the deflection model and also validate the implemented SQP optimization approach. In the future, this optimization approach can be extended to more complicated and different design cases to optimize geometries for acoustic pressure and deflection.
Figure 4-4: Deflection profile with one electrode for maximized deflection and acoustic pressure
Chapter 5

Fabrication

From the modelling and past PMUT work, a simple, general design is chosen for the first generation device. Since optimal clamped plate deflection occurs with approximately 60% electrode coverage, the top electrode will be patterned according to this constraint. Residual stress is also recognized as a potential design problem - for now, it will be addressed by extending the top platinum layer to the edge of the element. A center electrode covering 60% of the diaphragm will be used for actuation and sensing and an outer ring of platinum separated by a few microns from the center electrode will extend platinum coverage to element edge. This additional platinum is introduced to remove the potential residual stress discontinuity introduced by incomplete platinum coverage, and to prevent plate bending with no applied voltage. From previous discussion, it is clear that residual stress will still influence plate bending, bandwidth and sensitivity, but for now, the performance of a first generation design will be used to determine the necessity of more detailed residual stress engineering.

5.1 Current Progress

The first generation PMUT device is currently under construction following the process guidelines provided in the Table 5.1. A schematic view of the process flow is shown in Figure
5-6. For the steps that have yet to be executed, process parameters are based on previous steps and similar fabrication steps that have been successfully completed [27].

5.1.1 Initial Processing Steps

The starting base material for the PMUT fabrication is a double-side polished, six inch, silicon-on-insulator (SOI) wafer. An SOI wafer was selected because of the buried oxide layer, which serves as an etch stop for the final DRIE release. With the variety of geometries available for purchase, the thickness of the device layer can also be selected for the appropriate diaphragm resonant frequency, and the oxide layer thickness should be chosen considering residual stress. The base layers are also selected to be thick enough so that the diaphragm deflection is dominated by plate bending and not membrane mode; otherwise, the developed analytical models will not be valid. For the first generation of the device, a 4μm device and 1μm buried oxide layer were selected for plate bending corresponding to a resonant frequency in the MHz range for plate radii between 10 and 100μm. A double-side polish was selected so that the front and back side alignment marks could be more clearly visible on the wafer surface; however, the back side polish is not critical. SOI wafers are generally more expensive than standard wafers. If price is a problem, a boron implanted layer could be used as an etch stop with a standard p- or n-type wafer [52].

PMUT fabrication begins with cleaning and wet oxidation. The thermal oxide serves as an electrical insulating layer separating the electrodes from the silicon. Since the oxide is not serving as a structural support, the thickness is not critical and has been chosen based on standard procedures. For the 22min oxidation time, the average oxide thickness (taken from four measurements on each wafer) ranged from 1995.62 ± 22 to 2072 ± 22Å for a given wafer. Thus, the chosen processing time was appropriate to achieve ≈ 2000 Å of oxide thickness. The oxide provides a compressive contribution to the overall device residual stress; however, given its small thickness as compared with the buried oxide (200nm as opposed to 1μ m) its overall contribution is negligible. For more advanced residual stress engineering, it might be necessary to consider the compressive contribution of the thermal oxide, but for the first
Shallow alignment marks are then etched into the surface for reference in subsequent deposition and etching steps. Front and back side photolithography is performed simultaneously. First, photoresist is coated on the wafer front side and then hardened with short post-bake before coating an additional layer on the back side. The post-bake is then completed for the front and back side photoresist. Exposure is performed similarly, first with the front side, and then the resist is developed for a short time until the alignment marks are barely visible. The back side alignment mark exposure is then performed referencing the barely visible front side alignment marks. Development for the full development time is then finished for the front and back side marks. For the etching process, a short etch with buffered oxide etchant (BOE) is performed to remove the surface oxide before a shallow plasma etch of the silicon. The resulting alignment marks are a few microns deep, which is sufficient for the alignment marks to be visible for later processing. Pictures of alignment marks are provided in Figure 5-1.

5.1.2 Bottom Electrode Lift-Off

Many iterations of the lift-off process were performed before a finally recipe was decided. It is likely that sidewall coverage of the photoresist as a result of rounded edges in the resist profile created lift-off problems (Figure 5-2a). Because of the sidewall coverage, residual metal flecks were present after cleaning along the edge of the electrode (Figure 5-2b), preventing further processing. Spin coating speeds, initial and flood exposure time and contact
type, image reversal baking conditions, contact aligner and electron beam physical vapor deposition equipment, development time, and cleaning procedure were all varied to improve lift-off. Since adhesion was not a problem, the thickness of the titanium adhesion layer and deposited platinum was kept the same for all processing. In some cases, oxygen plasma ashing was performed to de-scum the developed areas and improve adhesion before electron beam deposition. For de-scumming, 3min at a power of 800W in the TRL Asher was used. However, for the final successful lift-off process, the features looked the same between an ashed and non-ashed sample, so the ashing step can be omitted for this particular process.

From literature on the image reversal photoresist and experience, the post-bake conditions appeared to be the most critical to success. The best results were achieved through uniform heating of the wafer on a metal slab in a 120°C oven for 90s before flood exposure. The thickness of the photoresist, set by the 2000rpm spin speed, also seems to be important as a thinner resist is less able to form a break in the metal layer between the developed feature and top of the resist. With the same image reversal baking and other conditions, a 3000rpm spin speed creating a thinner layer of resist resulted in poor lift-off. The final, successful lift off process is included in Table5.1 and pictured in Figure5-2c.
5.1.3 PZT Processing

With the difficulties experienced during lift-off, the PZT processing steps were tested without patterning steps. A titanium adhesion layer and platinum was deposited, using the same electron beam deposition process, on a thermally oxidized, blank wafer. It is preferable to use platinum as the bottom electrode because its crystal orientation is compatible with the perovskite phase of PZT. Without a platinum seed layer, a properly oriented PZT film cannot be formed, which is visually evident through widespread cracking on the surface. 15 wt% PZT sol-gel purchased from Mitsubishi Materials is spin coated onto this fully platinum coated wafer. ≈2mL of PZT sol-gel is removed from the container with a syringe, and a 0.45 μm PTFE filter is screwed onto the end of the syringe before deposition. The filter removes larger, excess particles from solution, which can nucleate cracks in subsequent drying and pyrolysis steps when inadvertently deposited onto the substrate. The full volume of sol-gel was deposited from the syringe directly before or during the initial low speed sping at 500rpm for 5s.

Spin speeds and drying and pyrolysis temperatures in Table 5.1 are consistent with those in [27]. PZT is pyrolyzed at 350°C; however direct ramping to this temperature should be avoided as thermal shock will stress the film inducing cracking. Moisture can also modify film stress and induce cracking, so a lower temperature drying step is critical. Temperature steps between room and the pyrolysis temperature were included for film drying and to reduce the effects of thermal shock to form a smooth, crack-free film (Figure 5-3b). On the test wafer, the coating process was completed twice to confirm robustness of the coating process. No patterned was performed on the test wafer, so the PZT etch process described in Table 5.1 has yet to be performed; however, the same etching process was successfully used in [27].

Annealing was performed in a high temperature furnace. The test wafer was placed in a Quartz wafer boat sandwiched between two dummy wafers to protect the PZT surface from collecting unwanted particle contaminants present in the furnace environment. Oxygen flow to the machine was initiated and the temperature was ramped to the annealing temperature of 700°C. After approximately 2 hours, the furnace reached 700°C and the temperature
was set back to room temperature (20°C) beginning the ramp down process. The wafer was gradually cooled overnight and the final annealed substrate was removed the following morning. Again, a slow ramp down and ramp up were used to prevent cracking.

After annealing, the quality of the PZT test films was analyzed using x-ray diffraction crystallography (XRD) at the Center of Materials Science and Engineering (CMSE) at MIT. The Rigaku Cr-Source RU300 Rotating Anode X-ray Generator 185mm Radius Diffractometer was used to record crystal structure data. Intensity data was collected between 20° and 90° 2θ. Since the film was deposited on a single crystal platinum substrate, an offset of 1° and monochromator receiving slit of 0.45mm were selected to prevent damage to the machine and reduce the K-beta intensity of the platinum peaks, making the PZT diffraction pattern more clearly visible. Standard values were selected for the divergence slit, scatter slit, receiving slit, which were set to 1mm, 2mm and 0.3mm respectively. The crystal peaks corresponding to the perovskite phase of PZT are superimposed on the XRD scan indicating strong agreement between the sample and the perovskite crystal structure of $Pb(Zr_{0.52}Ti_{0.48})O_3$. Remaining peaks line up with the platinum diffraction pattern. From the XRD, the PZT film is consistent with perovskite phase $Pb(Zr_{0.52}Ti_{0.48})O_3$ and should therefore be highly piezoelectric.
Counts

Figure 5-4: Square root intensity scaling of XRD raw data with superimposed platinum and perovskite PZT peaks. Numbers on peaks indicate crystal orientation of PZT.

Peak List

Figure 5-5: Peak matching data with measured XRD peaks on top and perovskite PZT and platinum peaks underneath. Perovskite PZT and platinum match all of the recorded peaks in the XRD data.
5.2 Next steps

With successful lift-off and PZT pyrolysis and annealing, the groundwork for over the half of the PMUT fabrication has already been set. Deposition and lift-off of the top electrode should be possible with the same recipe used for bottom electrode lift-off. The final back side etch could prove challenging but is a standard process that can be refined and finished.

Subsequent packaging steps have yet to be fully addressed; however, it is assumed gold wire bonding to the contact pads on either side of the elements will be interfaced onto a ceramic pin grid array (CPGA). New mask designs will include a layout with electrodes running to the edges of an approximately 15mm by 15mm chip (appropriately sized to fit the CPGA) for easier wire bonding and testing. Chips would be released via the final back side deep reactive ion etch (DRIE) already included in the process plan for element release.
## Fabrication Process Steps

The process steps described in the table below are part of the fabrication process for a specific device. The steps above the bold line have successfully been completed.

### Wafer Cleaning
- **RCA clean**

### Thermal Oxide Deposition
- **Deposit 2000Å SiO₂**
  - **Processing conditions:** Recipe 3W1000 for 22min
  - **Location:** ICL RCA Station

### Front and Back Side Alignment Marks
- **Hdms deposition**
- **Spin coat OCG825 resist onto wafer front side**
- **Pre-bake**
  - **Condition:** 95°C for 15 min
  - **Location:** TRL HMDS, TRL Coater
- **Spin coat OCG825 resist onto wafer back side**
- **Pre-bake**
  - **Condition:** 95°C for 30 min
  - **Location:** TRL Pre-bake Oven, TRL Coater
- **Exposure top side alignment marks**
- **Develop**
  - **Condition:** Soft contact, 2.3s exposure
  - **Location:** TRL Pre-bake Oven, TRL Coater
- **Expose back side alignment marks**
  - **Condition:** Front to back alignment: hard contact, crosshairs, 2.3s exposure
  - **Location:** TRL Wet Bench, TRL EV1
- **Develop**
  - **Condition:** 90s with 934 1:1 developer
  - **Location:** TRL Wet Bench, TRL Pre-bake Oven

### Bottom Electrode Deposition and Lift-off
- **Image reversal photo and develop**
  - **Condition:** Spin coat resist AZ 5214E onto wafer front side
  - **Location:** TRL HMDS, TRL Coater
- **Pre-bake**
  - **Condition:** 95°C for 30 min
  - **Location:** TRL Pre-bake Oven, TRL EV1
- **Expose**
  - **Condition:** Hard contact, 1.4s exposure
  - **Location:** TRL Pre-bake Oven, TRL EV1
- **Image reversal bake**
  - **Condition:** Place wafer on slab in oven at 120°C for 90sec
  - **Location:** TRL Pre-bake Oven, TRL EV1
- **Image reversal exposure**
  - **Condition:** Flood exposure, 60s
  - **Location:** TRL Pre-bake Oven, TRL EV1
- **Develop**
  - **Condition:** 90s with AZ 422 MIF developer
  - **Location:** TRL Pre-bake Oven, TRL EV1
- **Deposit Ti**
  - **Condition:** 200Å at 1Å/s
  - **Location:** TRL Pre-bake Oven, TRL EV1
- **Overnight**
  - **Location:** TRL Pre-bake Oven, TRL EV1
- **Develop**
  - **Condition:** Acetone / methanol / 2-propanol
  - **Location:** TRL Pre-bake Oven, TRL EV1

### PZT Deposition, Etch and Anneal
- **Spin coat 15 wt% PZT solution by Mitshubishi Materials**
- **Drying and pyrolysis on hot plates**
  - **Condition:** 500rpm for 5s, 2500rpm for 25s
  - **Location:** TRL PZT spin-coater
- **Repeat deposition**
  - **Condition:** Spin coat 15 wt% PZT solution by Mitshubishi Materials
  - **Location:** TRL PZT spin-coater

### Etch
- **Hdms deposition**
- **Spin coat OCG825 resist onto wafer back side**
  - **Condition:** 95°C for 30 min
  - **Location:** TRL HMDS, TRL Coater
- **Pre-bake**
  - **Condition:** Soft contact, 2.3s exposure
  - **Location:** TRL Pre-bake Oven, TRL EV1
- **Expose top side alignment marks**
  - **Condition:** 90s with 934 1:1 developer
  - **Location:** TRL Pre-bake Oven, TRL EV1
- **Wet etch**
  - **Condition:** 20s in 1:1 DI water, HCl, BOE
  - **Location:** TRL Pre-bake Oven, TRL EV1
- **Clean**
  - **Condition:** Acetone / methanol / 2-propanol
  - **Location:** TRL Pre-bake Oven, TRL EV1

### Top Electrode Deposition and Lift-off
- **Image reversal photo and develop**
  - **Condition:** Spin coat resist AZ 5214E onto wafer front side
  - **Location:** TRL HMDS, TRL Coater
  - **Pre-bake**
  - **Condition:** 95°C for 30 min
  - **Location:** TRL Pre-bake Oven, TRL EV1
  - **Expose**
  - **Condition:** Hard contact, 1.4s exposure
  - **Location:** TRL Pre-bake Oven, TRL EV1
  - **Image reversal bake**
  - **Condition:** Place wafer on slab in oven at 120°C for 90sec
  - **Location:** TRL Pre-bake Oven, TRL EV1
  - **Image reversal exposure**
  - **Condition:** Flood exposure, 60s
  - **Location:** TRL Pre-bake Oven, TRL EV1
  - **Develop**
  - **Condition:** 90s with AZ 422 MIF developer
  - **Location:** TRL Pre-bake Oven, TRL EV1
  - **Deposit Ti**
  - **Condition:** 200Å at 1Å/s
  - **Location:** TRL Pre-bake Oven, TRL EV1
  - **Deposit Platinum**
  - **Condition:** 2000Å at 1Å/s
  - **Location:** TRL Pre-bake Oven, TRL EV1
  - **Overnight**
  - **Location:** TRL Pre-bake Oven, TRL EV1
  - **Develop**
  - **Condition:** Acetone / methanol / 2-propanol
  - **Location:** TRL Pre-bake Oven, TRL EV1

### Electrode Deposition
- **Image reversal photo and develop**
  - **Condition:** Spin coat resist AZ 5214E onto wafer front side
  - **Location:** TRL HMDS, TRL Coater
  - **Pre-bake**
  - **Condition:** 95°C for 30 min
  - **Location:** TRL Pre-bake Oven, TRL EV1
  - **Expose**
  - **Condition:** Soft contact, 2.3s exposure
  - **Location:** TRL Pre-bake Oven, TRL EV1
  - **Image reversal bake**
  - **Condition:** Place wafer on slab in oven at 120°C for 90sec
  - **Location:** TRL Pre-bake Oven, TRL EV1
  - **Image reversal exposure**
  - **Condition:** Flood exposure, 60s
  - **Location:** TRL Pre-bake Oven, TRL EV1
  - **Develop**
  - **Condition:** 90s with AZ 422 MIF developer
  - **Location:** TRL Pre-bake Oven, TRL EV1
  - **Deposit Platinum**
  - **Condition:** 2000Å at 1Å/s
  - **Location:** TRL Pre-bake Oven, TRL EV1
  - **Overnight**
  - **Location:** TRL Pre-bake Oven, TRL EV1
  - **Develop**
  - **Condition:** Acetone / methanol / 2-propanol
  - **Location:** TRL Pre-bake Oven, TRL EV1

### Lift-Off
- **Sonicate in acetone**
  - **Condition:** 10min
  - **Location:** TRL Photo Wet Au, TRL Photo Wet Au
- **Clean**
  - **Condition:** Acetone / methanol / 2-propanol
  - **Location:** TRL Photo Wet Au, TRL Photo Wet Au

### Back Side Etch Release
- **Hdms deposition**
  - **Spin coat OCG825 resist onto wafer back side**
  - **Pre-bake**
  - **Condition:** 95°C for 30 min
  - **Location:** TRL HMDS, TRL Coater
  - **Expose top side alignment marks**
  - **Condition:** Soft contact, 2.3s exposure
  - **Location:** TRL Pre-bake Oven, TRL EV1
  - **Develop**
  - **Condition:** 90s with 934 1:1 developer
  - **Location:** TRL Pre-bake Oven, TRL EV1
  - **Deep reactive ion etch**
  - **Condition:** Recipe MITS6 at 3um/min [27]
  - **Location:** TRL Photo Wet Au
  - **Clean**
  - **Condition:** Acetone / methanol / 2-propanol
  - **Location:** TRL Photo Wet Au
Figure 5-6: Cross-section views of PMUT device throughout fabrication process. Starting from top left and proceeding left to right down the page: oxide growth, bottom electrode lift-off, PZT after wet etch, top electrode lift-off, and back side DRIE.
Chapter 6

Summary

6.1 Conclusions

Design and modelling of a circular clamped plate PMUT is detailed focusing on maximizing deflection and volumetric displacement for high acoustic pressure output. The design framework for a small form factor, high resolution medical ultrasonic transducer is outlined based on an axiomatic design approach with emphasis on high sensitivity, low power consumption, and high acoustic power. Functional requirement and design parameter hierarchies are established creating decoupled design matrices that can be used to guide the PMUT design process. In the proposed design process, the effects of residual stress pertaining to bandwidth, resonant frequency and sensitivity are considered and analyzed for effective PMUT design.

A general form of the plate vibration equation is derived that applies to any stack of piezoelectric and supporting material, and is particularly solved for the bimorph and unimorph PMUT. From basic acoustic principles, the radiated acoustic pressure is determined based on the plate deflection profile and is compared to that of an rigid piston with equal area operating at the maximum and average displacement rate of the plate deflection. The acoustic pressure output is shown to scale with the volumetric displacement rate of the plate deflection and can be approximated by a rigid piston with equal area moving at the average
A novel Green’s function approach is used to solve the plate vibration equation. An auxiliary function, known as the Green’s function, is defined based on the homogeneous solution of the plate vibration equation and an assumed arbitrary driving force singularity. From the piezoelectric layer, there is an applied piezoelectric moment that is added to the plate vibration equation in terms of a forcing function. With the Green’s function and the plate vibration equation including the piezoelectric forcing function, the plate deflection equation is explicitly solved for an arbitrary number of circular or ring electrodes. The resulting solution is less complex, more easily computed, and more flexible than existing solutions based on an energy minimization approach. For validation, the Green’s function solution is compared to the energy-based solution [21], [44] for the case of a single electrode cover 60% of the plate surface with strong agreement.

Preliminary optimization with an SQP approach shows that the optimum coverage for a single electrode is approximately 60% as previously reported in [44]. For a two electrodes, opposite applied voltages at a center electrode with 60% coverage and an outer ring electrode covering the remainder of the plate results in twice the displacement proving to be the optimum configuration for the two electrode case. Maximizing for volumetric displacement and thus high acoustic pressure output results in slightly larger center electrode coverage to increase the average displacement rate.

Based on the modelling results, an initial PMUT design is proposed with 60% electrode coverage with an outer electrode ring that serves to minimize the effects of residual stress. Fabrication of the device is currently underway. Successful bottom electrode deposition and PZT pyrolysis and annealing has been completed. With XRD, the appropriate perovskite crystal structure is shown to be present in the fabricated PZT, which will be critical to later device performance.
6.2 Future Work

The most immediate next step is the completion of a fabricated first generation device. Once complete, the acoustic pressure and deflection can be quantified and compared with the proposed models for experimental validation. From the deflection profile, the equivalent circuit model can be analytically determined for the clamped, circular plate PMUT and compared to impedance data.

Based on the performance of the first device, the established axiomatic design framework can be used to design a second generation device hopefully improving on the performance of the first device. The Green's function solution process can also be used to analyze the effects of residual stress on the PMUT design and further work can be focused on minimizing this residual stress. With the Green's function approach and outlined optimization tools, a high acoustic pressure design can be optimized for more complicated design scenarios including the second generation design. When a successful transducer design has been fabricated, future work will focus on integration into one- and two-dimensional arrays for two- and three-dimensional imaging.
Bibliography


