On the Fluidic Forces and Shape Optimizations of Resonant Curved Cantilever Wings

On the Fluidic Forces and Shape Optimizations of Resonant Curved Cantilever Wings

By Andrey D. Goussev, B. Eng.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree of Master of Applied Science

McMaster University MASTER OF APPLIED SCIENCE (2018) Hamilton, Ontario
(Engineering Physics)

TITLE: On the Fluidic Forces and Shape Optimizations of Resonant Curved Cantilever Wings

AUTHOR: Andrey Dmitrievich Goussev, B.Eng. (McMaster University)

SUPERVISOR: Professor Rafael N. Kleiman

Number of pages: xiii, 113

ABSTRACT

Artificial flight on millimeter size scales has been a major challenge due to the difficulty in making a feasible flight mechanism in terms of fabrication, thrust and power used. Many have tried to copy animal flight but there has been little success at such sizes. One proposed solution is to make small thrusters out of resonant curved cantilevers which act as wings that follow a simple 1 degree-of-freedom motion. Such wings are free of joint friction, can be planarly fabricated using well documented techniques, can be predictably scaled to different sizes, and have been shown to generate a net thrust.

In this thesis, the work investigates the nature of the wings' thrust through thorough studies of computational fluid dynamic simulations to understand how they interact with the surrounding fluid and how exactly the forces are generated. Specifically, it considers the role of unsteady lagged fluid waves generated by the wings and explains how the wing-fluid interactions relate to drag coefficients at low to high flapping amplitudes and Reynolds numbers ranging from 10^2-10^5. It then studies the effect of different wing aspect ratios on the net force and power efficiencies. The results are then extended to a general dependence on the wings' aspect ratio which allows for this parameter to be used in optimizing the wings' net force/power used. Test wings are then made using an updated fabrication method and Molybdenum as the curve-inducing material in an attempt to produce more environmentally-stable wings with important successes, failures and improvements discussed. Results show that such Molybdenum-based wings are practical for flight, and that resonant curved cantilevers wings can be made more feasible by simple changes to their shape.

ACKNOWLEDGEMENTS

Completing this thesis took many hours of work and dedication, yet no amount of my own effort would have sufficed to replace the tremendous support of several other people. I am enormously grateful for the help from my supervisor Dr. Rafael Kleiman for not only funding this research and giving me office space, but also for his advice, his always friendly and motivating attitude to test new theories, and his trust and patience with me. Our weekly group meetings were engaging, and I would like to thank the rest of Dr. Kleiman's PV group for the many useful discussions and suggestions for my work. Thank you to Dr. Abhi Rampal for *always* being around to help me with lab work and questions, and for letting me into the building on cold, winter nights when I did not have my keys. Thank you to Dr. Kevin Boyd for helping me with the FTIR and proposing its use in the first place.

A lot of the results and fabrication would not have been possible without the help from Dr. Shahram Tavakoli and I am thankful for the many hours he spent fixing the RIE, teaching me how to use the sputtering machine, and his general guidance during my lab work. Thank you to Doris Stevanovic for giving me advice on how to measure my samples, for her help in the cleanroom and help in preparing my samples. Thank you to Jim Garret for teaching me about the lab equipment necessary for developing shadow masks, and his advice on sputtering. I must also thank my close friends and family for their encouragement, love and limitless support. They helped me stay goal-orientated and supported me in this endeavour while providing me with peace, love and food during difficult and frustrating times.

Now I would like to thank Dr. Matthew Minnick. He introduced me to his PhD work late in my undergraduate career that led to the pursuit which resulted in my own thesis. He has been a wonderful mentor since our first meeting and introduced me to the field, the theory, lab equipment, staff, software and much more. Thank you for believing that I could do this from the beginning, and I am grateful for the amount of knowledge and guidance given to me – regarding this thesis work and beyond.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

A – Wing/cantilever area

a – proportionality constant for A_r and $\sqrt{V}$

A_sR – Aspect ratio defined as 2L/B

B –Wing/cantilever width

C_{DN} – General transient force drag coefficient

C_{DN0} – Force drag coefficient particularly for a wing with $A_sR = 8$

C_{DP} – General transient power drag coefficient

C_{DP0} – Power drag coefficient particularly for a wing with $A_sR=8$

C_m – Added mass coefficient that relates to the relative magnitude of the virtual mass to real mass

f – frequency (measured in Hz)

Fr – Froude number

g – Gravitational Acceleration

h – Wing/cantilever thickness

K – Air damping constant

l – Characteristic length scale used in the definition of Fr

L – Wing/cantilever length, measured from the base to the tip.

m – mass

m_a – virtual mass

MEMS – Micro-Electro-Mechanical Systems

P – Pressure

Q – Quality Factor

Re – Reynolds number

Re_8 – Reynolds number definition of a wing with $A_sR=8$

t – Time

TFM – Transient Force Magnitude. The magnitude of the vector sum of the z and y force components on the wings, taken over the wing area.

V – relative velocity between the object and the fluid infinitely far away from the object's surface.

V_{rms} – Spacial Root Mean Square of the velocity of the resonating cantilever/wing as it passes through its equilibrium point.

γ – Frequency correction factor

θ – phase difference between the wing/cantilever position and the lagged mass/wave

μ – Dynamic Viscosity

ρ – density

ϕ – bending angle formed by the curved wing (measured in degrees)

ω – Resonant frequency (rad)

1 INTRODUCTION

For over a century, scientists and engineers have dominated the sky with large flying vehicles, yet when it comes to millimeter size scales, we have largely been unsuccessful. The main difficulty comes from the different forces coming into play when scaling large-scale technology like planes and rockets. Such propeller or jet-driven flight becomes impossible, or at least impractical, due to the large relative mechanical friction, increased difficulty in fabrications, decreased power efficiency and very different aerodynamics effects to consider. To overcome this, different flying mechanisms need to be used while considering current limitations of MEMS (Micro-Electro-Mechanical Systems) and which aerodynamic phenomena to exploit.

The most obvious starting point is the usual case: to observe and learn from nature. This is reasonable since flying animals commonly operate on the millimeter size scale with some animals like the fairyfly reaching lengths as low as 158um [1]! Unfortunately, the aerodynamics of flapping-wing flight cannot be fully explained by current steady-state (i.e., airfoil type) theories, although many have tried. Only recently has the complex aerodynamics of such flight been thoroughly investigated and so-called, unsteady mechanisms, discovered, although exact descriptions of animal flight are still highly debated.

Instead of fully understanding the nature of animal flight, many have tried to simply replicate it. The problem here is that it is very difficult to replicate the wing motions usually seen in birds and especially in insects. The joints and tissues connecting the wings allow for timely rotating,

pitching and heaving motions while the structure of the wing itself can add to the efficiency through hairs initiating turbulent flow or vortex generations from their elasticity [2]. Furthermore, it is not obvious which animals to imitate since flapping motions vary between fliers and can depend on their size, wing shape, flying style or on some environment-specific evolutionary factor (like high winds). Thus, engineers are left with deciding which flight mechanism is sufficient for the task and how to recreate these fast, complex joint movements without breaking.

Minnick proposed in [3] that a much simpler solution could be used that may not replicate the high efficiencies found in nature but would work sufficiently well to power very small flying vehicles. The idea is to use resonating curved cantilevers at high amplitudes to generate a net thrusting force. This would involve no rotating parts and the only movement would come from the high amplitude vibrations of the cantilevers. Some major benefits would be

- Relatively easy construction
- No joints (so low friction loss and longer lifetimes)
- Easily scalable to different sizes (from centimeter to below millimeter-scales)
- Monolithic fabrication that allows them to be fabricated in large quantities in parallel

The ground work for these devices was developed in terms of construction, resonant properties, dimensional analysis and fluidic force expectations based on look-up tables from CFD simulations. Minnick's work on this technology has also shown that theoretically such wings can generate forces higher than their own weight, and enough to support onboard power supplies and sensors.

Initial results supported the theoretical predictions, yet at the time, Minnick was only able to create wings that lifted 2-7% of their body weight. Much of this was due to poor materials and possibly an inefficient shape design based on insects' dimensions. Also, little was understood about how the wings generate force, since the completely different flapping motion not seen in nature suggests that there are new thrust mechanisms at play.

In this thesis I will improve on Minnick's wings by first explaining the force phenomena seen in these curved cantilevers undergoing large oscillations by examining simulation results. I will then expand these results by simulating wings of different aspect ratios to predict optimal wing

designs and present results for fabricated wings that I made using more stable materials. With these results and explanations, one can bring these devices to generate forces higher than previously thought possible, with a more complete knowledge of the underlying fluid dynamics that can be considered in future work.

Literature Review

Much of the work leading up to NAVs (Nano-Air Vehicles) began with the understanding of steady-state airfoil theory developed in the early 20th century. Unfortunately, scaling of airplanes and helicopter-like machines to centimeter-sizes becomes very difficult due to the complexity of fabricating small components and large relative friction at smaller sizes that is associated with rotating parts. Naturally, NAV scientists tried to overcome this by taking lessons from nature's best fliers: insects and birds.

Early attempts at modelling these animals used the same steady-state aerodynamic theories used for airplanes/helicopters and assumed that wings would function exactly like airfoils; thus, by dividing the wing's flapping cycle and knowing its position/velocity from quickly captured images, one could apply steady-state solutions and reproduce the forces/power used by the wings over a full flapping cycle. Indeed, for a fast travelling bird or insect, where the wing flapping speed is slower than the forward flight speed, this approximation held well [2], [4], [5]. What this method failed to account for were unsteady transient effects that occur specifically for airfoils/wings undergoing large and rapid oscillations. As a result, most approximations using this "quasi-steady-state" theory led to large underestimations of the forces generated in flight, and the phenomena of animal hovering was left unexplained [4], [6]. Lots of work, speculation, and experiments went into characterizing and explaining how some of the most unlikely flyers like bumble bees and fruit flies can manage to stay in the air.

Weis-Fogh [6] presented a theory to account for the extra lift force via his "clap and fling" mechanism, where large bound vortices are created at stroke reversals which can add significantly to the lift as shown in Figure 1

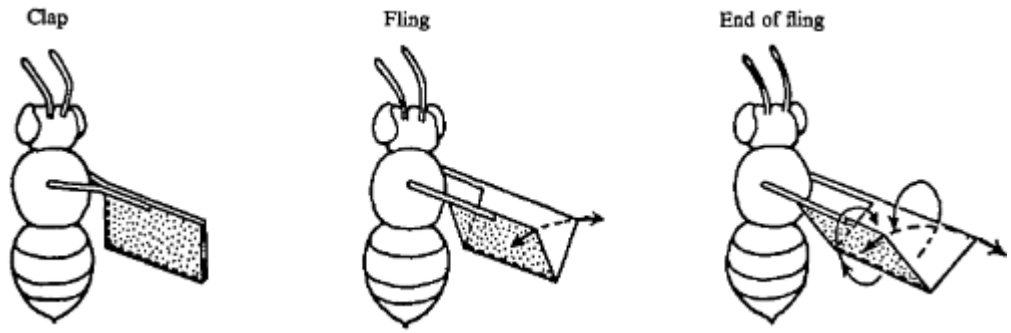


Figure 1: Clap and fling mechanism presented by Weis-Fogh to account for extra lift generation in some animals [6]. As the wings peel away from each other, large bound vortices are created which follow along with the wing during the flapping cycle. These vortices can be large enough to account for the extra lift in some animals.

However, this idea was not enough to explain flight in animals that flapped at low amplitudes and those that fly at low Reynolds numbers. Furthermore, the entire theory was still based on quasi-steady state theory, so it did not account for unsteady transient effects that were likely happening at points in the cycle of the oscillating wing [6]. In continuation of his work, Ellington explains in [7] that the leading-edge vortex (LEV) is the primary lift generating mechanism that most animals use. The phenomena described is that of delayed stall seen in airplanes where large lift occurs above the stall angle that can last for several chord lengths through the LEV created on the airfoil. Normally, in planes, this effect quickly dissipates but due to the cyclic flapping nature of animal wings, this effect can be used throughout most of each half of the flapping cycles. It is now generally agreed that the major contributions to lift come from some combination of the clap and fling, delayed stall, rapid pitching of the wings at stroke reversal and potentially some wake capture from vortical flow, although the amount of contribution from each mechanism is still a major subject of debate [8]. Most importantly, all these described mechanisms required that the wing still have a forward velocity relative to the fluid; that is, the wing never simply moves up and down, but heaves, pitches and performs complex flapping patterns mimicking those of flying animals.

Eventually, engineers attempted to incorporate and replicate flapping wings at large scales to produce machines that could benefit from the high efficiency of a steady, forward flying airfoil, with the potential of hovering. These ‘ornithopters’ typically failed to increase the efficiency by

flapping and produced little vertical force when not moving forward in the fluid. In [9], Bennet et al. shows an early example of a 10-foot ornithopter undergoing large amplitude flapping motions at different pitching angles as seen in Figure 2, although this design suffers from the issues mentioned earlier. Smaller designs were eventually made as in [10]-[11], which were able to perform fast maneuvers such as rapid pitching and flapping, and successful flight was achieved on the size scale of birds – although no hovering yet.

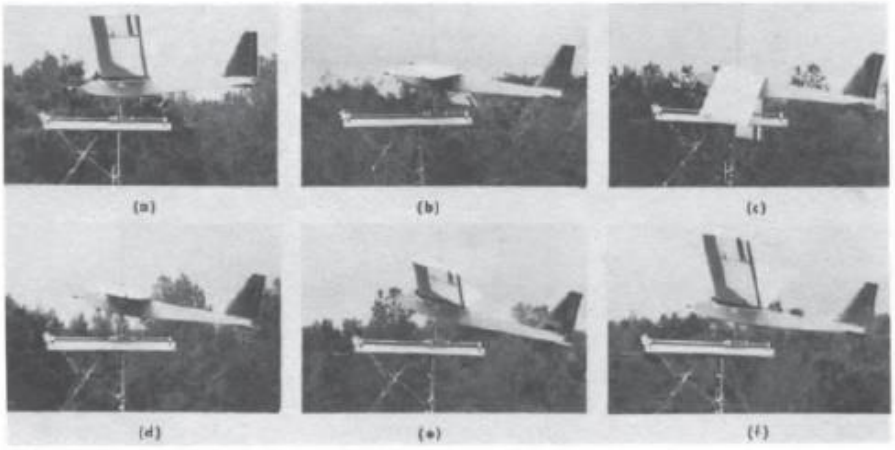


Figure 2: A typical large scale ornithopter where the different subfigures show the movements of the airfoils [9].

Karpensen et al. [12] did a thorough review of certain actuation mechanisms that could be used to power flapping-wing flight for NAVs and smaller aerial vehicles. Among thermally and statically driven options, they considered piezoelectrically actuating wings using cantilever bimorphs and unimorphs. Using known resonant frequency values for cantilevers in small oscillations, along with other dimensional parameters and the blocking force, they predicted the maximum force output of such devices, and provided a dimensional analysis. However, since no exact actuating mechanism was presented nor were any aerodynamics considered, they assumed the blocking force (minimum force to stop cantilever from curving) as the maximum thrusting force, which ignores many of the losses and aerodynamic complexities an oscillating wing in a fluid can experience. Also, no actual attempt to make this a reality was done - at least in using pure resonating cantilevers for thrust. That being said, they did present some interesting

frameworks, initial ideas and rough estimations for the practicality of resonating cantilevers in NAVs, leaving future engineers to come up with a way to use these actuators for thrust.

Later, Harvard's Robobees project successfully demonstrated flight of the smallest functioning ornithopter to date (about 3cm wing span and 100mg weight), by using piezoelectric actuators at the wing joints to control and imitate the flapping motions of similarly sized animals [13]. Although hovering has been achieved, 4 main problems still exist:

1. The fabrication process is difficult
2. The joint structure is brittle
3. It requires an external power source
4. It requires an external computer to control the wing

Also, the hovering mechanism employed will likely need to be adjusted for different sizes, as it is known that insects' flap differently and use more of certain mechanisms depending on their relative size and physical ability [5].

A couple of years later, Minnick [3] proposed that resonating cantilevers *alone* can be used to generate a net thrust by making them initially pre-curved. Having the base layer be made of a piezoelectric material, the cantilever could be actuated electrically at large amplitudes controlled by the generated electric field through the contact layers, with the Reynolds number proportional to resonant frequency which is governed by the cantilevers' material properties. He laid out an extensive groundwork regarding the mechanical-electrical-fluidic coupling of such a wing, including their expected resonant shape, expected resonant frequency, and expected force (found through Computational Fluid Dynamics (CFD) simulations). He showed that it could theoretically be used as a practical flight mechanism and built a test model that verified parts of the theory. This method has the benefit of easy fabrication/parallelization, less brittle joints and easily scalable from centimeter to millimeter length scales.

This radical design showed that it was not necessary to replicate nature to obtain useful, working NAVs at such size scales. He laid out a strong framework for analyzing, testing and fabricating these wings, although the aerodynamic effects regarding net force generation and power used were not well understood; that is, the CFD drag results only presented a look-up table of values for a specific rectangular wing shape. Furthermore, the loose tolerances and material choices

during fabrication resulted in inconsistent results that made those exact wings impractical for use in NAVs (discussed in further detail in the next section, “background information”).

In earlier work, many authors did investigate fluidics around a resonating cantilever, made possible by approximating the cantilever as a series of oscillating spheres or as an infinitely long cylinder. The spherical solution uses already solved analytical expressions for a sphere oscillating in a fluid, and considers a cantilever as a ‘string’ of such spheres, where fitting constants are used for numerically accurate results [14]. The cylindrical method accounts better for the drag created by a solid body and is analytically solved by some authors directly from Navier-Stokes equations, as in [15]. Both assume small amplitude vibrations, which allows for the use of the simpler, linearized Navier-Stokes equations to obtain the full solution of the fluid velocity field around the spheres or cylinder. I have not found analytical solutions for large amplitude vibrations, which is probably due to the complexity in solving the non-linearized Navier-Stokes equations - a problem usually tackled through numerical means as done by Minnick. Also, I found that, generally, authors are not interested in fluid force generation in resonating cantilevers, but instead focus on the expected resonating frequency shift and the change in the quality factor Q – understandably, due to the common use of thin cantilevers as timing and sensing devices.

There has been work regarding flat plates oscillating at large amplitudes in fluids, exploring eddy vortices [16], drag coefficients [17], and fluidic responses for oscillating plates undergoing timely pitching motions [8], [18]–[20], although any reference that considered force on the plate shows that the force over a cycle canceled out if the plate is oscillated with 0 pitching angle, due to the symmetry in the motion. As a result, for practicality, any investigation of aspect ratio or wing shape on large oscillation motions is done with non-zero pitching angle, where an actual cycle-average force can be seen. Also, even for large oscillations, many models considering pitching motions use small-amplitude oscillation concepts from aeroelasticity theory from [21] and [22], due to the similarities in motion for elastic airfoils experiencing ‘fluttering’.

Since Minnick’s wings generate force with a simple oscillating motion with no pitching, undergoing large oscillations and a resonating mode that differs greatly compared to pitching and heaving motions (see background information for detail), it is not obvious which results, if any,

can be extended to Minnick's work. Thus, to progress Minnick's piezoelectric resonating cantilever for use as practical NAV wings, I chose to investigate:

1. General unsteady, non-linear transient force effects on the wing by investigating a simple flat-plate and lagged wave model, which potentially are extendible to some part of the force generation theory in the wings.
2. The patterns in force generation and power used by these cantilevers by thoroughly examining forces and fluid motions occurring throughout the flapping cycle.
3. Different aspect ratios of the wing to find more optimal wing designs and extend the known parameter space.
4. Fabricate the wings using different materials, fabrication processes, and tighter tolerances, while confirming results with theory.

Background Information

In this section I will present the most important and relevant aerodynamic concepts and other theory necessary to follow the subsequent chapters. The more specific equations and figures will be presented in the chapters as needed. Also, as this thesis focuses on advancing knowledge of Minnick's resonating curved cantilever devices, much of the background information will summarize important points from his work; for further detail, the reader is encouraged to read his work in [3].

1.1.1 Wing Structure and Operation

The wing is built as a multimorph structure essentially containing a piezoelectric center followed by a tensile/compressive layer and a conductive layer. The structure used by Minnick is shown in the following table in order of layers, along with the description and thickness:

Table 1: Stack structure for the wing made in [3] in top-bottom order.

Stack Material (from top to bottom order)	Description	Thickness
SU-8	Tensile Layer	7.6um
Au	Conducting Layer	150nm
Cr	Adhesion layer for Au and Quartz	25nm
Quartz (x-cut)	Piezoelectric Layer	7.6μm
Cr	Adhesion layer for Au and Quartz	25nm
Au	Conducting Layer	150nm

The conductive layer is necessary to activate the piezoelectric strain from the electric field, while the SU-8 layer is used to create tension relative to the piezoelectric layer, to give an initial 'pre-curvature' (i.e., curvature with no applied field) and thus bending angle, ϕ. The asymmetry about the quartz layer is also necessary for the wing to bend when the quartz expands/contracts in response to the electric field. An illustrative example of a released wing with the initial bending angle is shown in Figure 3, where ϕ is defined as the angle made by the circular arc, h is the thickness and L is the wing length.

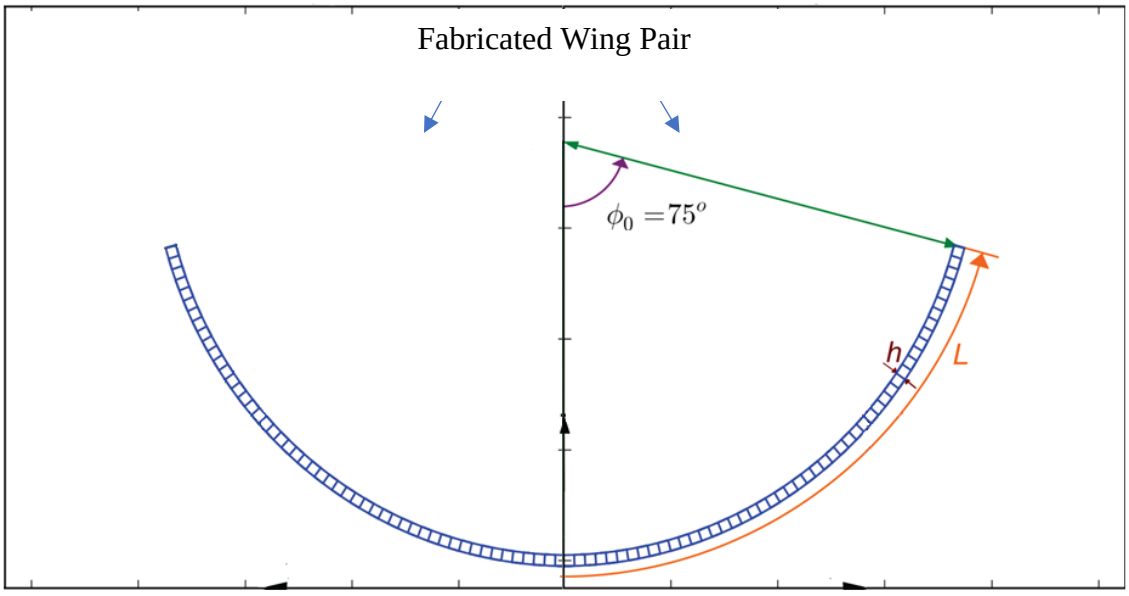


Figure 3: Example of fabricated wing pair with an initial bending angle, ϕ, of 75°.

In his analysis of resonating curve cantilevers, Minnick postulated that for a wing with thickness (h) much less than length (L), the infinitesimal sections of the cantilever behave as those of a flat cantilever, and during resonance, their local curvature changes the same amount as a flat cantilever at an instant in time. This allowed him to derive analytical expressions for the first resonant mode shape and transient behaviour of thin cantilevers undergoing large oscillations at any given pre-curvature. In the solution, for a given cantilever length, the oscillation amplitude (A_r), along with ϕ, define the motion completely. An example of the mode shape is given in Figure 4, where the cantilever has a ϕ of 75°, and the shapes at the different A_r represent the mode shapes at the cycle maxima; that is, a wing resonating with A_r=0.7 will curve into that shape seen at A_r=0.7 at ¼ cycle (starting from the static bending shape), and the shape at A_r=-0.7 at the ¾ cycle point.

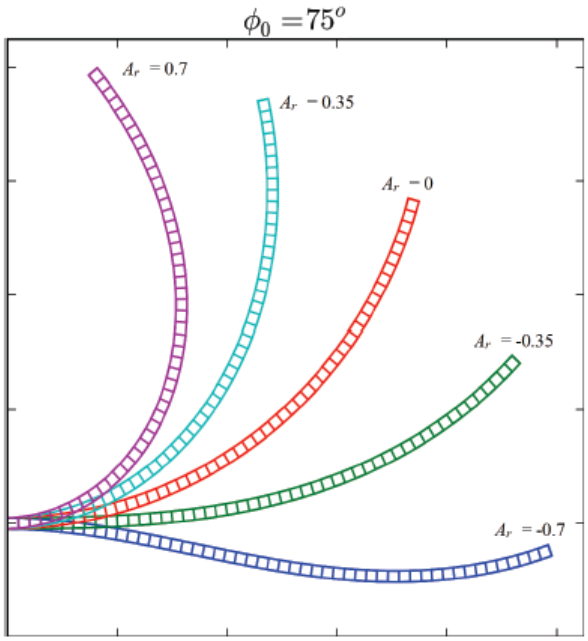


Figure 4: Resonant modes for a wing with ϕ=75° at different A_r (oscillation amplitude) values [3].

Regarding resonant frequency, ω, he found that higher ϕ has a stiffening effect, increasing the resonant frequency by a factor dependent only on ϕ. For a sense of the magnitude of this effect, for a 75° bending angle, the resonant frequency would be about 1.035 times the expected frequency for a straight cantilever. In a fluid, however, it is known that resonant frequencies decrease for resonating cantilevers due to the added virtual mass of the fluid that follows

"attached" to the cantilever, which also has the effect of decreasing the Q. For small oscillations, the frequency change is approximated by the following equation [15]

$$\frac{f_f}{f_v} = \sqrt{\frac{m}{m+m_a}} = \frac{1}{\sqrt{1+C_m \frac{\rho_f}{\rho_b}}} \tag{1.1}$$

$$\frac{f_f}{f_v} = \sqrt{\frac{\rho_b hLB}{\rho_b hLB + \rho_f LB^2}} = \sqrt{\frac{A_s R * h\rho_b}{2L\rho_f + A_s R * h\rho_b}} \tag{1.2}$$

Where f_f is the frequency in the fluid, f_v is the frequency in vacuum, *m* is the mass of the cantilever, m_a is added virtual mass of the fluid, C_m is an added mass coefficient, ρ_f is the density of the fluid and ρ_b is the average density of the beam. Here, we see that as a first order approximation, virtual mass alone governs the change in oscillation frequency, even when the authors solved the problem to account for viscous damping. For small oscillations, this virtual mass can be considered as a cylinder of air with diameter and length as the width (B) and length (L) of the cantilever, respectively [22]. Thus, for a cantilever, the above equation becomes

$$\frac{f_f}{f_v} = \sqrt{\frac{\rho_b hLB}{\rho_b hLB + \rho_f LB^2}} = \sqrt{\frac{A_s R * h\rho_b}{2L\rho_f + A_s R * h\rho_b}} \tag{1.3}$$

where for convenience, I rewrite B in terms of the aspect ratio: $A_s R = \frac{2L}{B}$. The $A_s R$ is a useful parameter which is investigated in Chapter 5 and is a common wing parameter used in airfoils and the study of flapping wings (see Figure 5 for details).

Now, although $\rho_f << \rho_b$, the length is much larger than the thickness, so we cannot ignore the $2L\rho_f$ term in (1.3). In the case of Minnick's wing that has an $A_s R$=8, (under the assumption that this relationship is accurate for large amplitudes) then

$$f_f = f_v \sqrt{\frac{8*15.6\mu m*\rho_b}{2*7.6mm*\rho_{air} + 8*15.6\mu m*\rho_b}} = 0.969\, f_v \tag{1.4}$$

So, fortunately for Minnick's chosen A_sR, the decrease in resonant frequency is small and was acceptably excluded from the framework, but it is important to note the dependence of the expected decrease in frequency on A_sR.

Interestingly, Kirstein et al. show in [15] that the Q is inversely proportionate to this frequency decrease because of the added mass effect at higher Re; however, for lower Re and high viscous damping relative to mechanical damping, the Q is expected to decrease. Note again, that these are first order approximations because the formulas presented were derived for small oscillations of a straight cantilever. In the case of Minnick's cantilevers, the Q does decrease, but stays relatively large (around 40), resembling an underdamped system; thus, effects to the mode shape are considered negligible, while fluidic effects on frequency are ignored for the given A_sR.

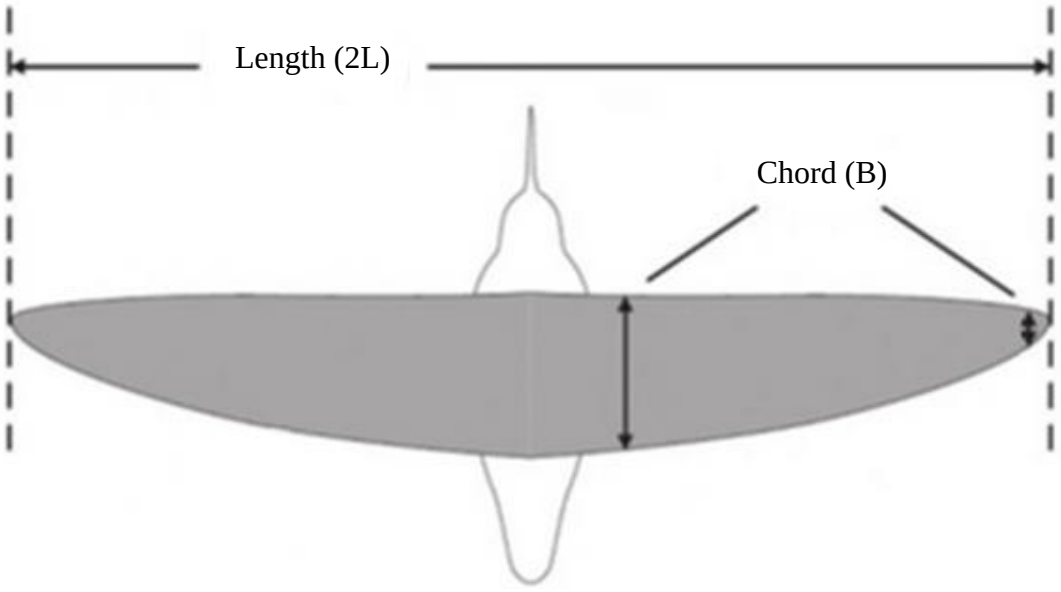


Figure 5: Dimensions typically used in studies of animal flight and airfoils. Image taken from [23].

1.1.2 CFD Simulations

Knowing how the cantilever is expected to move allows one to predict the fluidic reactions to the motion and shape. Such reactions are completely governed by the Navier-Stokes equations. The non-dimensionalized version accounting for incompressible flows and no heat transfer is:

$$\frac{\partial \vec{V}^*}{\partial t^*} + (\vec{V}^* \bullet \vec{\nabla}^*)\vec{V}^* = -\vec{\nabla}^* P^* + \left[\frac{1}{Fr^2}\right]\vec{g}^* + \left[\frac{1}{\mathrm{Re}}\right]\vec{\nabla}^{*2}\vec{V}^* \tag{1.5}$$

where,

$$\mathrm{Re} = \frac{Inertial\ Forces}{Viscous\ Forces} = \frac{\rho V L}{\mu}$$

$$Fr = \frac{Inertial\ Forces}{Gravitational\ Forces} = \frac{V}{\sqrt{gl}}$$

$\vec{g}$ = Gravitational Acceleration Field

g = Gravitational Acceleration

L = Body Length Scale

l = Characteristic Length Scale

μ = Dynamic Viscosity

P = Pressure

ρ = Fluid Density

t = Time

$\vec{V}$ = Fluid Velocity Field

V = Fluid Velocity Relative to Body

* denotes the dimensionless value

The dimensionless values Fr and Re result from the non-dimensionalization process, referring to the Froude and Reynolds number, respectively. The benefit of using the equation in this form is that one can use the solutions at different dimensional scales, so long as the dimensionless values Fr and Re remain equal. For the case of a flapping wing, the inertial force is much greater than the gravitational force, and so Fr is large and the respective term in (1.5) is negligible; thus, Re alone governs similar flow solutions. In other words, the problem can be scaled to any dimensions defined by Re and as long as Re is the same, the solution to the dimensionless Navier-Stokes equation will be the same as well.

This concept of Re scaling is used with steady-state drag coefficients for objects in motion. Such an object travelling at a velocity V in a fluid of density ρ_f will experience a drag force given by

$$F = \frac{1}{2} C_{DN}(\mathrm{Re}) \rho_f V^2 A \tag{1.6}$$

Where C_{DN} is the drag coefficient and A is an area describing the object. In terms of Re, this becomes

$$F = \frac{1}{2} C_{DN}(\mathrm{Re}) \frac{\mu^2}{\rho} \mathrm{Re}^2 \tag{1.7}$$

Thus, we can expect that the drag force on an object is independent of its size, if Re is the same.

Another useful metric is the power used by the travelling object to move the fluid. This can be similarly defined as

$$P = V \frac{1}{2} C_{DP}(\mathrm{Re}) \frac{\mu^2}{\rho} \mathrm{Re}^2 \tag{1.8}$$

Where C_{DP} is the power drag coefficient.

It is important to note at this point that equations (1.6)-(1.8) are typically used only for steady-state cases. Some authors have extended the use of these steady-state drag definitions to transient problems involving objects oscillating in fluids [3], [17], [24]. Minnick did this to both the force and power by averaging them over a cycle and defining the parameters in Re as

$$\mathrm{Re} \equiv \rho_{fluid} \frac{V\sqrt{LB}}{\mu} \tag{1.9}$$

$$V = V_{rms} \equiv \frac{A_r \omega L}{2\gamma} \tag{1.10}$$

where ω is the resonant frequency, L is the cantilever length, B is the cantilever width, γ is the frequency correction factor (i.e., the stiffening factor mentioned earlier), and A_r is the oscillation amplitude constant that governs the resonant shape of the cantilever for a given ϕ (as shown in Figure 4). The V_{rms} definition comes from taking the RMS velocity over the wing length, when it passes the equilibrium (maximum velocity) position.

Unfortunately, finding analytical solutions to the C_{DN} and C_{DP} dependence on Re and the flapping shape (governed by A_r and ϕ), is extremely difficult due to the non-linearity in (1.5) and the unsteady nature of the problem. Instead, 3D Computation Fluid Dynamics (CFD) simulations

were done for a cantilever of $A_sR=8$ for a wide range of Re, A_r, and ϕ. Minnick also confirmed through simulations that the concept of Re scaling still works for these unsteady, transient cases, as long as the cantilever's A_r, A_sR, ϕ and Re remain fixed. The exact CFD solver, mesh parameters and simulation setup information can be found in 'Appendix K' in [3].

The contour plots for the C_{DN} and C_{DP} values are shown in Figure 6 and 7.

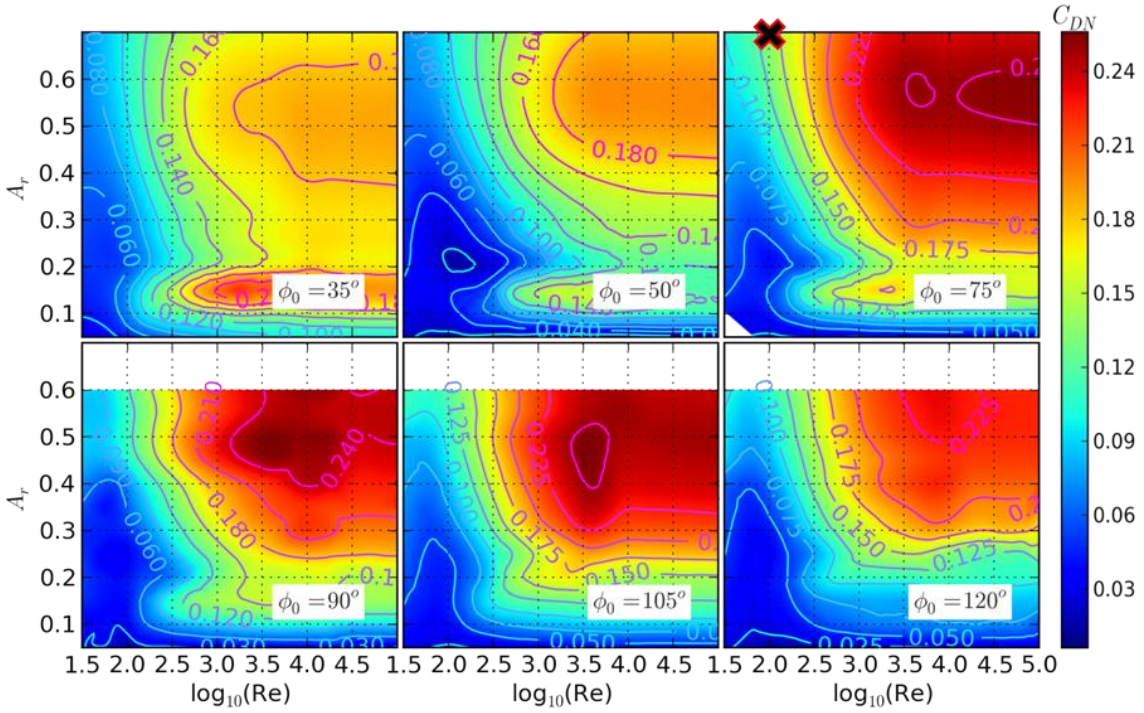


Figure 6: C_{DN} plot showing the force drag coefficient values as a function of A_r, Re and ϕ. Higher values of C_{DN} correspond to higher net force generation (for equal A_r, Re and ϕ).

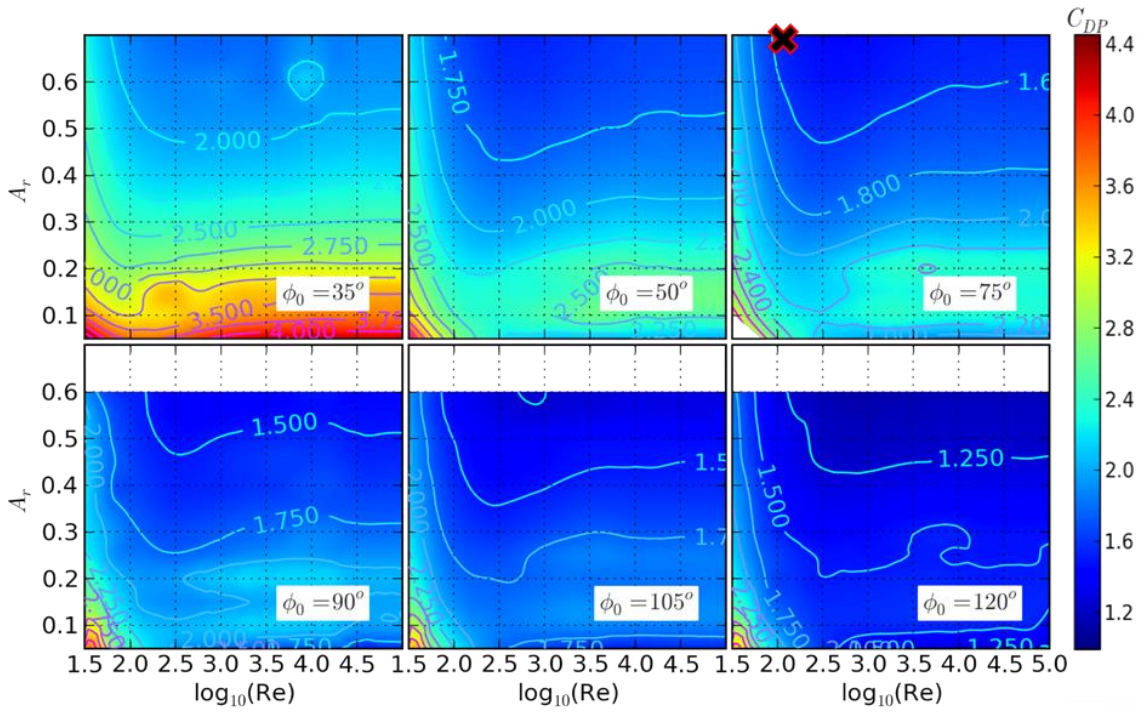


Figure 7: C_{DP} plot showing the power drag coefficient values as a function of A_r, Re, and ϕ. Higher values of C_{DP} correspond to higher cycle power used (for equal A_r, Re and ϕ).

The following is an example on how to use these plots:

Let us imagine that we can build a wing with the characteristics shown in Table 2 and that we can actuate it with an electrical potential high enough such that it resonates at an amplitude A_r=0.7 (this scenario is exactly the case for Figure 4).

Table 2: Example characteristics of a fabricated wing.

Characteristic	**Value**
A_sR	8 (must be this to use C_{DN} and C_{DP} plots)
ϕ	75°
f	340Hz
L	2mm
B	0.5mm

Using equations (1.9) and (1.10), the corresponding Re for this wing is calculated to be 100. At ϕ=75°, A_r=0.7 and $\log_{10}(Re)$=2, the corresponding C_{DN} and C_{DP} values would be about 0.138 and 0.16, respectively. These points are marked in both Figure 6 and 7 as well. Applying these to equations (1.7) and (1.8) results in an expected average force of 0.367μN (for 2 wings per flapping cycle) and an average power of 0.53μW.

1.1.3 Issues

Although a lot of groundwork has been presented for this wing device, there is room for improvement in the fabrication method/stack structure materials, and in the general understanding of the fluid-dynamics around such wings.

Specifically, the drag coefficient values only exist as 'look-up' tables, and little is understood about why the trends seen in Figure 6 and 7 exist, and what would happen if a different wing shape was used. Also, the stack structure presented in Table 1 was easy to fabricate and resulted in very high curvatures during Minnick's testing, but the use of SU-8 as the tensile material was impractical because its material properties would respond strongly to changes in temperature, pressure and humidity. As a result, the curvature of the wings would change daily, and the curvature would be permanently removed when exposed to a vacuum due to outgassing. These issues and gaps in knowledge need to be addressed to make these wings a practical reality for MAVs.

2 THE WAVE-WING INTERACTION

When examining the transient forces on a set of wings for any given A_r, ϕ, Re and A_sR, it can be seen that the forces *always* lag the wing position. The amount of lag is also dependent on some of these parameters at certain extremes. It seems that the wings drag fluid along during the flapping motion and later 'push' against it, resulting in a lagged force generation. This chapter thoroughly explains this phenomenon by investigating and modelling the interaction between a flapping wing and the air 'wave' surrounding it. We will see that this wing-wave effect dominates the force time dependence in flapping wings and is expected to be a general effect for flat shapes following oscillatory motion.

The Axial Components of the Force

In all the simulations performed in this thesis, the frame of reference is such that the x direction is chord-wise (i.e., in direction of the width of the rectangularly shaped wings) and the z direction is upward and positive in the direction of net force. The origin of all three lies at the center of the two-wing or one-plate system. This is illustrated in the following Figure 8.

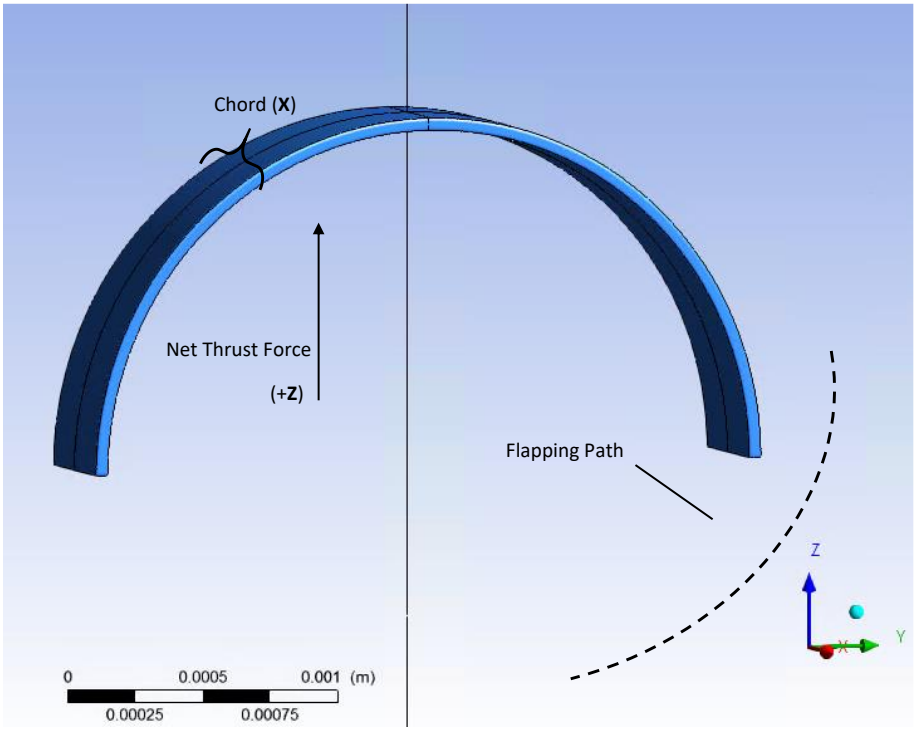


Figure 8: Visual representation of the simulated wing and components, taken in the reference frame used throughout the research. To preserve simulation time, only one quarter of the above wing system is simulated, with symmetry planes normal to x and y directions.

When considering the net force generated for a wing pair, force components in the x & y directions are zero by symmetry; thus, the net force is the force in the z-direction (multiplied by 4 if only considering the force of a half of one wing, which is what was typically simulated for computational efficiencies). Similarly, when considering only one wing (as in the left half in the above figure) the x-component of the force is zero, but the y-component remains, meaning the total force is a vector sum of the differential y and z component of the force vector over the surface of the wing. This one-wing investigation (still simulating wing-pairs flapping but looking at just one wing) presents more information about the transient force distribution and hence, it is of primary concern for this chapter.

It is important to note that the y and z force components will naturally be different for different wing shapes at an instant, even if the magnitude of the vector sum is equal; thus, to group the y and z force components and compare flapping shapes with naturally different positions, I use the

magnitude of the vector sum of the forces for the following figures. I will refer to this as the transient force magnitude (TFM).

Wing Simulation Observations

For the remainder of this chapter, wing simulation results with a ϕ of 90^o are investigated because this parameter demonstrates a good ratio of net force produced to power required (for a rectangular wing of A_sR=8), and therefore is an interesting shape to investigate [3] . It is also a good space to begin wing design optimization through variation of other parameters like the A_sR which will be done in Chapter 4. Although most of the work focuses on this area, we will see that other ϕ follow similar trends.

The following are plots of the normalized TFM of the wing force throughout a flapping cycle, for a ϕ of 90^o. The plots were normalized by dividing all values by the maximum found in the respective data to bring attention to their relative phase and the wing position (the actual forces on the wings of Re=$10^{3.5}$ are 3 orders of magnitude higher than those of Re=10^2).

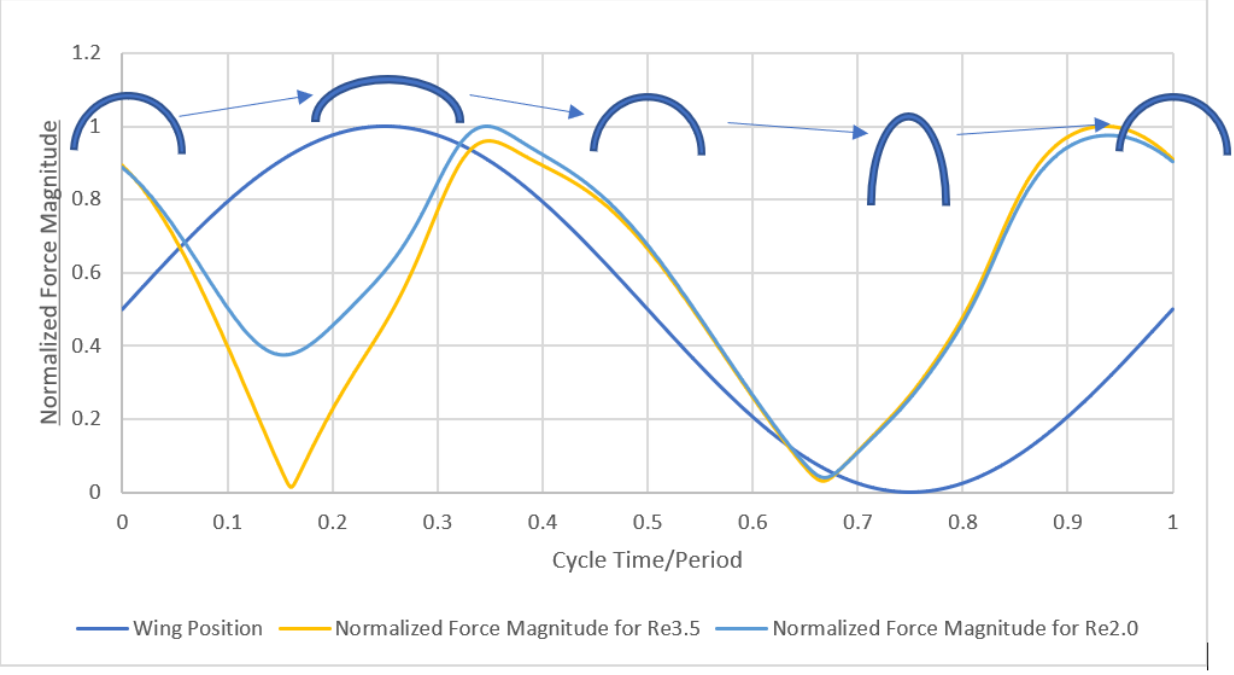


Figure 9: Normalized TFM for Reynolds numbers $10^{3.5}$ and $10^{2.0}$ for the yellow and blue curves, respectively, with an A_sR of 8. Included is a wing visualization to show the wing shape corresponding to the wing position.

From the above plot we can see that both normalized TFMs are in-phase with each other, with quite consistent overlap in the latter ¾ of the cycle, and slight deviation before that. Both transient curves are out of phase from the wing position by about 10% of a cycle. For the convenience of the reader, a visual representation has been added to this plot to clarify the meaning of 'wing position'; this convention will be used for the rest of the thesis.

The following are two more plots examining $Re=10^{3.5}$ vs. $Re=10^{2}$ for wings with identical ϕ and A_r but different A_sR: the first has $A_sR = 1$ and the second has $A_sR=0.25$.

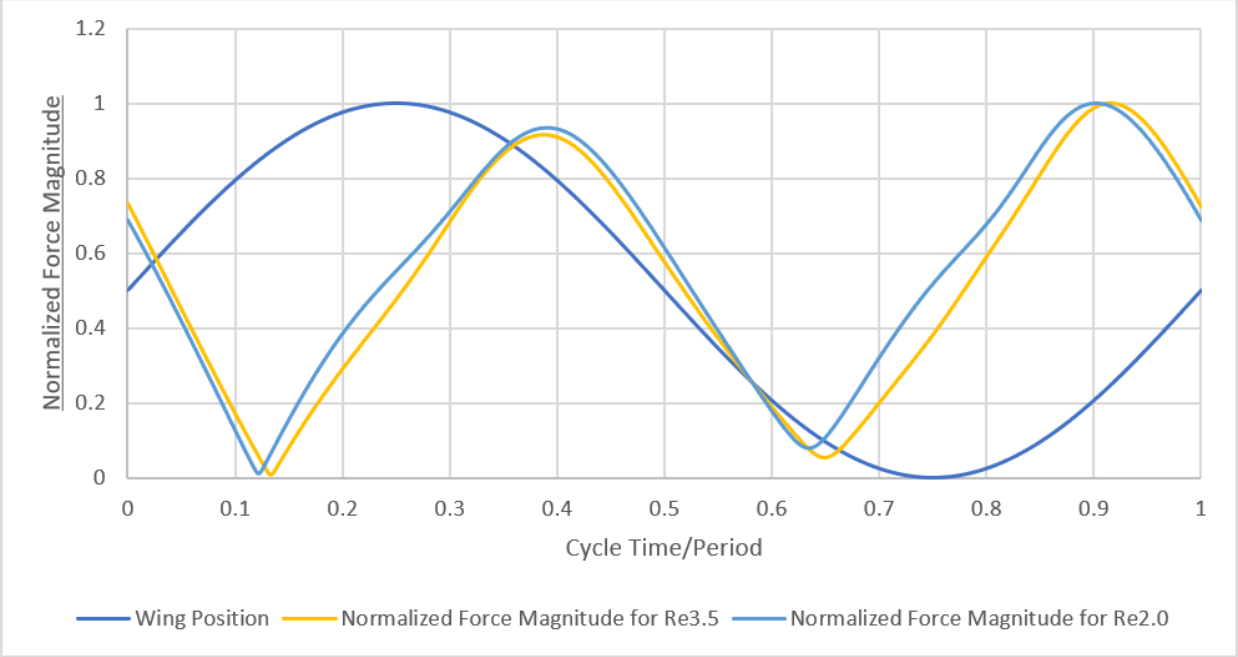


Figure 10: Normalized TFM for wings of $A_sR=1$.

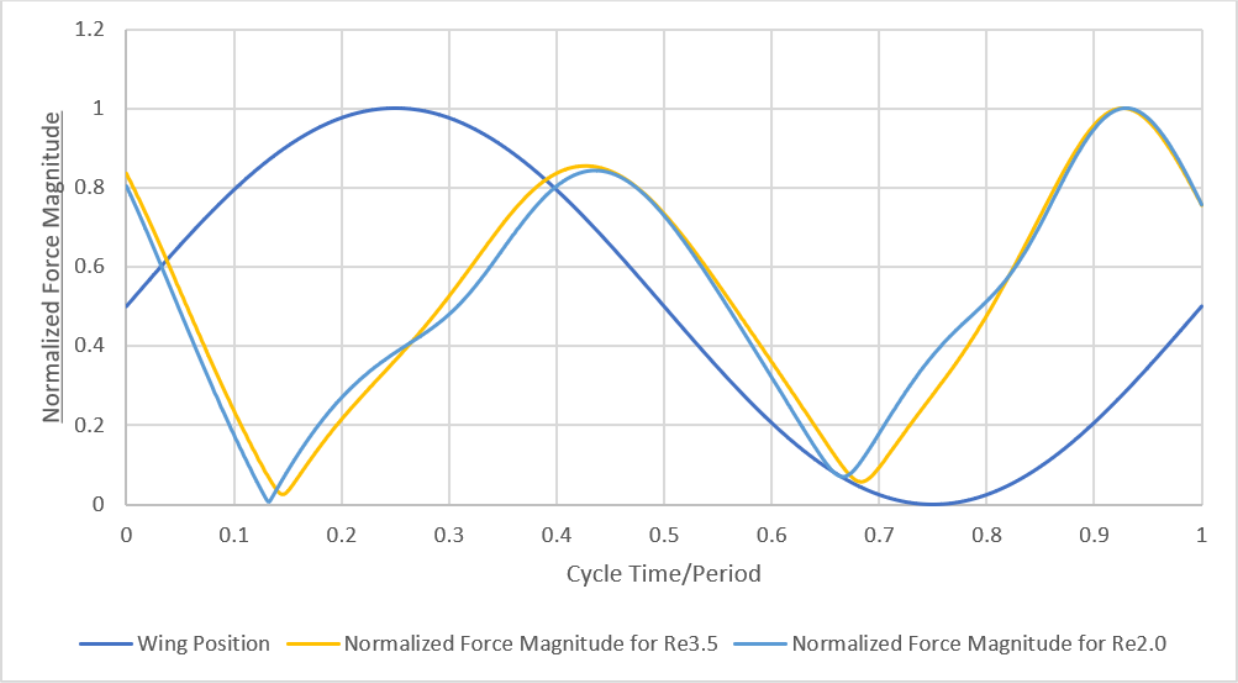


Figure 11: Normalized TFM forces for wings of A_sR=0.25.

Without concerning ourselves with the changes in the shapes and the relative differences in the TFMs (discussed thoroughly in the next chapter), we notice again that the force generated on the wings always lags behind the wing position by about 15-20% of a cycle (making it lead the wing speed by about 5-10% of a cycle). As mentioned earlier, this is true for all cases and might be unexpected whether one considers that an object accelerating a mass should have its acceleration in phase with the force, or that the point of highest drag force should be a quarter cycle out of phase (making it in phase with velocity). The former case is related to what many authors consider as the virtual mass when studying objects undergoing small oscillations, defined as a mass of air attached to the body which always moves in phase with the object [15], [17], [21], [22]. This virtual mass seems to always be included in calculations regarding fluid-cantilever interactions, yet unsteady fluidic effects are often ignored. The most related fluidic effect that I have found is the wake-capture effect proposed in insects where vortices generated in previous half cycles can help to generate lift during the reverse stroke [25], but this is for the complex wing movement case. At any rate, the phase lag effect is very noticeable and is a general occurrence in the simulations; thus, we begin the breakdown of the thrust generation on the wing with an explanation of this phenomenon.

2.1.1 Position-Force Phase Lag

The reason that the force on the wing lags its position is likely from the fact that the wing does not simply push a mass, but it oscillates and interacts with earlier 'pushed masses'. The mean mass flow is the wave created by the wing that leads it. This wave can then perhaps be considered as a sheet or a point mass (at least to explain the interference with the wing). To visualize this in the simulations, the mass-flow weighted average position was found at an instant in time around the wing. This average position was taken along several straight lines of varying z_i coordinates along the yz-plane that intersects the wing, and calculated as follows:

$$\bar{y} = \frac{\sum |v(z_i, y)| * y}{\sum |v(z_i, y)|} \tag{2.1}$$

Where $v(z_i, y)$ is mass flow. The result is shown in the following figure for an instance in the flapping cycle, with $\bar{y}(z_i)$ represented with black dots. Note that simulations have shown that the position of these mass-weighted averages exists away from the wing when it is at 0 velocity, and so this does not simply represent the virtual mass of fluid attached to the wings (which would be 0 and overlaid with the wing).

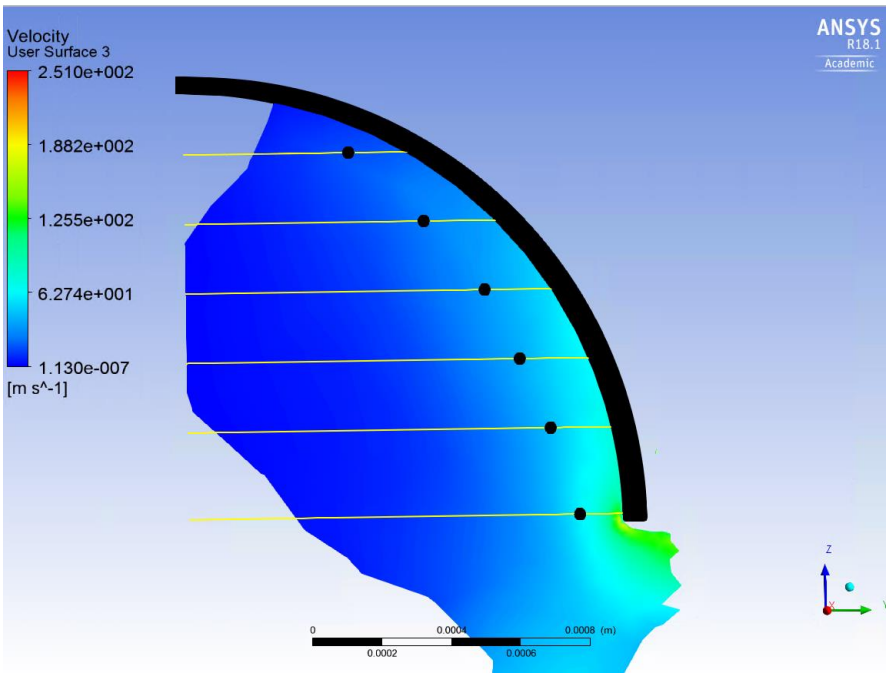


Figure 12: Mass-Flow weighted position mean that demonstrates the wave lag behind the wing. Note that here the wing is in motion opening upward. The black dots represent the mean positions, and when connected, give an idea of the wave sheet that follows the wing at this moment in the cycle.

This same mass-sheet then interacts with the fluid between it and the wing, effectively feeling drag (and adding to the drag force on the wing). Despite this approximate explanation, we can nevertheless attempt to model the system as a mass-damper interaction, with a moving wall representing the wing, as shown in the following figure.

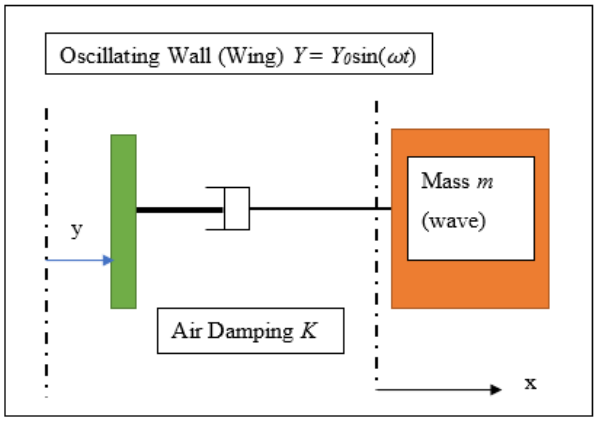


Figure 13: Equivalent mechanical system approximation of the wing-fluid interaction. Note that the dashed lines represent the equilibrium position for the respective functions.

In the above model, the damper accounts for the added drag force felt by both the wave and the wing, based on the relative velocity of the two.

For simplicity, we will assume laminar flow (linear drag force dependence on velocity) and so the differential equation for this system is

$$m\ddot{x} = -L(\dot{x} - \dot{y}) \tag{2.2}$$

At high *t*, this model has a steady sinusoidal solution for the motion of the mass with the added phase offset θ as shown below

$$\begin{aligned} x(t) &= \frac{Ky_0\left(-m\omega\cos\left(t\omega\right) + K\sin\left(t\omega\right)\right)}{K^2 + m^2\omega^2} \\ &= a\sin\left(t\omega + \theta\right) \end{aligned} \tag{2.3}$$

where

$$a = Ly_0\sqrt{(m\omega)^2 + K^2}$$

$$\theta = \tan^{-1}\left(\frac{-m\omega}{K}\right)$$

The following is an example plot of the effect which shows the plate position and the force.

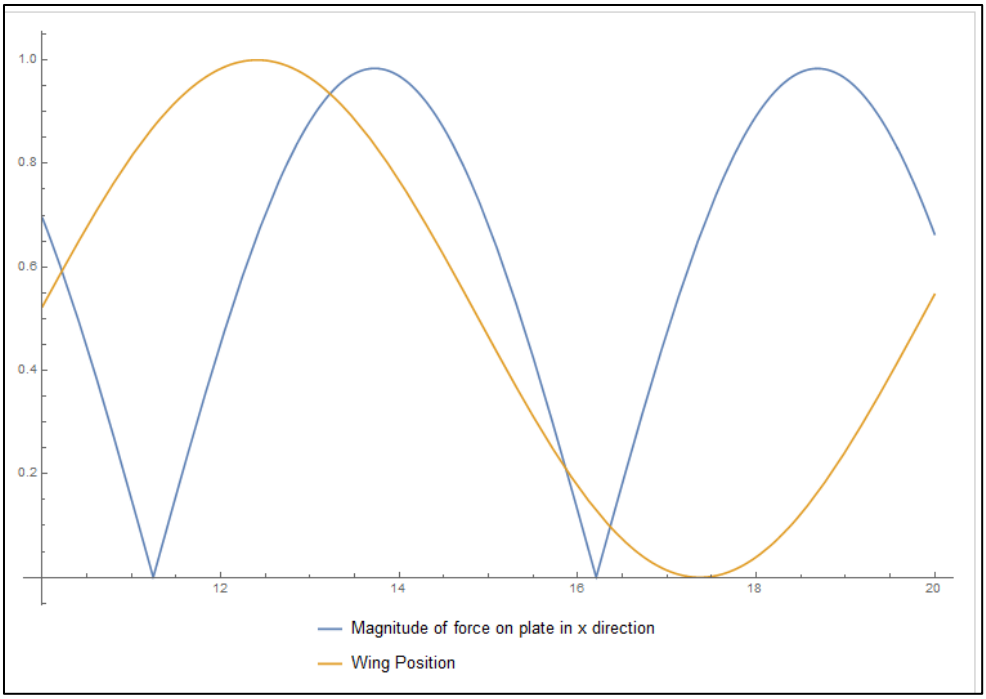


Figure 14: Plot of analytical solution to the mass-damper model.

Although this model was based off a simple approximation, we can see from Figure 14 that it does match the simulated results (Figure 10 and 11) well, once correct L, ω and m values are chosen. This suggests that there is a phase lag dependence on the amount of fluid drag, the 'mass' of the wave, and the frequency of the plate oscillation.

2.1.2 Frequency Dependence

We will begin with investigating the frequency of the analytical model as it is the most obvious parameter that relates to the actual simulated wing, while the mass and damping are still loosely-based concepts. The Reynolds number in the C_{DN} and C_{DP} plots (Figure 6 and Figure 7) is frequency dependent; thus, it makes sense to try and get a sense of the effect of frequency in the simulation, and how well it matches this model. In an attempt to remove other force-contributing factors such as the flapping motion and initial curvature, we simplify the case to that of a square plate, sinusoidally oscillating in air. This can be thought of as an approximation of a wing with $\phi=0$, and $A_sR=2$. The transient force simulation results are shown in the following figure, for varying frequencies (and thus, Re).

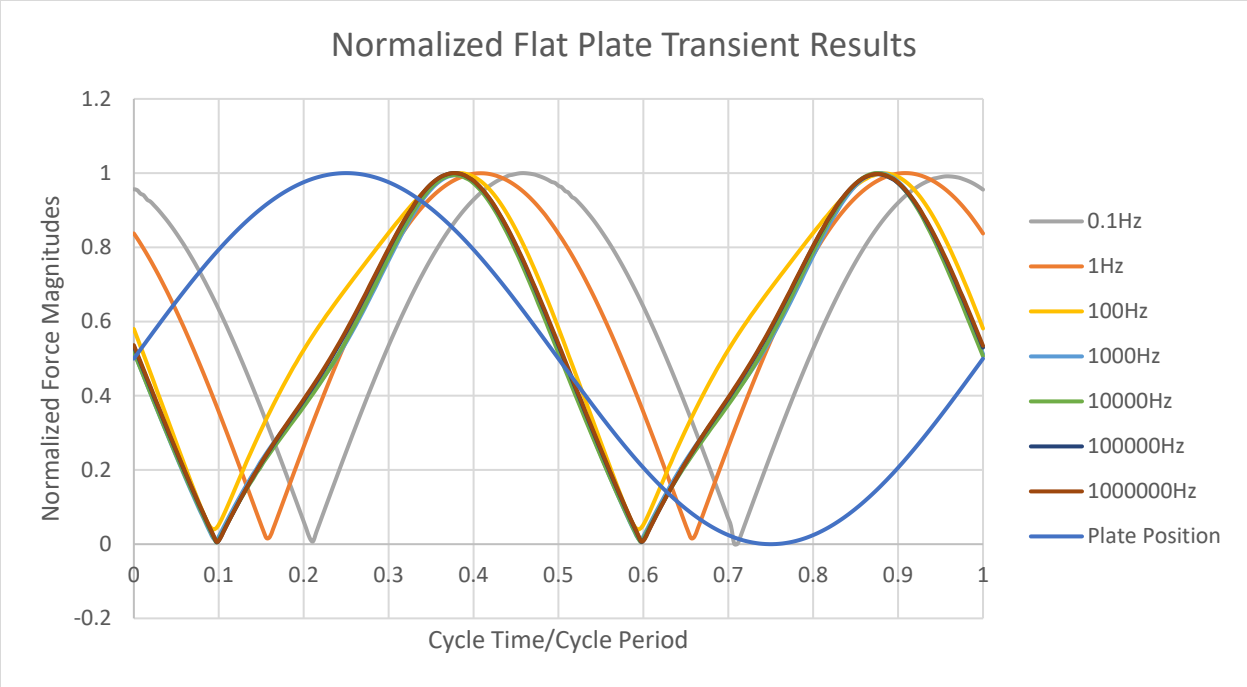


Figure 15: Normalized magnitudes of transient forces flat plate oscillations at varying frequencies. Note that the amplitude for oscillations here is 1.0mm, or half the wing side length.

Here we see the results of the flat plate motion at different frequencies, with Reynolds numbers calculated ranging from 0.13 to 130 000 corresponding to frequencies of 0.1Hz to 1 000 000Hz, respectively. Below 100Hz, there is a clear phase shift toward the right, with maximum force occurring close to the beginning and middle of the cycle. At this point, the force is in phase with the velocity of the plate. This is consistent with the concept of Stokes' drag where at low Reynolds numbers, viscous forces dominate over inertial forces, and the drag force is in phase and linearly proportionate to the relative velocity of the system as shown by the following equation:

$$F_d = S\mu Bv \tag{2.4}$$

Where C is a drag constant, μ is the dynamic viscosity of the fluid, B is the characteristic length of the object and v is the relative fluid velocity. Although (2.4) is derived for steady-state systems, this is a good assumption at the limit of low oscillation frequencies. In fact, this is related to the laminar flow assumption that gave the second term in (2.2), although it must be stated that the model is incomplete since the mass-damper system becomes in-phase with

position at low frequencies. At this point of low Re, the model changes to a simple wall-damper-plate system as in the following figure:

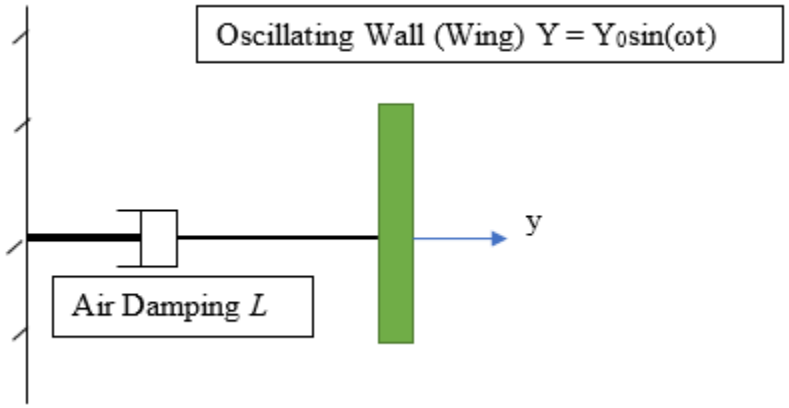


Figure 16: Adjusted model for the case of low frequencies/Reynolds numbers.

In any case, the most relevant Reynolds numbers are those between 13 and 130000 (i.e., 100Hz-1 000 000Hz) that span the Re parameter used in the C_{DN} and C_{DP} plots. We can see from Figure 15 that there is little difference in the plot shapes except for cases below 100Hz where the corresponding Re is less than 13. Here, due to the lower Re, the plate has more viscous forces present which have a greater effect on the overall TFM plot shape than those of higher Re. That being said, to a first order approximation, the plot shapes do not deviate much from one another, and we will see in Chapter 3 that these deviations are negligible when considering the main reasons behind net force generation.

We can also note the simple force distribution about the flat plate simulation during the cycle. An example surface plot of where force is generated maximum velocity is shown in the following figure:

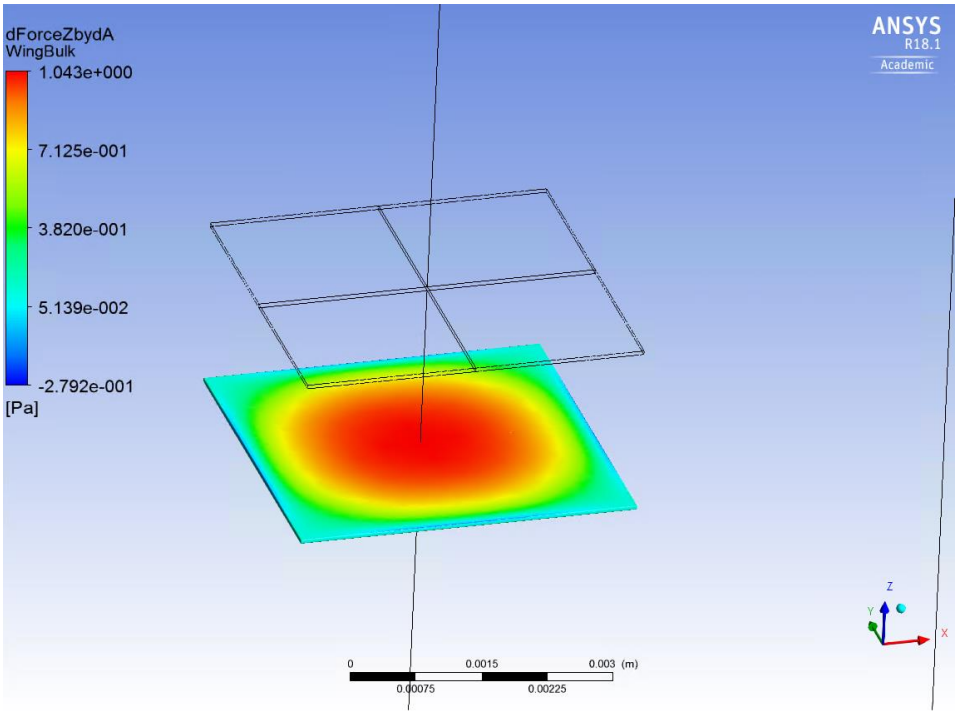


Figure 17: Surface pressure plot for half a cycle on a face of the oscillating flat plate.

We can see that the majority of the force differential is centralized on the plate due to symmetry and the force decreases near the edges due to pressure escaping on the sides. Note how the corners provide two sides through which pressure can equalize, and so we see low force per area in that vicinity (an important note that I will return to when I discuss A_sR effects in chapter 4).

At this point, it is important to make three observations:

1. The results from the flat plate transient simulation show very little phase-frequency dependence for the Reynolds numbers of interest (Re>100). This is also true for the transient results for a chosen wing A_sR and ϕ (90^o) as shown in Figure 9 - Figure 11.
2. The results of the flat plate simulations are again out of phase with the wing/plate position which is consistent with the model and wing transient force data.
3. The force plot shapes of the flat plate simulations are very similar to those of a flapping wing (comparing Figure 15 to Figures 10-11).

Point 3 is particularly interesting as it suggests that the transient force magnitudes are not very dependent on how the wing flaps. This result is peculiar considering that each part of the resonating wing at a distance from the center always moves following a different path and

velocity, while for an oscillating plate, all points follow the same movement. Most likely, this relationship has to do with the fact that the majority of the force is generated near the parts of the wings that move at highest velocity: the area near the wing tips. This fast moving area more closely parallels a flat plate motion.

Considering that the simple oscillating square plate produces qualitatively similar force plots to that of the complex wing resonant flapping shape presented in "Background Information", we can continue to use this model to decouple the complexities of the force generation of the curved, resonating cantilever wing.

2.1.3 Oscillation Amplitude Dependence

The second, major variable in the C_{DN} and C_{DP} plots is the amplitude of the wing 'flap', A_r, defined as the ratio of the path swept by the wing tip to the wing length. I ran similar plate simulations for Reynolds numbers of interest of varying A_r, which in this simple case is just the amplitude of the motion. Note that the results of the previous subsection use an A_r of 1, and at around *f=100 000* Hz, the transient force results consistently overlapped (i.e., there was low frequency dependence). The following are the force results of different A_r at 100 000Hz:

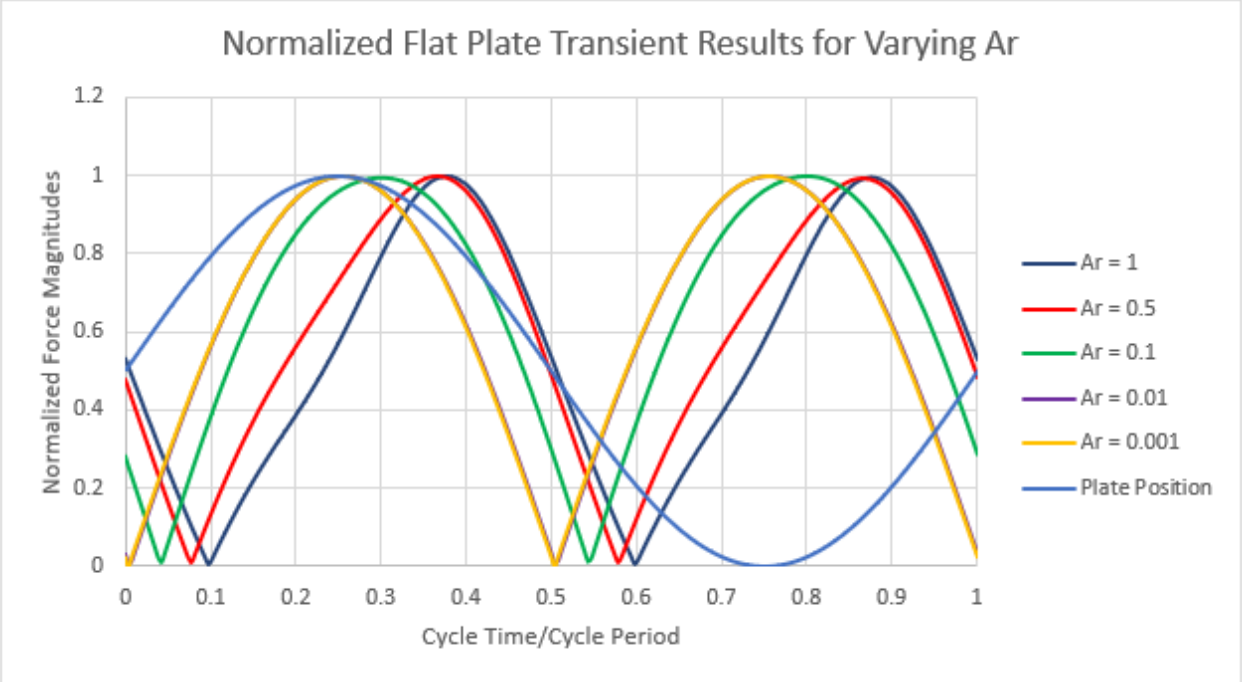


Figure 18: Normalized magnitudes of transient forces flat plate oscillations at varying amplitudes. The oscillation frequency here is 100 000Hz.

Notice how the phase shift for smaller oscillations tends toward the wing position curve where it matches the function shape and is almost exactly in-phase at amplitudes less than 1% of the wing size (Ar=0.01 and A_r=0.001 which overlap in the above plot). In terms of Re, at constant frequency, the phase shifts to the left with decreasing Re.

This is what is referred to as the “virtual mass” effect mentioned previously, where for small oscillations, the air near a cantilever will be displaced by the moving cantilever and can be considered as an inertial force term, proportional to the cantilever acceleration only. Continuing with the mass-damper analogy, this would be equivalent to the solution in the limit of high air damping that brings the wing and mass in-phase.

Notable Conclusions

In this chapter, the wing-wave interaction was thoroughly inspected with regards to varying frequency and flapping amplitude, by simulating the simple case of an oscillating plate. The interaction between the wing and the moving wave was explained by analogy to a mass-damper

system. For further investigation into the nature of the force generation of these 1-dimensional flapping wings, it is important to note the following conclusions:

1. There is little frequency dependence for the force lag for wings above Re=100.
2. Flat plate simulation results show that the transient drag force is very similar to that of simulated wings (when normalized), which suggests that the transient force magnitudes of the wing is not dependent on the flapping shape of a wing (to a first-order approximation). That is, a wing can be curling or oscillating up and down yet similar TFM plots will be seen.
3. The force phase of the simulated plate becomes more in phase with lower oscillation amplitude (shifting the force plots to the left).

In the next chapter we will see how these results, along with effects from varying wing flapping parameters can explain such different net thrusts and C_{DN} values for varying parameters Re, A_r and ϕ.

3 FORCE GENERATING EFFECTS

In the previous chapter we looked at the Wing-Wave interaction: How the wing drags a wave of air along with it leading to phase-lag for the air's force on the wing. In those results, the entire force magnitude was used (rather than just force in the thrusting direction, z) for comparison to simpler cases/models and to note relationships between parameters; this let us notice bulk "total-force" effects that occur on the wing. In this chapter, we will build upon the observations and explanations of Chapter 2 to explain how a net-thrust force (on average over a flapping cycle) is generated on the wing as it interacts with the transient wave created in its wake. We will see that the TFM is mostly the same for different ϕ with all other things equal, and that actual force and efficiency is related to the wing positions during the cycle.

Results of Investigation of the Y & Z Force Components

To examine different possible dependencies of the net force over a cycle, we begin by selecting points of interests based on the earlier C_{DN} plots. Looking back at that plot in Figure 6, there appear to be a few major phenomena occurring:

1. In general, above A_r=0.2, the C_{DN} tends to increase with increasing ϕ, up to a peak at ϕ=90°, around where Minnick in [3] found that generally efficient and practical wings exist (given an A_sR=8).
2. For a given ϕ, the C_{DN} generally increases with increasing A_r and increasing Re.

3. An exception to phenomena 2 is an anomaly that occurs at 0.1<Ar <0.2 and Re>$10^{2.0}$, where high C_{DN} is found around Re=$10^{3.5}$. This manifests strongly at low ϕ and the C_{DN} in the area gradually reduces with increasing ϕ.

Based on these observations, 5 points (P1-P5) are chosen for investigation that fall into the extremes of the phenomena, as shown in Figure 19.

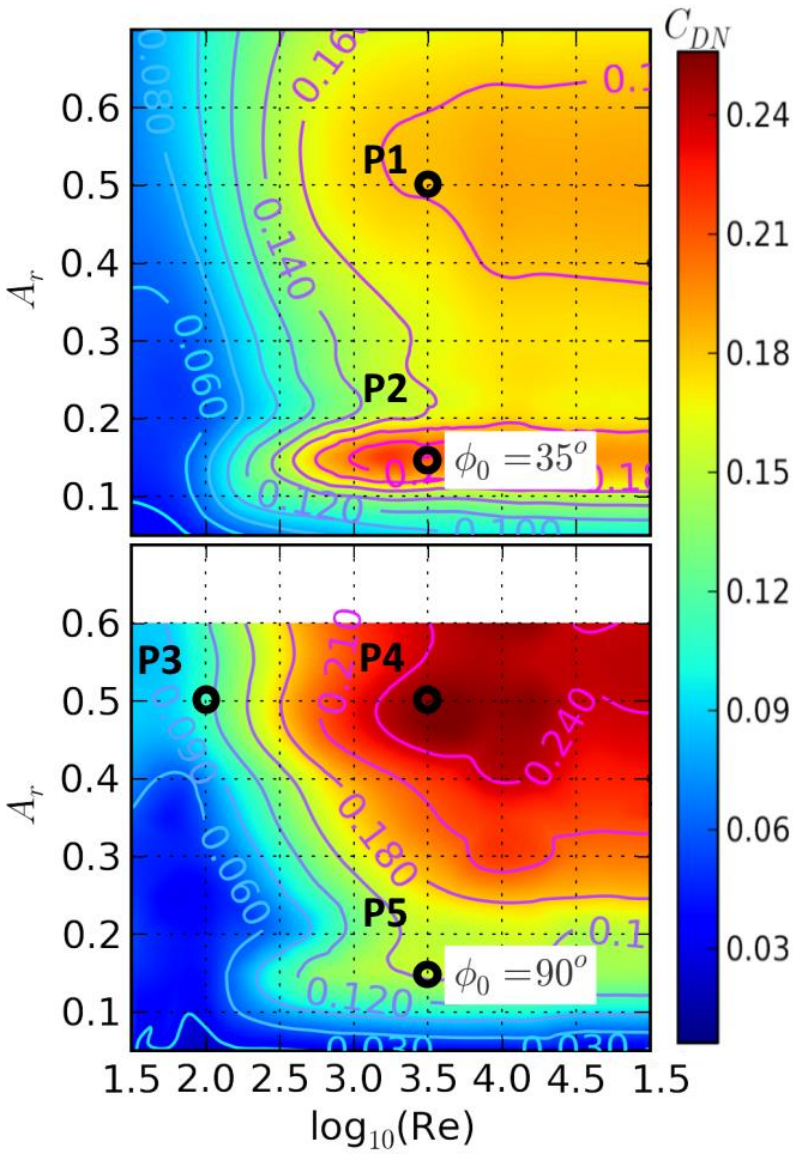


Figure 19: C_{DN} plots at ϕ=35 and 90°, with points P1-P5 representing points of interest to be examined.

The corresponding force plots for each point, along with the wing position are presented in Figure 20.

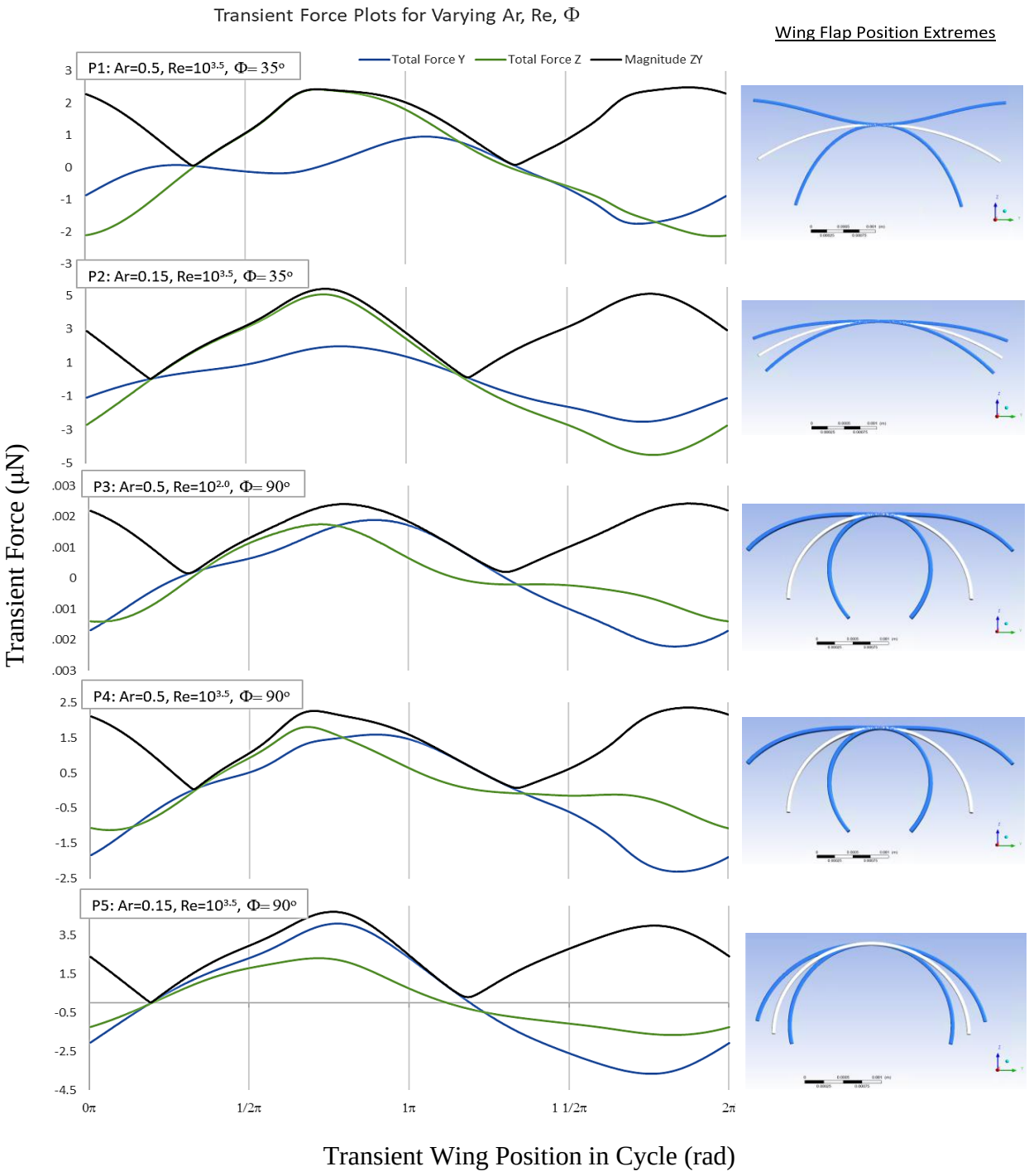


Figure 20: Stacked plots of transient forces on the wing (left) and the wing flap position extremes. The white wing outline represents the resting wing position for that ϕ, while the top and bottom blue shapes represent the positions at 1/2π and 3/2π of a cycle while flapping, respectively.

General Net Force Generation and Efficiency

Before we begin breaking down the effects seen in Figure 20, one concept must be pointed out regarding the generation of net z force of a wing or any oscillating object:

Over an oscillation cycle, there must exist an asymmetry between the negative and positive z force contributions, such that there is more positive force, in order to get a net positive thrust.

The above statement may seem obvious, but it sets the fundamental fact upon which all further explanations build. It follows that, other things equal, more symmetry in the z-force over a cycle means less net force is generated, while higher asymmetry in favor of the positive force will generally lead to higher efficiency in wings. Looking at the magnitude plots of Figure 20, we see that they all appear to roughly follow the function

$$F(t) \approx A\left\|\sin \omega t + \theta\right\| \tag{3.1}$$

where $\sin \omega t$ is the normalized wing position. This is consistent with the flat plate oscillation and lagged wing-wave interaction described in earlier. The main differences are in the phases and distribution of the force magnitude between the y and z axis (shown in blue and green, respectively). Considering this, we can see that the absolute limit in efficiency (and maximum *net* force for a set of parameters), is when all the force magnitude during the first half-cycle (the downstroke) is projected onto the wing as it is most parallel to the y axis, while the other half cycle that generally contributes to negative force (the upstroke) is projected in the z-direction. This way, the resultant force of the upstroke has no negative contribution in the thrusting direction z. This limit is equal to the average of the above equation for half a cycle:

$$F_{net_{Max}} = \frac{1}{T}\int_0^T A\left\|\sin \omega t + \theta\right\| dt \tag{3.2}$$

which is equivalent to

$$F_{net_{Max}} = \frac{2}{T}\int_0^{\frac{T}{2}} A\left\|\sin \omega t + \theta\right\| dt = \frac{2A}{\pi} = \frac{2F_{\max}}{\pi} \tag{3.3}$$

Where F_{max} is the maximum instantaneous TFM. This theoretical max can be useful in measuring the efficiency of the wing, comparing to nature's fliers, or testing if certain parameters can even achieve the minimum force necessary for flight.

Note that it can be argued that the best scenario would be where all the force throughout the entire cycle translates to positive-z force. This has never been seen in the simulations and is in contradiction with the flat plate model which suggest that there will always be a counter force during one half of a cycle for such a simple flapping motion as shown in Chapter 2. That being said, one of the reasons that many insects and some birds have evolved to produce complicated, 3 degree-of-freedom wing flapping motions is to achieve this type of scenario where both halves of the flapping cycle result in positive z force and higher power efficiency (a common force plot for the flapping cycle of animals can be found in [8], [26]).

Key Observations and Explanations

In this section we will discuss the force results shown in Figure 20 in greater detail to understand not only the C_{DN} plots, but in general, the factors affecting the wing force on a first order scale. We can later extrapolate this information to make better parametric decisions in the creation of more powerful, efficient wings. As done earlier, first the TFM will be discussed based on phase similarities, function shapes and amplitude differences. Later, we will see how this magnitude is projected differently onto the y and z axis of the wing, which dictates the actual net thrust on the wing.

3.1.1 Phase Similarities

Looking at the points of minimum TFM in Figure 21, we see that certain plots have the same phase (as shown by the dashed lines). Note that I have rearranged the order of plots to make this easier to see; now we compare each static bending angle at $Re=10^{3.5}$, first at $A_r=0.5$ then $A_r=0.15$.

The similar phases support the idea that the TFM is independent of the ϕ, and strongly dependent on the A_r, as shown by the flat plate oscillation results. We can also note the clear negative phase shift for $A_r=0.15$ when compared to $A_r=0.5$. This too is consistent with the flat plate simplifications in Chapter 2 where inertial mass effects became more prominent at smaller oscillation amplitudes for a fast moving wing (high Re). Note that although the Re is equal in all four cases, the simulations done at $A_r=0.15$ oscillate at roughly 0.5/0.15=3.33 times the

frequency of those at A_r=0.5 which is a result of the V_{rms} definition (equation (1.10)); however, as in the flat plate simulations, this small frequency change has no noticeable effect on the phase.

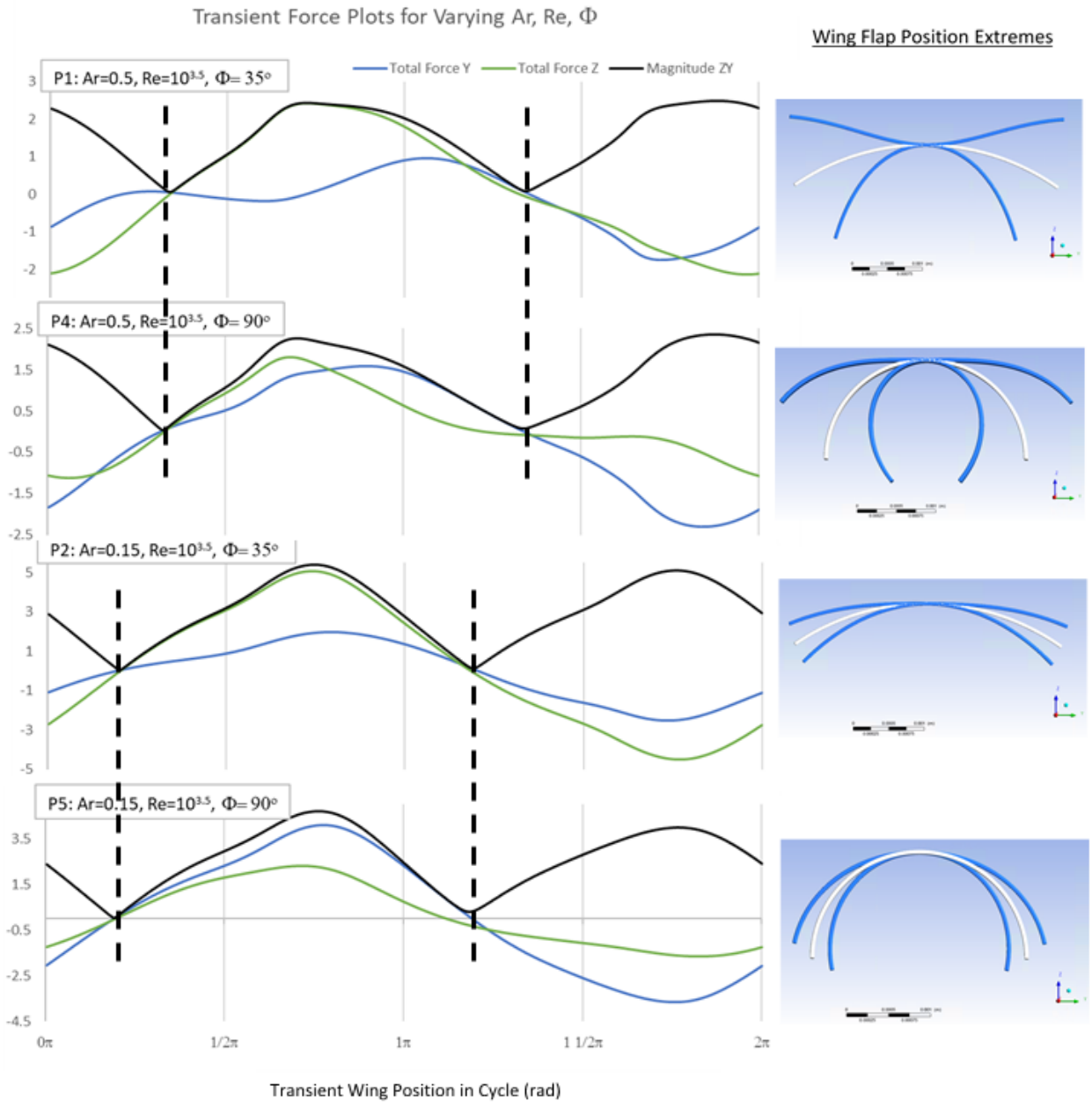


Figure 21: Transient force plots and corresponding wing positions for P1, P2, P4 and P5.

3.1.2 Function Shapes

Another curious result is in the similarities in certain TFM shapes. Although earlier I claimed that the magnitude can be roughly described by that of a sinusoidal function, we can see that even certain supposed second order effects remain constant between simulations. In the above figure, we see that the same pairs of wings with equal A_r and Re (P1&P4 and P2&P5) have a similar TFM dependence on the cycle time, even though the flapping shapes vary so drastically. Consequently, we may approximate that

$$F(t,\varphi_1) = cF(t,\varphi_2) \tag{3.4}$$

for equal Re and A_r, where C is a constant.

It is likely that changes in the function shape are due to complex wing interactions with the surrounding air and the induced wave. It has been seen that changes in A_r result in different magnitude vortices occurring near the edges of the wings. Vortices occur from large gradients in local pressures and mass flows, and such large gradients in this situation occur near the extremes of the wing motion, where the induced wave continues to flow past the wing boundary, while the flow near the surface follows the wing motion (in the opposite sense). An example is shown in the following figure, in which the wing is in the downward motion after the first extrema. The vortex phenomenon shown is more dependent on the transient wing velocities near the boundaries, which are constant for wing motions at constant A_r and ω (and hence, Re).

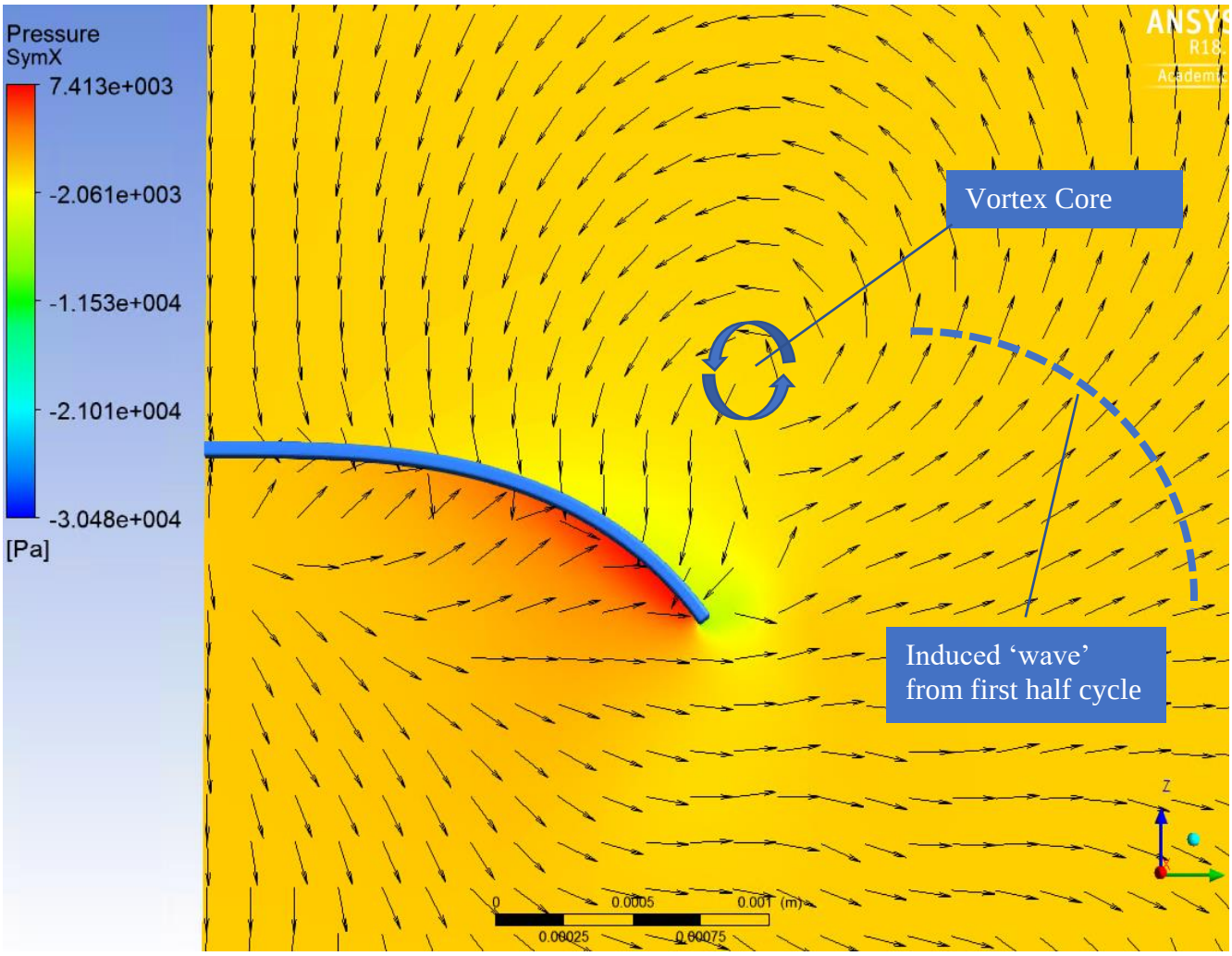


Figure 22: Example of vortex shedding of a wing just after the first quarter cycle for Phi0=90, A_r=0.5, and Re =$10^{3.5}$. The arrows show the normalized local velocity vector, and the colours denote the pressure. Here, the wing is beginning its downward motion 'dragging' air behind, yet the induced lagged wave (seen by the velocity vectors on the top right half of the plane) continues upward. The result is a gradient in mass flow and a vortex above the wing.

3.1.3 Amplitude Differences

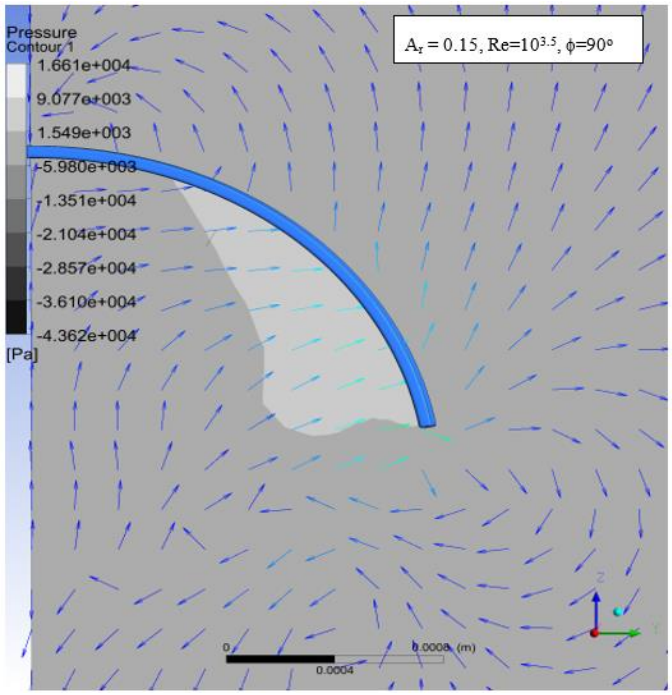
We have seen how transient force magnitudes at the same A_r behave similarly during a cycle, yet there is a noticeable amplitude difference. Using (3.4) and comparing P1 and P4 of Figure 21 such that the RMS of the residuals between the plots is lowest, we find that

$$F(t,\varphi_1 = 35^o, A_r = 0.5) = 1.07F(t,\varphi_2 = 90^o, A_r = 0.5) \tag{3.5}$$

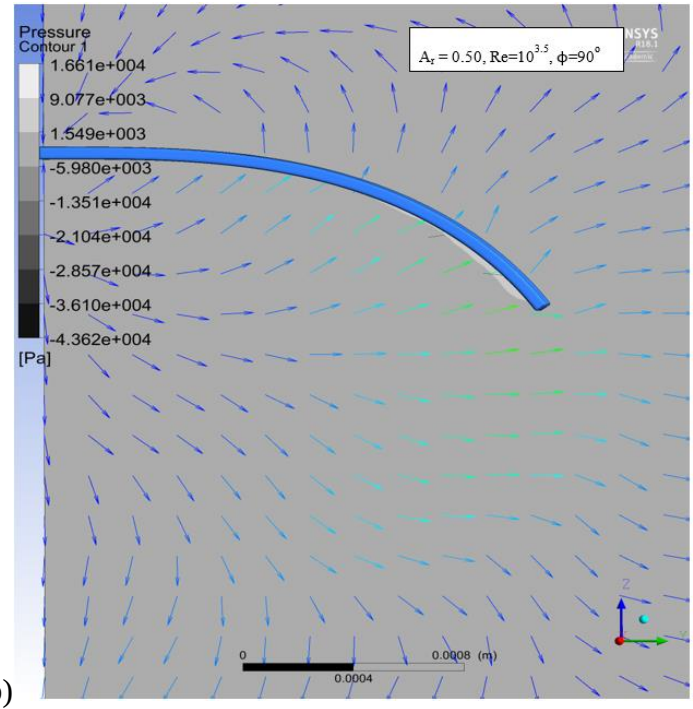
while P2 and P5 have the relation

$$F(t, \varphi_1 = 35^o, A_r = 0.15) = 1.15F(t, \varphi_2 = 90^o, A_r = 0.15) \quad (3.6)$$

It is also true that in general, the TFM of both wings with A_r=0.15 is higher than the counterparts at A_r=0.5 (at constant Re). This is a consequence of the inertia of part of the 'wave' flowing behind the wing. When the wings move at small amplitudes, a larger portion of the built-up mass flow (i.e., the lagged wave) is encountered on the reverse stroke, since the wing flaps between more parallel extremes (i.e., low A_r flapping wings will be pointing in almost the same direction at the end of the downstroke as they are at the end of the upstroke). At larger flapping amplitudes, the wing curves largely while the mass flow tends to follow a straight path; thus, the flow escapes the vicinity of the wing, contributes less to the pressure build up on the reverse cycle, and ultimately produces less instantaneous force. We can see this clearly in Figure 23 where vector plots of the mass flow around the wing are presented for both the A_r=0.5 and A_r=0.15 cases (near the extremes); these plots show the same greyscale of pressure, so the more extreme shading on the A_r=0.15 case demonstrates the increased magnitude of this wing-wave interaction, while the direction of the green (i.e., high) velocity vectors gives an impression of the source of this pressure difference.



a)



b)

Figure 23: Pressure contour and mass flow vector plot taken on the symmetry plane normal to the x axis near the end of the first quarter cycle for a) A_r=0.15, ϕ=90° and Re=$10^{3.5}$ and b) A_r=0.50, ϕ=90° and Re=$10^{3.5}$. The colour scale shows the magnitude of the vector, increasing from dark blue to green.

At A_r=0.15 (Figure 23a), we can see that most of the mass flow follows closely behind the wing at relatively normal directions to the surface. At A_r=0.50 (Figure 23b), a lot of the mass flow escapes past the wing in the +y direction, and fluid close to the wing follows a more parallel trajectory. This latter case leads to less normal mass flow near the wing during the wing-wave interaction, and thus less force (as seen by the higher-pressure contour at A_r=0.15). Note that between both cases, the V_{rms} is almost identical at all instances throughout the cycle.

A similar explanation can describe the amplitude differences shown by equations (3.5) and (3.6). A resonating wing with a higher initial curvature tends to bend away from its surface normal, more so than a straight, ϕ=0° wing – see Figure 24 for an exaggerated visual. This effect can be quite pronounced as seen in Figure 14 where the transient normal surface Vrms is shown. Here, at each point on the surface of the wing, the local wing velocity was projected onto the local surface normal - qualitatively, this is a representation of how the wings moves relative to its normal. The Vrms was then taken over the entire wing. If a wing is moving in a direction generally normal to its surface, then one can expect that, as previously stated, a stronger interaction between the wing and wave would take place. It turns out that resonating motions at smaller ϕ tend to follow more closely to the surface normal, for a given Ar. I must note that this explanation is harder to justify, as the differences in the TFMs at constant A_r are not very large, and other second order effects from local vortices and interferences might also play a factor. In any case, I consider this to be a second order effect and, as we shall see shorty, it is not necessary in describing the phenomena observed in the C_{DN} plots.

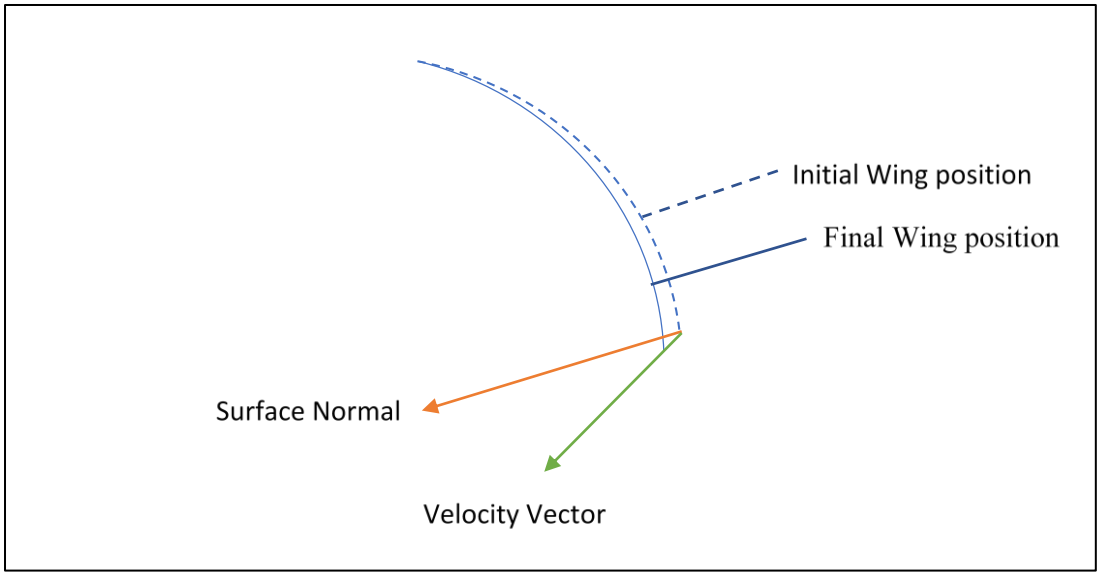


Figure 24: Exaggerated visual showing the differences in wing velocity of a part of the wing and the surface normal at that point.

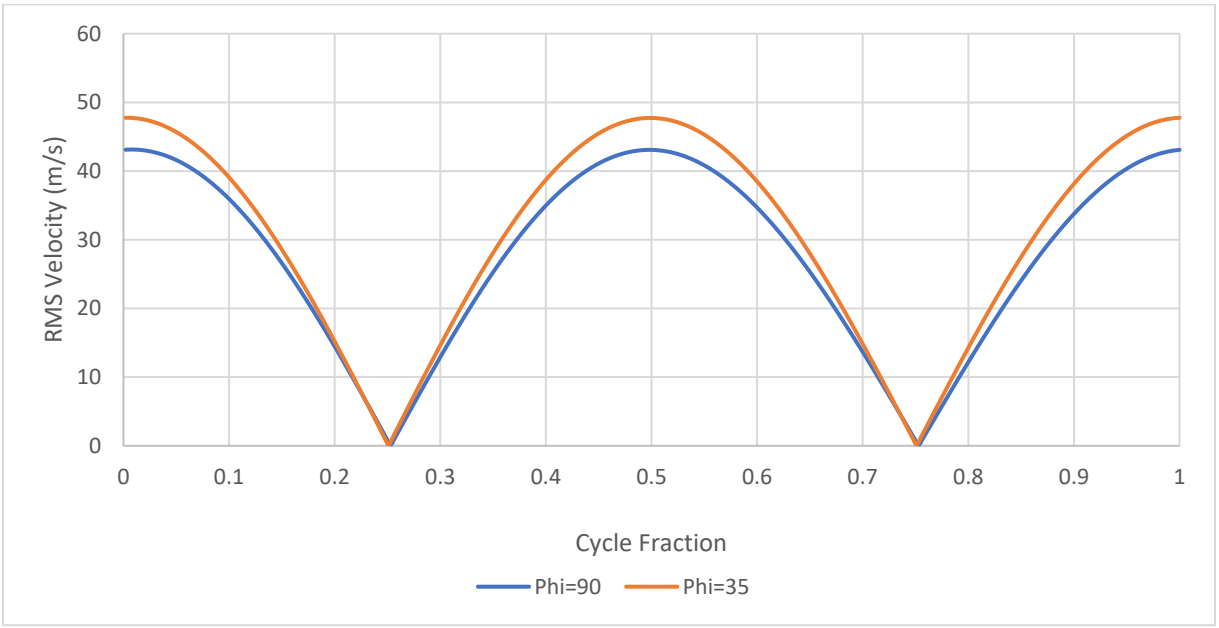


Figure 25: RMS over the wing surface of the normal-to-the-surface component of the wing's velocity throughout a flapping cycle.

3.1.4 Y and Z Force Projections

An important and unexpected observation that was implied in the approximation equations (3.1)-(3.3), is the remarkable symmetry about the minima in each plot of the TFM. It was suggested by

Minnick in [3] that the main reason for the net force generation in resonating curved cantilevers is that a closed wing has less drag than an open wing, as in steady state flows; thus, the closed half of the cycle would have less drag force than the open half (in opposite directions) and the net effect is positive thrust. However, we see very strong similarities in the TFM generated for both halves of the flapping cycle in each case, even at the high C_{DN} values of P2 and P5. In fact, in the case of P5, if we integrate the force magnitude between the dashed lines of Figure 21, the second half (corresponding to the negative z force) is about 2% *larger* than that of the first. This begs the question: how do the wings really generate a net thrust?

When considering a flat plate oscillating like in the previous chapter, the net force over a cycle was always 0 since the force in the z direction was the same in both halves of the cycle due to the plate shape and orientation remaining constant. In the case of a flapping wing with an initial curvature, the shape of the wing changes with time such that sometimes it is more parallel to the z axis (considering the wing as flat and straight). Depending on its transient position, the forces on the wing surface will be projected differently on the y and z axis. For example, in P4, when the wing is open at $\pi/2$, it is generally more *parallel* to the y axis and hence a larger proportion of the force is in the z-direction (see corresponding wing flap position extreme on the right of the Figure 21). Conversely, at $3\pi/2$, when the wing is closed and more *perpendicular* to the y-axis, the total y force proportion is larger. Looking at P1, at $\pi/2$ the z force also dominates, while at $3\pi/2$, the wing is roughly at a 45° angle to the horizontal, and so the forces are projected almost equally in both directions.

This effect is seen even more clearly when looking at P2 and P5, which both flap at the same amplitude. Here, the y and z transient forces of P2 and P5 are reversed; at P2, most of the force is projected in the z direction, due to the transient wing shape consistently being more perpendicular to the z axis, while the same is true for the y-axis for P5.

In the ideal case, the TFM would have a large amplitude and a phase such that for a given wing flapping shape, the maximum positive force magnitude occurs at the wing position most parallel to the z axis, and the maximum negative force occurring at the position most perpendicular to the z axis. This would approach the set limit for this flapping mechanism given by (3.3).

3.1.5 Skin Drag

One last significant observation is regarding the considerably low TFM amplitude of P3 when compared to P4 in Figure 20. This can be explained by noting how force scales with Re^2 based on the drag force equation presented in Chapter 1 (equation (1.7)). Since the only differences in these cases are that at P3, Re=2.0 and at P4, Re=3.5, the TFM of P3 should be scaled down by a factor of $\left(\frac{10^{3.5}}{10^{2.0}}\right)^2 = 1000$, under the assumption that drag force is still proportional to $Vrms^2$ over this range. If we account for this and overlay the two plots as in Figure 26, we see that they match closely.

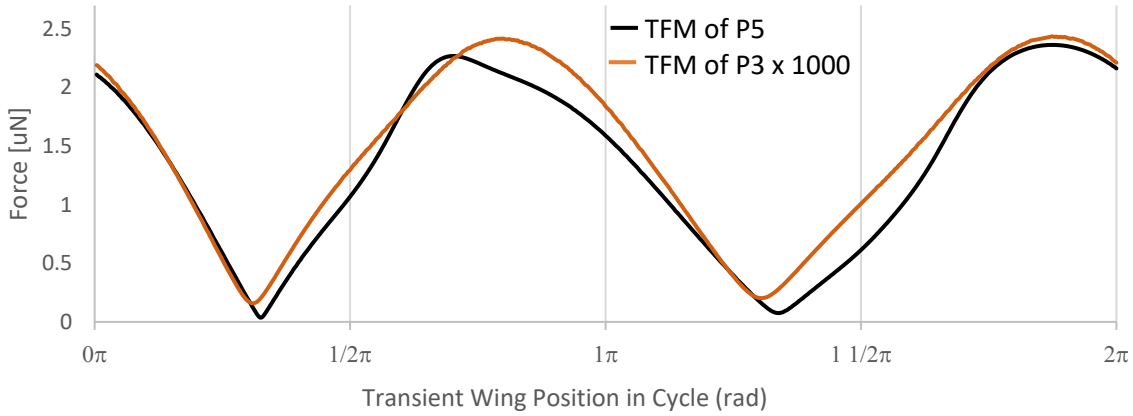


Figure 26: TFM of P3 and P4 with the former scaled by a factor of 1000.

At this Re, the presence of vortices and turbulent effects is small, and we can see this effect from the smoother, sinusoidal shape of P3. (Note how the consistency in the transient magnitudes also justifies Minnick's fortunate choice in defining the velocity scale by the spatial Vrms, for this steady-state analogue.)

There is, however, a contradiction in the above result: since both transient wing flapping shapes (i.e., their resonant modes) are exactly the same, and the TFM are similar, then the Y and Z projections and the C_{DN} value should be very similar as well; however, the actual C_{DN} value for P3 is about half of that for P4 (where $Re=10^{3.5}$) . If we look closely at the overlaid z force plots shown in Figure 27, we see that there is a discrepancy at the second cycle where more negative force occurs at $Re=10^{2.0}$, which contributes to a net lower force (and significantly lower C_{DN}).

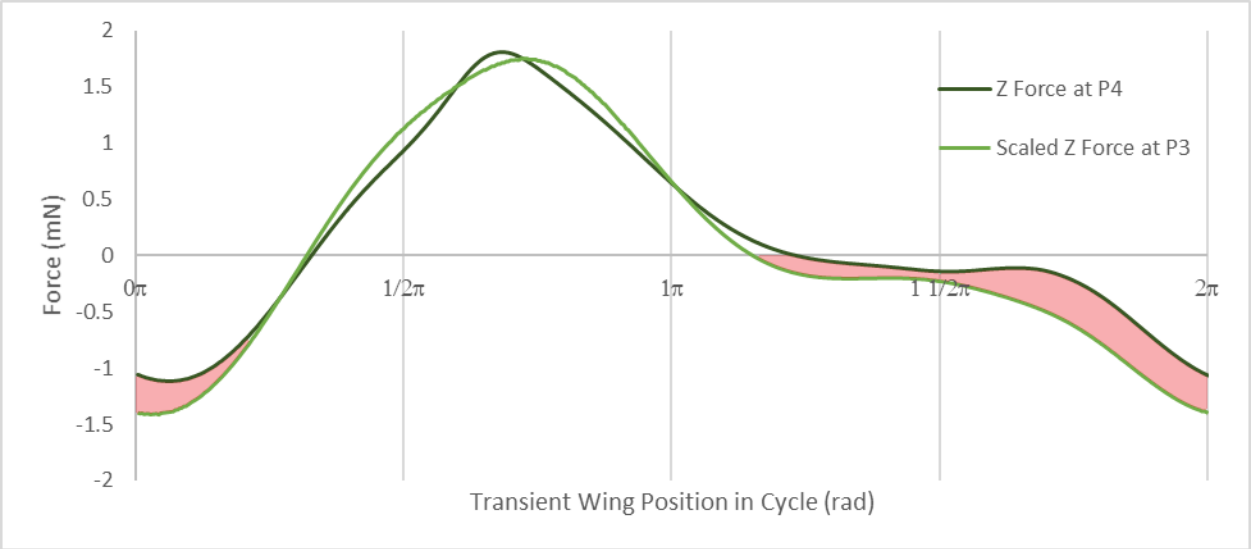


Figure 27: Transient z-force for P4 and P3 (scaled), where the red area shows the larger amount of negative z-force throughout the cycle at P3.

If we look deeper into the origin of the transient z force by splitting it up into the tangential force (i.e., shear force due to viscosity effects), and normal force (i.e., inertial force effects), we see that at parts of the cycle, the tangential forces become significant (see Figure 28). This is expected since at low Re, the ratio between inertial and viscous forces decreases and viscous effects begin to play a more noticeable role. We see that the viscous forces also follow a sinusoidal shape, only shifted down such that there is a more negative contribution in the second half of the cycle (adding to the overall negative z force).

It follows that if we disregard viscous effects, then we are left with only the inertial fluid-wing interactions and thus the inertial forces should match up more closely. Looking at Figure 29, we see that this is the case when only normal force components are considered (which roughly represent inertial forces on the wing).

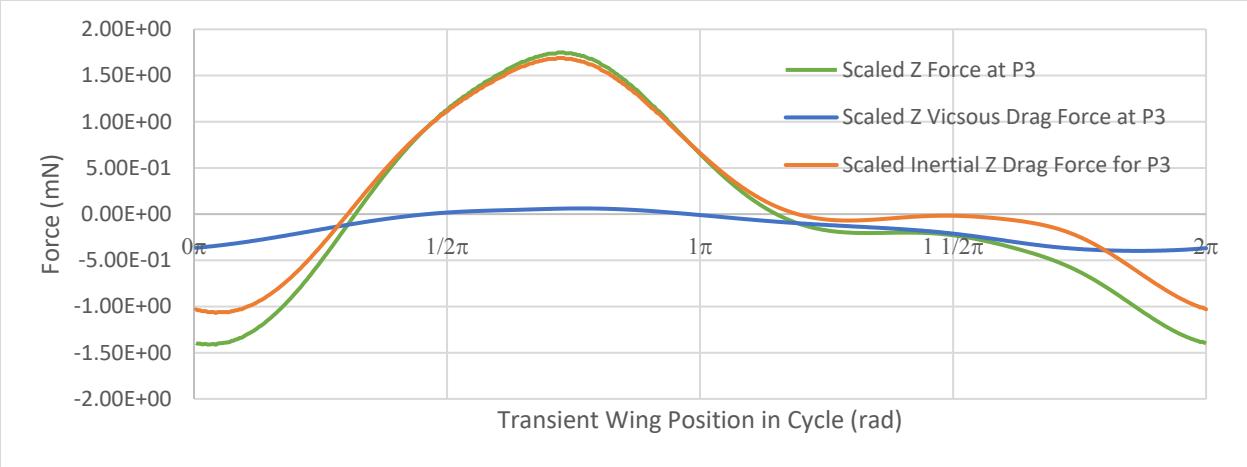


Figure 28: Transient z forces of P3, including the inertial, viscous and total force.

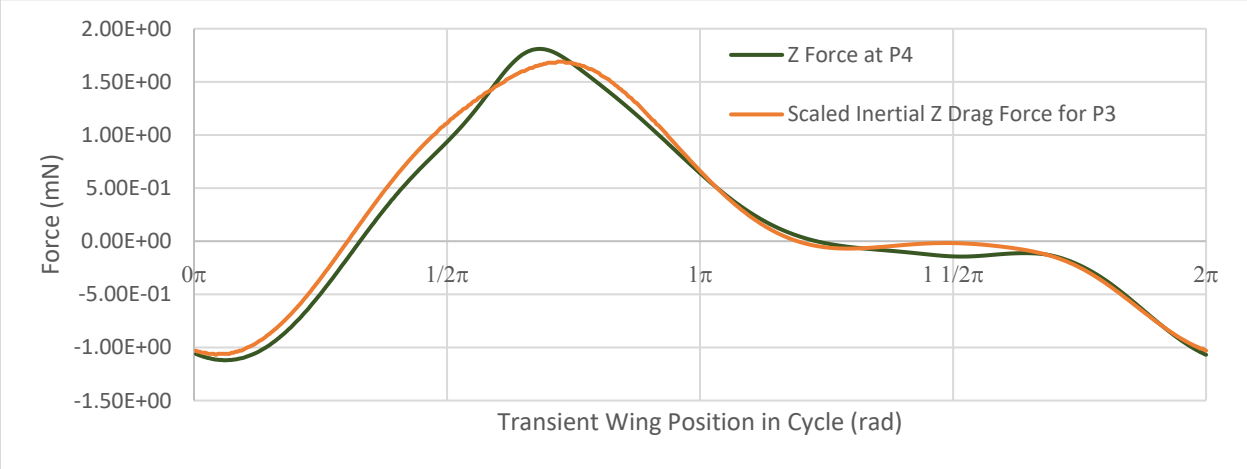


Figure 29: Transient total z-force of P4 and the inertial component of the z-force for P3 (scaled).

The viscous force seen in Figure 28 is referred to as skin friction drag and arises from friction between the boundary layer of the object and the fluid passing tangentially. The size of a boundary layer is related to the area of the object and its surface roughness and so skin friction drag generally increases with increase in object area (where the area normal is perpendicular to the flow). In the case of these wings, at $Re<10^2$, the skin friction drag becomes more comparable in scale to the inertial forces, but its effect related to the wing position is in the opposite sense: when the wing is parallel to the z axis, skin friction drag projects more onto the z axis and 'drags' the wing down. This is shown by the negative shift in the transient viscous drag in the above plot. An illustration of both the inertial and viscous drag contribution is shown below.

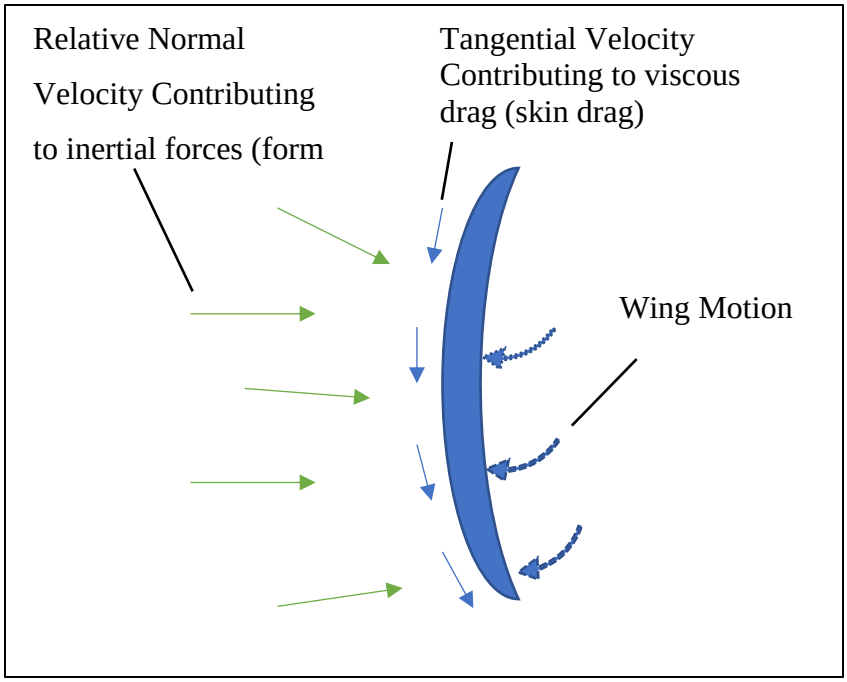


Figure 30: Visualization of the interaction of the viscous and inertial drag on the wing during the end of the downstroke.

Summary of First-Order Effects

Summarized effects on force are presented in the following three tables with each proceeding table showing the relationship to the previous table's dependent variables, eventually leading to the net z force generation. The noticeable and moderately noticeable (second order) effects are highlighted for clarity in green and yellow, respectively. The effects in each cell assume that all other variables stay constant.

Table 3. Tertiary dependencies related to the C_{DN}.

Dependent Variables	Dependencies on C_{DN}			
	$A_r \uparrow$	$\phi \uparrow$	$\omega \uparrow$	$\xi \uparrow$
TFM Amplitude	Decreases	Small Decrease	Low Effect	Low Effect
TFM General Function Shape	Low Effect	Low Effect	Low Effect	Low Effect
TFM Function Asymmetry about Inflections	Low Effect	Low Effect	Low Effect	Low Effect
TFM Phase	Small Decrease	Low Effect	Low Effect	Low Effect
ξ	Increases	Increases	No Effect	-
Inertial Drag/Skin Drag (Re)	Increases	Low Effect	Increases	Low Effect
$\uparrow$ Z-Force Projection Amplitude (at Low Ar)	-	Decreases	Low Effect	-

In Table 3 I have defined the parameter ξ as a scalar quantity describing the transient flapping shape of the wing. Qualitatively, ξ is the physical asymmetry between the flapping extremes about this z axis. Quantitatively, I define it as the ratio of the xy-plane-projected wing area between instantaneous wing shapes at $\pi/2$ and $3\pi/2$.

Table 4. Secondary dependencies related to the C_{DN}.

Dependent Variables	TFM General Function Shape $\uparrow$	TFM Function Asymmetry about minima $\uparrow$	TFM Phase $\uparrow$	$\xi \uparrow$	Inertial Drag/Skin Drag (Re) $\uparrow$
Z-Force Projection Efficiency	Low Effect	Increases	Increases	Increases	Increases

Table 5. Primary dependencies related to the C_{DN}.

	TFM Amplitude ↑	Z Force Projection Efficiency ↑	↑ Z Force Projection Amplitude (at Low A_r)
C_{DN} (Normalized Net Z Force)	Increases	Increases	Increases

Neglecting all variables related to small effects, we arrive at the simpler C_{DN} dependence shown in the following figure.

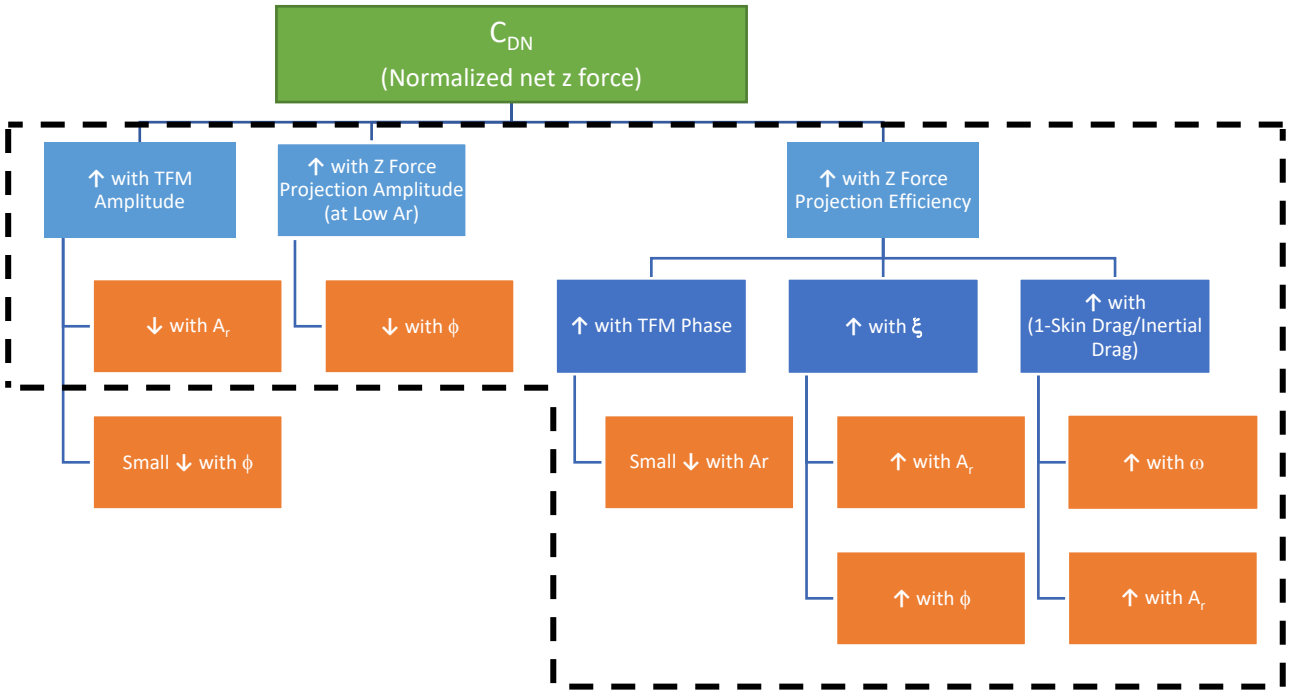


Figure 31: Simplified chart explaining the C_{DN} dependence on fundamental wing flapping parameters A_r, w, and ϕ. The dashed border surrounds the highly correlated effects. The up and down arrows mean increases and decreases, respectively.

Using Figure 31, the 3 phenomena mentioned at the beginning of this chapter can be explained as follows:

"1. In general, above A_r=0.2, the C_{DN} tends to increase with increasing ϕ, up to a peak at ϕ=90°...)."

As ϕ increases, so does the ξ and consequently the z force projection efficiency. At this point, the A_r is relatively high and so the z force projection amplitude effect is small.

"2. For a given Phi0, the C_{DN} generally increases with increasing A_r and increasing Re."

As A_r increases, the TFM magnitude decreases while ξ increases. The result is less transient z force magnitude but higher z force projection efficiency – the latter effect is dominant at higher Ar.

As Re increases (and A_r is held constant), the oscillation frequency ω increases, the skin drag becomes less prominent and so the overall z projection efficiency increases as well. Note that once the skin friction drag becomes negligible, increasing the frequency will not increase the z force projection efficiency since it is not related to the amount (or amplitude) of inertial drag. This explains why at high Re, the C_{DN} plateaus.

"3. An exception to phenomena 2 is an anomaly that occurs at 0.1<Ar <0.2 around and Re of $10^{3.5}$, where relatively high C_{DN} is found, which manifests strongly at low ϕ and gradually reduces with increasing ϕ."

In this region, A_r is relatively low and so the TFM amplitude is much higher. At this point ξ is lower and thus so must be the force projection efficiency, yet the effect of higher TFM amplitude is more prominent here. This also explains why the power used at these regions is generally higher while efficiencies are lower since symmetric z-force projections in both halves of the cycle result in more force cancellation (highest possible efficiency comes from highest asymmetry which was described earlier by equation (3.3)).

Finally, this 'island' region disappears at increasing ϕ because at low A_r and high ϕ, the wings generally oscillate more perpendicularly to the y axis, and so less z force amplitude is projected throughout the cycle. This strongly counteracts the TFM amplitude increase at low Ar. Also, we can note that the TFM amplitude is mainly dependent only on A_r so its increase with A_r is

constant for all ϕ – this may explain why said ‘island’ region remains roughly in the same bounds for each ϕ.

Concluding Remarks

In this chapter we investigated all the main phenomena related to thrust generation in wings of $A_sR = 8$. In the initial work done on such high-amplitude, curved, resonating cantilevers, it was theorized that the main effect responsible for the thrust generation was that the cantilevers had a more aerodynamic shape during the half-cycle responsible for negative thrust generation, when compared to the other half, thus giving a net higher drag in the positive z direction. Here, I have shown that the steady state C_{DN} has little effect on the TFM of the wing, and the strongest net force dependence is actually related to the timely wing flapping shape for efficient z-axis projections, force magnitude dependence on small sweeping motions, and the relative magnitude of viscous effects. The former two appear to be unrelated to the relative wing dimensions and so we can predict that different wing shapes should benefit from these effects similarly at optimal points of Re, φ and A_r (given all other effects remain equal). The magnitude of viscous effects, however, are area and texture dependent, which may present themselves more prominently given designs with larger areas at equivalent Re, and rougher surfaces.

4 WING SHAPE OPTIMIZATIONS

In the previous 2 chapters, we have seen how the transient wing shape and flapping speed are the main parameters affecting C_{DN} and flapping efficiency. In those simulations we varied Re, ϕ and A_r, although one important parameter was held constant: wing shape. The wing shape investigated so far has an aspect ratio A_sR of 8 (i.e., 4:1 length to width ratio). This ratio was chosen simply based off the common wing A_sR of flying animals; however, many of these insects follow complex flapping maneuvers including wing rotation about the joint, which is facilitated by the low moment of inertia of such high A_sR wings. In fact, animals like bats, butterflies and moths that flap in a simpler, up-down motion, tend to have wings of lower A_sRs (see [27] for common flying animal dimensions). Given this and the results of the previous chapter, we will investigate the effects of different wing A_sRs on the force and power to characterize this new parameter that can easily be incorporated into an optimal wing design.

Parameters of Interest

It took thousands of hours to complete the simulations required to create the C_{DN} and C_{DP} plots presented by Minnick [3] so it would be impractical to redo these simulations for varying wing shapes. Fortunately, based on the results of the previous section, it appears that many of the points of optimal force generation can be expected to carry on to different wing dimensions (like the strong dependence on the wing flapping shape). This also implies that results over a few points may suffice to extrapolate the effects at other wing configurations.

Aside from characterization, an important end-goal is to physically create wings that can be used for micro-robotic flight. I chose to investigate the wing dimensional dependence at points P3, P4 and P5 of the previous chapter, shown again in Figure 32. In general, this region was found to have high C_{DN} and power efficiency, so it is a good parameter space to measure the effects of A_sR.

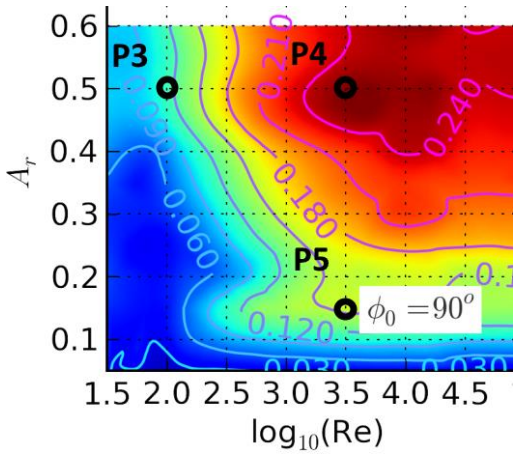


Figure 32: Plot 4 of the C_{DN} plots with points P3-P5 chosen for investigation.

These 3 points also vary in extremes in terms of flapping frequency and amplitude and may present more information to help generalize the effects about the entire parameter space (A_r, Re, ϕ, A_sR).

Note again that the Reynolds number definition in this plot is defined as

$$\mathrm{Re} \equiv \rho_f \frac{V_{rms}\sqrt{LB}}{\mu} \tag{4.1}$$

Where $\sqrt{LB}$ is an arbitrary definition of the Re length scale and is the same as the following equation, considering that the simulations always used a wing with a length to width ratio of 4:1.

$$\mathrm{Re} \equiv \rho_f \frac{V_{rms}L}{2\mu} \tag{4.2}$$

Since $A_sR = \frac{2L}{B}$, one might then express the Re in terms of the A_sR like so:

$$\mathrm{Re} = \rho_f \frac{V_{rms} L}{\mu} \sqrt{\frac{2}{A_s R}} \tag{4.3}$$

The issue with equation (4.3) is that the length scale definition for the Re in (4.1) was originally arbitrarily chosen and not meant to consider wing shape variation. It also shows that if the same Re and A_r values are chosen for simulations of varying A_sR wings, each simulation would have a different V_{rms} (and hence, flapping frequency ω); however, we simply cannot assume that Re scaling works with different shapes as it does with different μ, ρ, A_r and ω. We thus ignore this definition of Re for the moment and run the simulations at equal frequencies based on the frequencies used in points P3, P4 and P5 for the originally simulated wing of A_sR=8. This gives a more fundamental comparison, which I will later relate to the current definition of Re in Chapter 5. For each point, 7 wing shapes are simulated ranging from A_sR=0.25-16 – for clarity, the shapes of the wing pairs are shown in the following figure.

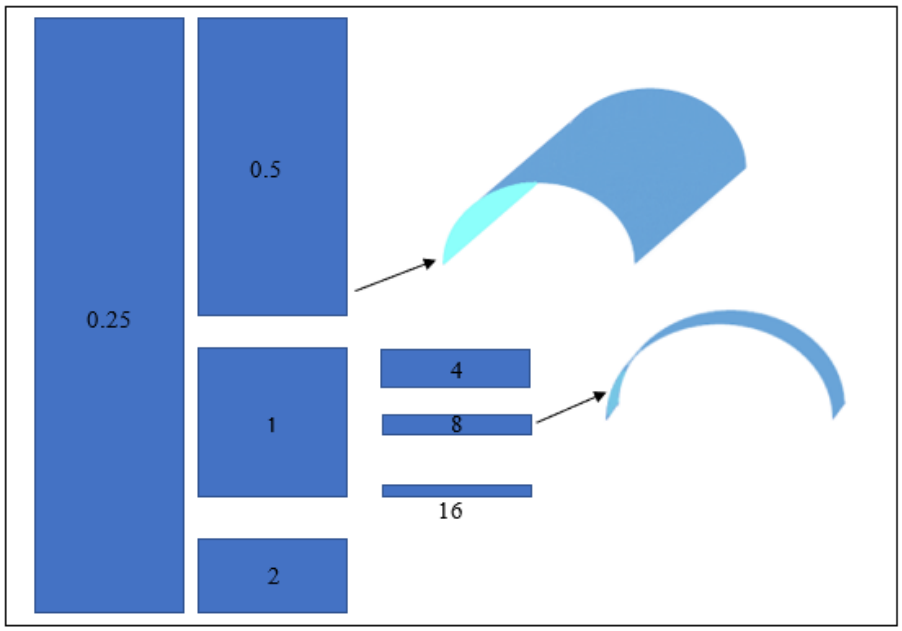


Figure 33: A_sR representation of the simulated wings showing the non-curved wing shapes with the numbers corresponding to the wing pair A_sR. Two isometric views of the curved wings are shown for clarity.

Results of Varying Wing Dimensions

The following are qualitatively important plots regarding force, power and efficiency. By understanding patterns and consistent relationships in the data, we may also extrapolate this

information further to predict what results would be achieved at different points in the parameter space A_r, Re, ϕ, and A_sR. At the very least, these results may be used to select the optimal A_sR for a system with specific force and power requirements.

4.1.1 Force

Figure 34 shows the raw cycle-averaged force results for different wing A_sRs at the parameter points mentioned above. Here, we can see the general result that average thrust decreases with increasing A_sR (i.e., decreasing width-to-length ratio). Figure 35 shows more clearly that this increase is not linearly proportional to the increase in wing width, but approximately follows an $A_sR^{-3/2}$ relationship.

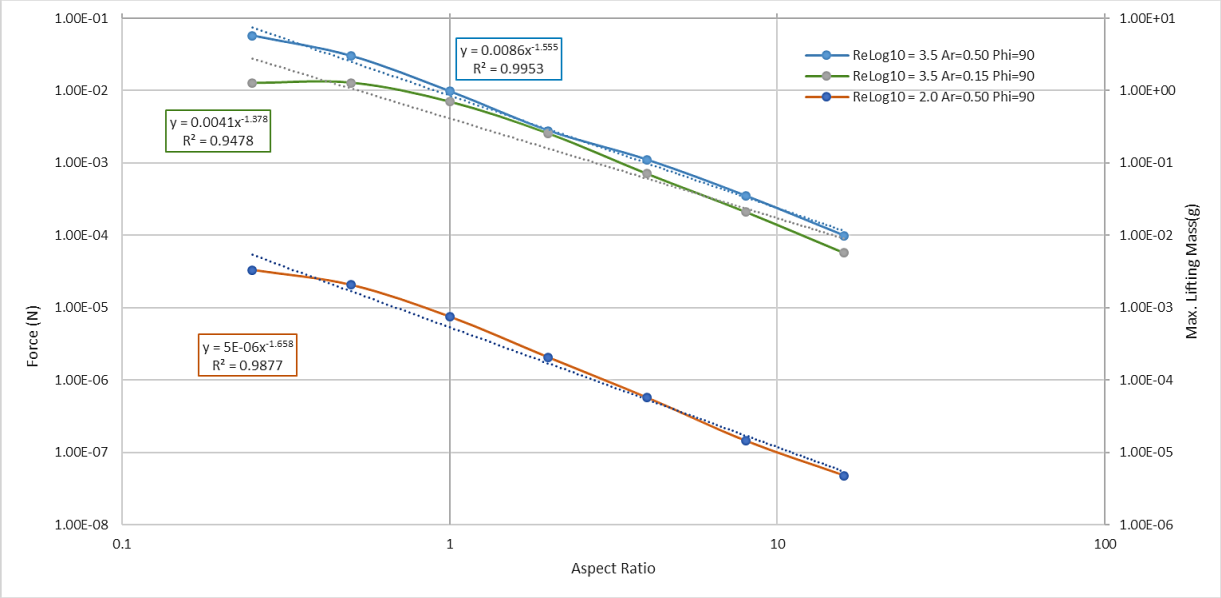


Figure 34: Force and theoretical lifting mass of a wing pair at different aspect ratios, A_sR.

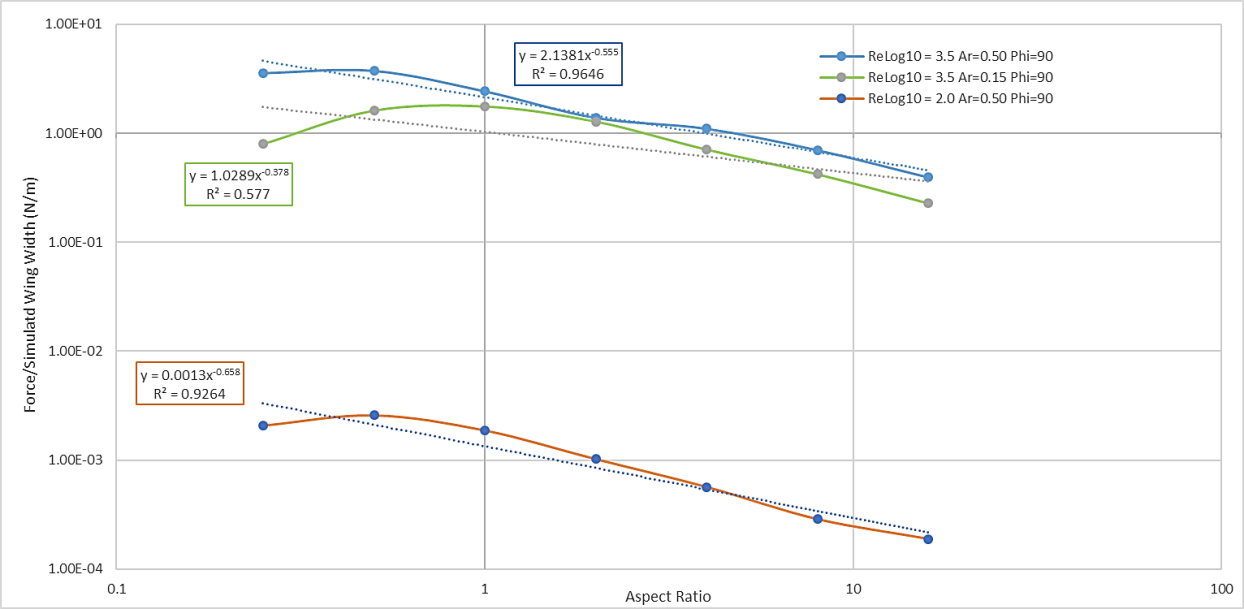


Figure 35: Wing force normalized for wing width at different A_sRs.

To better sense the scale of the forces, Figure 36 shows how many more times a pair of wings could lift their weight when fabricated using length, thickness and materials taken from [3]. Here we can see that no wing design at P3 (low Re) has enough force to lift itself, while P4 and P5 (high Re) are well beyond that threshold. We can also see that at low aspect ratios below 0.5, the added thrust per wing width begins to decrease in all cases, which suggests that there is at least one other z force-related mechanism that is dependent on the wing dimensions appearing at lower A_sRs.

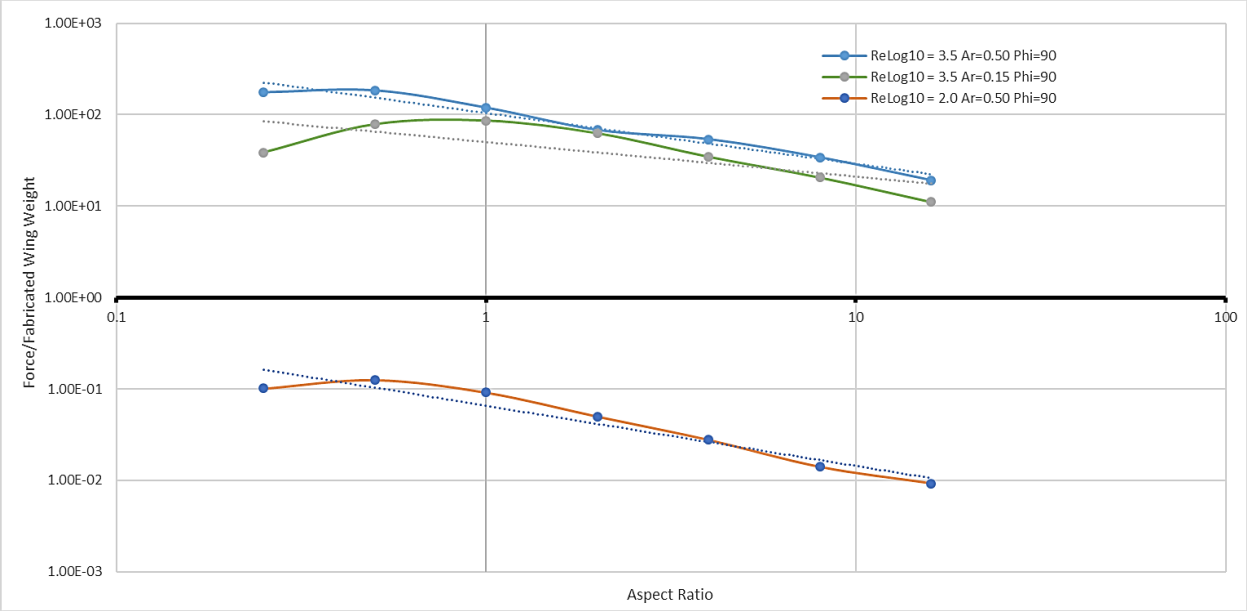


Figure 36: Wing force per theoretical wing weight based on current wing thickness, scale and mean density. Note that the bold line represents the value 1, at which point the wings generate exactly enough force to lift just themselves.

4.1.2 Average Power

Plots showing power-related dependencies on the aspect ratio follow. Figure 37 shows how the average power used by the wings increases and strongly follows the same $A_sR^{-3/2}$ relationship as the wing net force (even more so if we ignore $A_sR > 4$).

Figure 38 shows the efficiency in each wing design in producing a net thrust for the power used. We can see here that the data spread is small with no obvious function dependency. For point P4, the optimal wing design when considering force per power used is around an A_sR of 0.5, although this is only 1.75 times better than the lowest efficiency at this parameter. We can see that for point P5, the optimal region is around an A_sR of 2 and 4.3 times more efficient than the lowest result. P3 has a similar result to P5 except for the optimal A_sR being at 1.

We can also see that the simulations at P3 overall show higher efficiencies when compared to all other points. This is directly related to the V_{rms} dependence of the power used by the wings in equation (1.8). We can divide the results by the V_{rms} to get the 'power force' vs. A_sR as shown in Figure 38. Here we see the higher efficiency of P4 for all A_sR, which is expected since it is

already known that this region has high efficiencies for $A_sR=8$ [3]. Note that Figure 38 shows Force/Power Force which is equivalent to C_{DN}/C_{DP}.

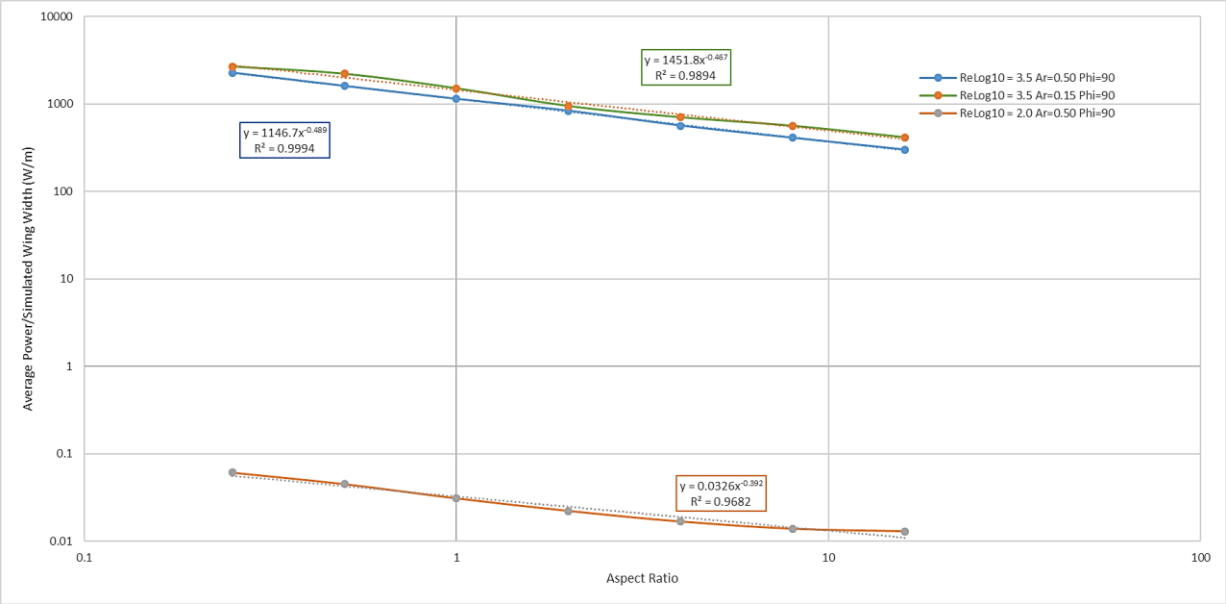


Figure 37: Wing power corrected for wing width, at different aspect ratios. Note that if the two highest A_sR data points are ignored in the power fits, the dependence on the A_sR for P3, P4 and P5 become $x^{-0.474}$, $x^{-0.496}$, and $x^{-0.507}$, respectively (all $R^2 > 0.99$).

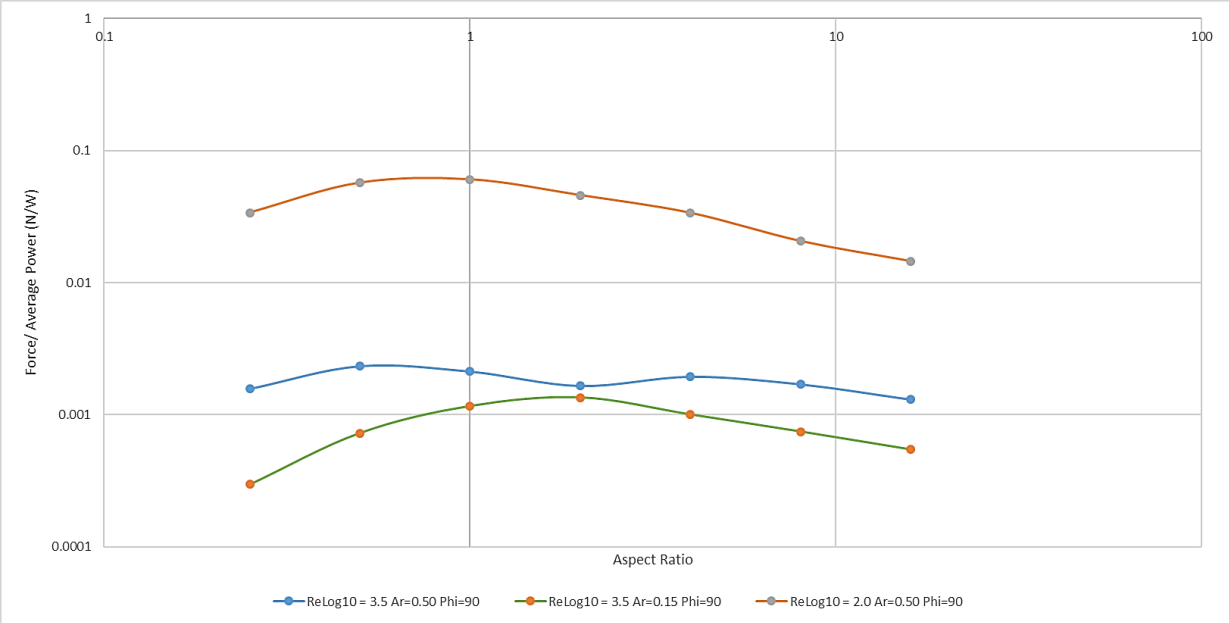


Figure 38: Force per Average Power used in a wing flap cycle at different A_sRs.

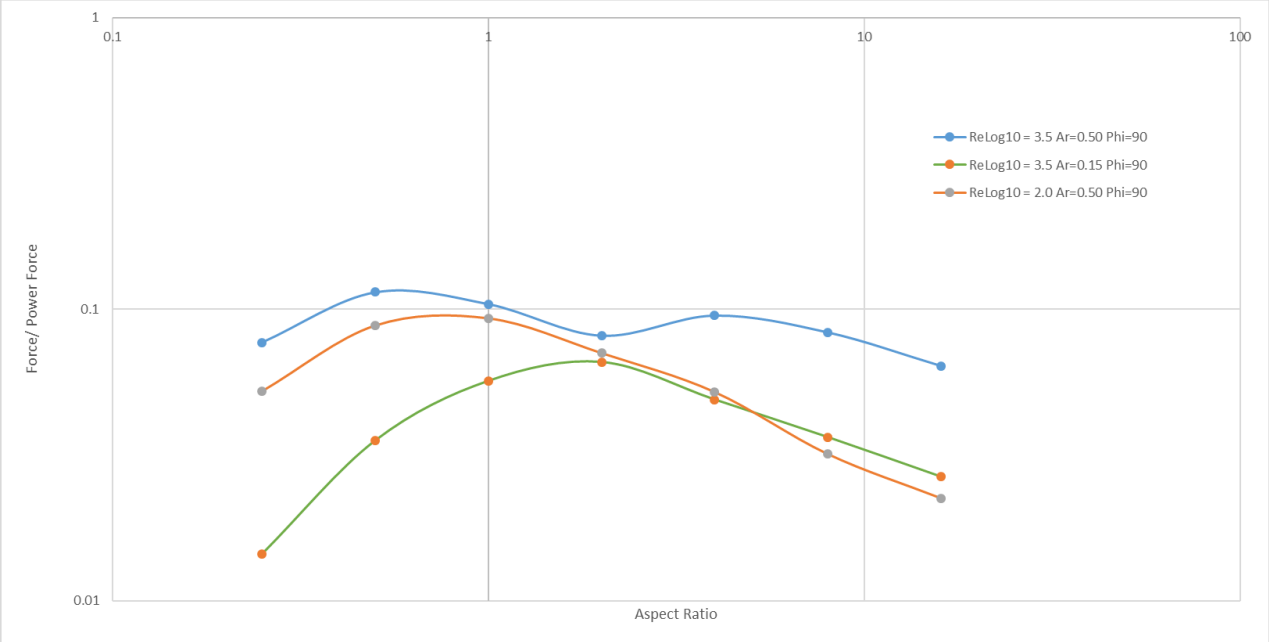


Figure 39: Force per 'power force' at varying A_sRs. Here, the power force is just the average power used by the wing but normalized for the expected V_{rms} dependence.

Discussion

4.1.3 Interpolating Results

The consistency between A_sR dependencies of some plots show that there are regions that can be confidently interpolated or stated as a general rule in the vicinity of that parameter space. For example, although no wing design with $Re=10^{2.0}$ is expected to be able to lift even the wings themselves, we can interpolate between the curves of P3 and P4 and expect that at an A_sR of 0.5 and $Re \approx 10^{2.3}$, the wings should pass this force threshold.

4.1.4 Force dependence and A_sR

If we imagine two pairs of wings producing forces such that the pair does not interact with the other, we can expect that the total force produced between them will simply be double the force produced by a single pair; however, the force results show that if we took these pairs of wings and attached them together, then given all else equal, the wings would generate about $\left(\frac{A_sR_1}{A_sR_2}\right)^{3/2} = \left(\frac{2A_sR_1}{A_sR_1}\right)^{3/2} = (2)^{3/2} \approx 2.8$ times the force.

This large force dependence on A_sR can be compared to a heat generating rod. A single, non-insulated heat generating cylinder dissipates heat in all directions on its surface. A medium surrounding the rod will then conduct the heat away, until a stable solution exists describing the temperature of the rod and the surroundings in 3-D space. If another rod is added to the system very far away from the first rod, then the average temperature of both remains the same; however, if the second rod is attached to the first along the circular faces, heat generated by that face no longer conducts away to the medium, and the average temperature of the two-rod system increases.

In the case of the wings, it is the pressure (instead of heat) that builds up around the wings throughout the cycle. This pressure is primarily responsible for the instantaneous force generation and so an increase in overall pressure increases the overall TFM and net force. During the wing flapping cycle, relative pressure is built up below and above the wing during the first and second halves of the cycles, respectively. This pressure can equalize by means of fluid transport around the wing, which I refer to as 'escape mechanisms' of air.

The three primary escape mechanisms are:

1. Air escaping through the sides (the chord-wise direction, x)

2. Air escaping through the wing tips (the length-wise direction, y & z)

3. Air escaping through the bottom and top planes of the wings (z normal plane), away from the wing tips.

From many simulations, it was found that in case 1 and 2, there is usually a high-pressure gradient near the edges which results in strong local vortices from pressure equalization about the edges. The rest of the pressure farther away from the edges equalizes through outward flow in the direction of the wing curve as seen in Figure 40. Case 3 captures the magnitude of the downwash which directly relates to the instantaneous force on the wing simply from Newton's 3rd law and is necessary for positive thrust.

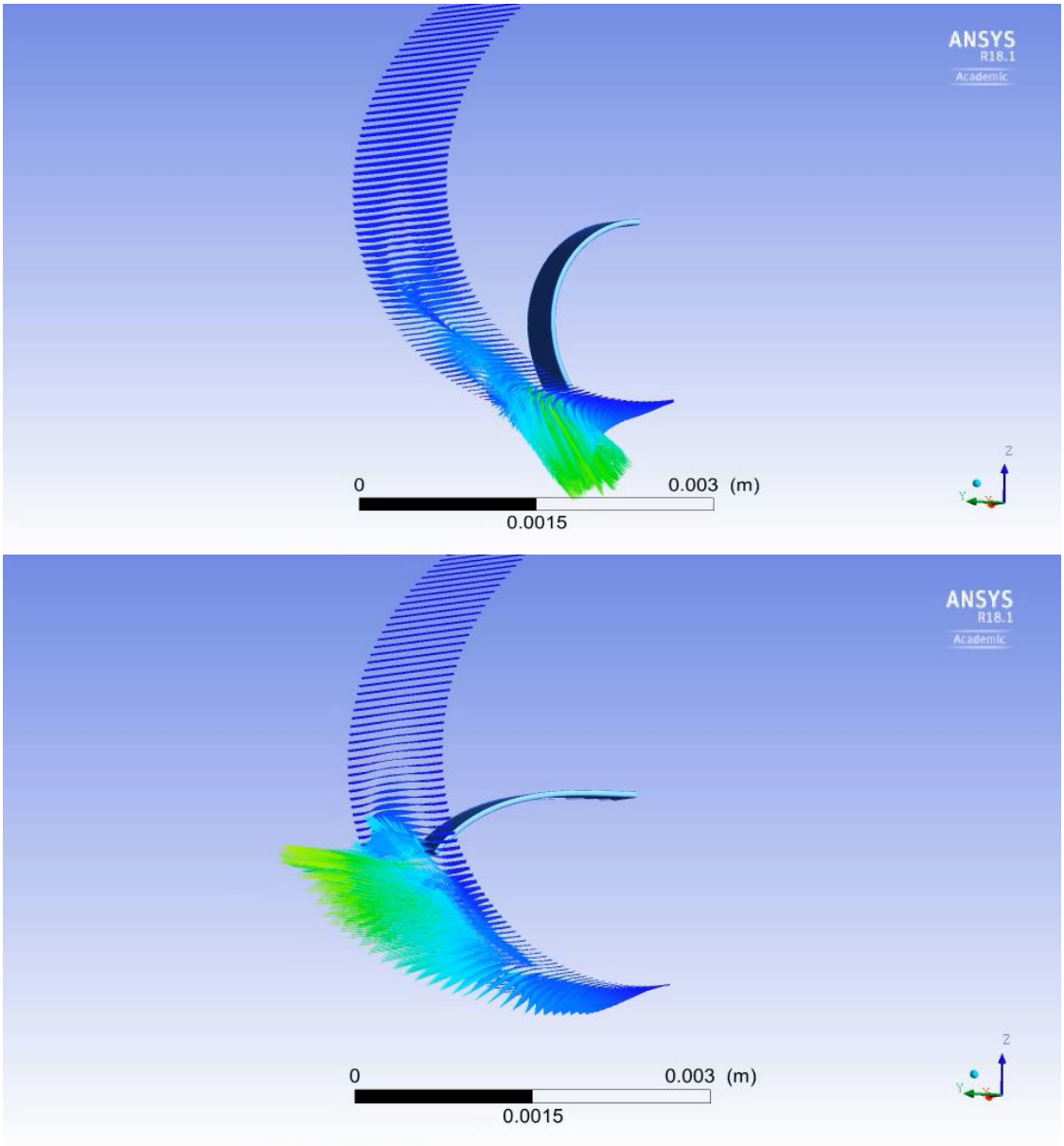


Figure 40: Air flow plane showing the 3D velocity vector along the sweeping path of the wing at the two extremes.

These pressure escape mechanisms are analogous to conduction surfaces in the heated rod example. Like the net temperature increase from joining two hot rods, we can capture more pressure under and above the wing by limiting the side escape mechanism by joining two wings, resulting in more instantaneous net force. Following the analogy, the highest pressure would then be reached near the center which has been verified through simulations and already shown for the square plate in Chapter 2.

To qualitatively verify these escape mechanisms, I averaged the air flow through the side plane of the wing, and through a plane attached perpendicularly to the wing tip. The results of the transient average are presented below.

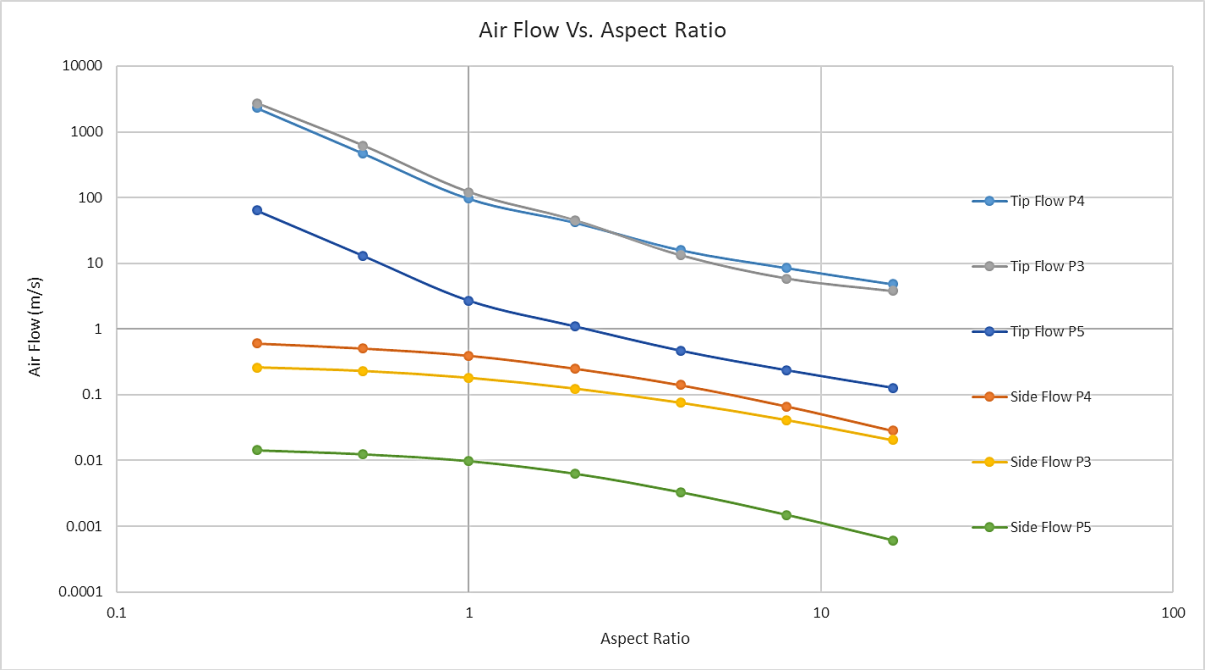


Figure 41: Absolute air flow through wing escape mechanisms at different A_sRs.

From the above figure, we can see that there is a gradual increase in the tip air flows with lower A_sR, while the side flows begin to plateau. This is consistent with the idea that joining wings together (or decreasing their A_sR), results in more pressure built up locally under the wing, yet the escape through the sides is sub-linearly proportionate to the increase in wing width. In the heated rod analogy, the equivalent result is a higher temperature near the center of the rod, while the average temperature of the circular faces increases sub-linearly with rod length.

These results suggest that transient forces should always be increasing with A_sR, although based off P5 in Figure 34, the heated rod analogy breaks down at $A_sR<0.5$, where no significant increase in force is seen with increase in width. If we look at the transient z force plots of P5 in Figure 42, we see that the transient forces appear to be scaling normally for different A_sRs but there is one major difference: the second (negative) half of the cycle disproportionally increases. Figure 43 shows this in the plot of the ratio of the absolute value of the integrated positive and negative force contributions. Here we see that the ratios approach 1 at lower A_sRs meaning that

both the negative and positive contributions are close to equal, even though the transient z force magnitudes increase. The exact reason that this happens requires further investigation, but I presume this to be an effect of the wing-wing interaction, where at higher pressure buildups, the interaction between the two wings becomes noticeable (note that up until now, the simulations included two wings flapping wings joined at the base). When P3-P5 parameters were simulated for a single wing of $A_sR=8$, the transient and net force differences were negligible when compared to the wing pair counterpart and so the effect of the presence of the second wing was deemed negligible; however, at lower A_sR, the single wing always showed a much smaller TFM. This can be seen in Figure 44 in the TFM plots of a wing pair at P4 with $A_sR=0.5$ along with the single wing result.

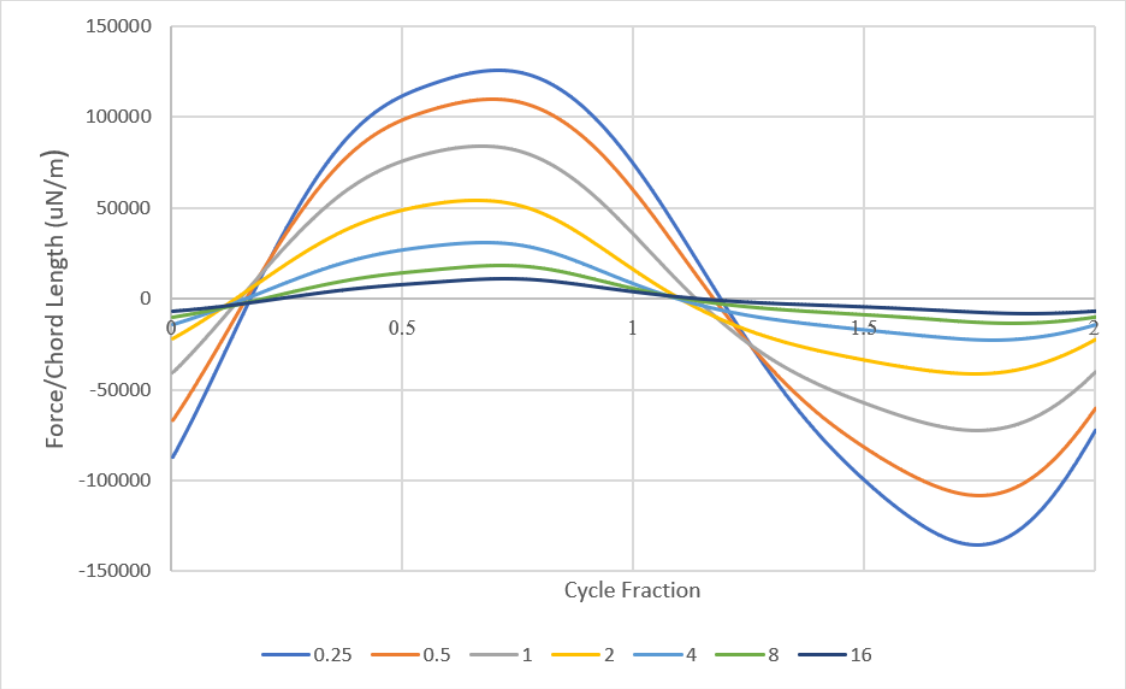


Figure 42: Transient Force/Chord Length plots for different A_sRs at P5 (Ar=0.15)

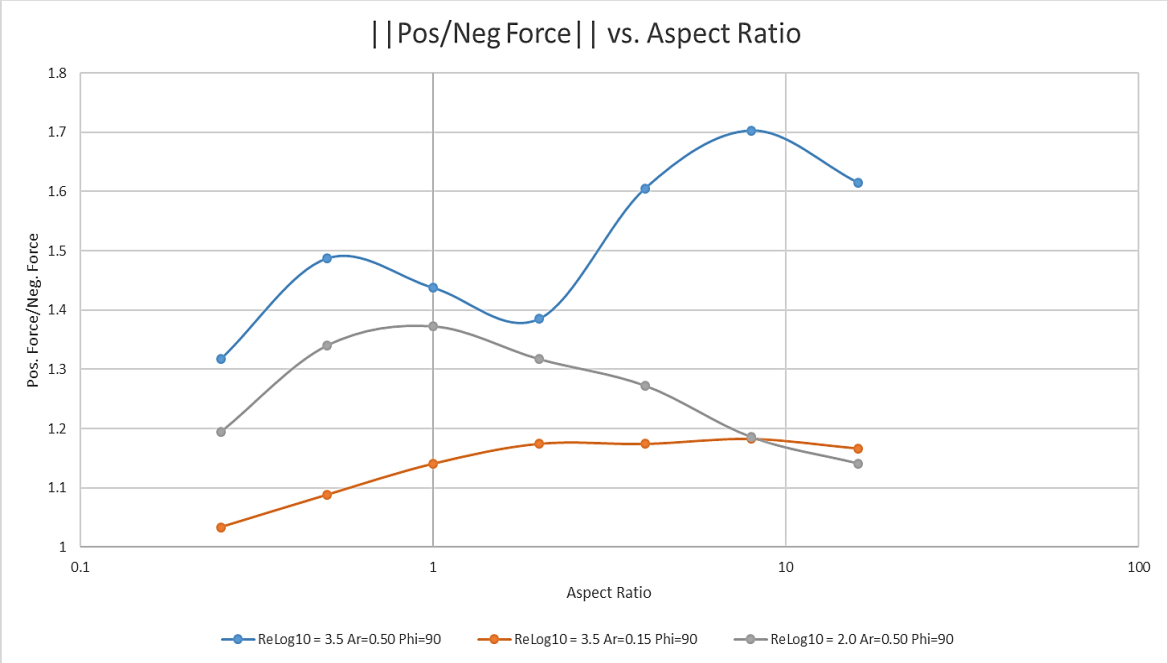


Figure 43: Magnitude of the positive to negative force ratio over a cycle at different A_sRs for P3-P5.

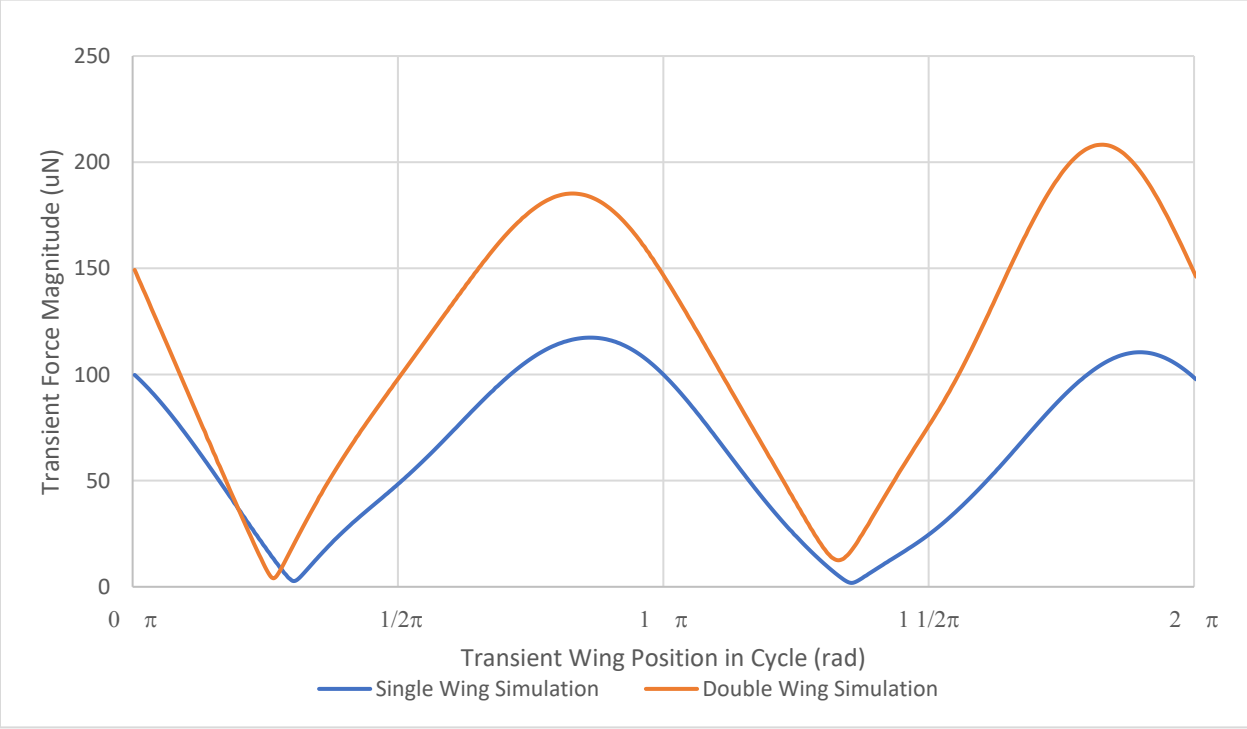


Figure 44: Transient force magnitudes for a single and double wing Simulations at A_sR=0.5 at P4.

The reason for the large discrepancy in the TFM plots of Figure 44 may be due to the possibility that another pressure escape mechanism is with the second wing. At low A_sRs, the wings might interact more by capturing more fluid and creating stronger induced waves and pressure gradients while near each other's vicinity – this would result in overall larger forces during the flapping cycle although more investigation is required. It could also be that extra low pressures occur during the second half cycle when the wings open upward on the reverse stroke and 'pull away' from each other, resulting in the increased negative force – similarly described by part of the "fling" mechanism in insects [6].

Conclusion

Regardless of what exact mechanism accounts for the decrease in efficiency and force at very low A_sRs, this data range is unlikely to be feasible for useful wing flight simply due to the highly unusual wing shape which would likely create mechanical problems and difficulties in the NAV's movements. In general, the results of this chapter show a surprisingly consistent $A_sR^{-3/2}$ dependence for both the force and power and thus can be confidently incorporated in a general way to be used in optimal wing designs. We can see that by simply increasing the wing area, we can achieve larger forces by orders of magnitude while keeping the efficiency relatively constant if actuated in the optimal region (P4), similar to the original C_{DN}/C_{DP} plots. In the next chapter I will use these results along with those of Chapter 3 to quantitatively incorporate wing shape dependence to the old theory and discuss wing designs, fabrications methods and finally show real results for a wing.

5 Designs, Fabrication and Results

In the previous few chapters, we discussed force and power dependencies on wing flapping parameters A_r, ϕ, ω, and A_sR. Conveniently, these parameters can be adjusted while maintaining a key benefit in this type of wing design: monolithic fabrication. In this chapter, optimal wing designs will be discussed, as well as updated fabrication methods that have been improved upon since the original work done by Minnick in [3]. I will also present certain results of the fabrication attempts, failures and key successful designs from which force data was collected.

Optimal Wing Characteristics

5.1.1 Including a new parameter: A_sR

Based on the work presented earlier regarding the net wing force and power used, as well as the wing characteristics affecting them, we can state some general guidelines for creating wings. If we ignore the extreme points of the simulated A_sR (i.e., 16 and 0.25), then the $A_sR^{-3/2}$ relationship matches well with force (and even better with power) and becomes a very good approximation for all 3 points simulated in the previous chapter. If we assume this as a first-order general rule, then we can rewrite the net force equation as

$$F_N = \tfrac{1}{2} C_{DN} \frac{\mu^2 \mathrm{Re}^2}{\rho}$$

$$= \tfrac{1}{2}\left(C_{DN_0} \left(\frac{AsR}{8} \right)^{-3/2} \right) \times \frac{\mu^2 \mathrm{Re}_8^2}{\rho} \tag{5.1}$$

$$= \frac{C_{DN_0} \rho A_r^2 \omega^2 B^{1.5} L^{2.5}}{2\gamma} \tag{5.2}$$

In equation (5.1), Re_8 is the Reynolds number for the originally simulated wing with $A_sR=8$, which served as the basis for simulations of other wing sizes. C_{DN_0} is the original drag coefficient value associated specifically for a wing of $A_sR=8$ (i.e., the force drag coefficient in Figure 6 in Chapter 1, 'Background Information').

Through simulations, Minnick found that the Reynolds number description of the drag equations worked, which allowed for Re scaling and made the results much more useful. With the results presented in Chapter 4, it would be very convenient to now include the newly-found A_sR dependence in the net force and power equations in a general way. By doing so, one would not have to recreate a C_{DN} and C_{DP} plot at specific A_sRs every time a new wing design needed to be made. (Otherwise, the results would only be meaningful for wing designs that are exactly 2mm long, in air, and at 101.325kPa.)

To try to incorporate this, let us first imagine that for a given A_sR, C_{DN} and C_{DP} increase by a constant factor which is a known function of A_sR. Then, for any A_sR, we can easily redraw the C_{DN} and C_{DP} plots exactly by multiplying all points by said factor. For each new C_{DN} and C_{DP} plot, we can expect that Re scaling can be applied the same way as before. If this is assumed to be true for wings for all A_sR, then we can simply group the A_sR dependence as an adjustment factor to the drag coefficients directly, and we have

$$F = \frac{1}{2} C_{DN} \frac{\mu^2 \mathrm{Re}_8^2}{\rho_f} \tag{5.3}$$

$$P = V_{rms} \frac{1}{2} C_{DP} \frac{\mu^2 \mathrm{Re}_8^2}{\rho_f} \tag{5.4}$$

$$C_{DN} = C_{DN_0}\left(\frac{A_s R}{8}\right)^{-3/2} = C_{DN_0}\left(\frac{L}{4B}\right)^{-3/2} \tag{5.5}$$

$$C_{DP} = C_{DP_0}\left(\frac{A_s R}{8}\right)^{-3/2} = C_{DP_0}\left(\frac{L}{4B}\right)^{-3/2} \tag{5.6}$$

Where $\mathrm{Re}_8 \equiv \rho_{fluid}\frac{V_{rms}L}{2\mu}$ retains Minnick's definition specifically for the $A_sR=8$ wings. Further investigation into the consistency of this relationship is necessary, but if found to be accurate in other Re, A_r and ϕ ranges, this definition will widely broaden the usefulness of the C_{DN} and C_{DP} plots from earlier. For the moment, we can be confident that it works for $\phi = 90°$ between the parameters bounded by the triangle formed by points P3-P5 from the previous chapter as shown in the Figure 45. The MEMS engineer can then use these simpler equations to carefully consider the dimensions of a wing, or in general for predictable fluidic force and power effects on differently shaped cantilevers (within the A_sR limits mentioned earlier).

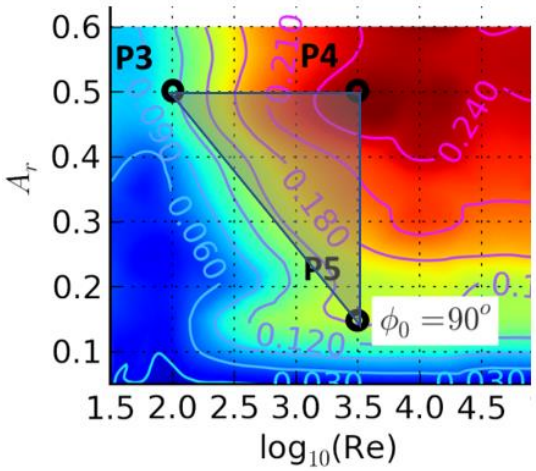


Figure 45: Parameter space where we can assume that the $A_sR^{-3/2}$ dependence works.

5.1.2 Optimal Parameters

Using the above equations (5.3)-(5.6), we can now discuss what values of A_r, ϕ, w, and A_sR should generally be chosen for flight. For simplicity, we will discuss the case where the robot has an unlimited power source and needs to generate the maximum force possible for hover-flight.

Based on this, the wing should have a high C_{DN} value and a high Re. Note that force is proportional to the square of the Re while linearly proportional to C_{DN}.

In general, the region covered by $0.4<A_r<0.6$ and $10^{3.0}<Re<10^{5.0}$ has an almost constant, high C_{DN} while the location of highest $C_{DN_0} * Re_8^2$ (i.e., force) lies around an $A_r=0.5$ and $Re=10^5$. In earlier work for wings of $A_sR=8$, it was difficult to reach this region because it requires that the wings have very high resonant frequencies or be very long. For example, 1cm wings flapping at $A_r=0.5$ in air at standard temperature and pressure (STP) would need to resonate at 173 Hz to be at $Re=10^3$ and 17300Hz to be at $Re=10^5$. The higher resonant frequency is difficult to obtain due to material and bimorph structure limitations when considering that the wing needs to flap at high A_r and initially be highly curved. The alternative would be to increase the wing from a 1cm to a 1m length, but this is counterproductive for creating NAVs.

Now, based on the equations (5.5) and (5.6), we can see that even if the optimal C_{DN0} or C_{DP0} region cannot be reached, the overall C_{DN} or C_{DP} values can be increased by decreasing the A_sR. Note that the resonant frequency is expected to decrease with decreasing A_sR as shown by equation (1.3), but this change will be low for wings with high densities relative to the fluid – for lower relative densities and low A_sR, the designer should keep this effect in mind in case the tradeoff for higher C_{DN} values is not worth the resultant lower resonant frequency. Thus, in general, to create wings that generate high force one should aim for:

a) a high ϕ, A_r, ω and low A_sR to achieve high C_{DN}.

b) a high resonant frequency to achieve high Re.

One caveat is that net power used also scales proportionally to V_{rms} (which scales with frequency) so there is a tradeoff between higher force and the likely heavier power supply/battery required to sustain it.

Minnick's work has also shown that points a) and b) are generally true for efficient power use; since it was found (to a first order approximation) that the power-drag scales like force for varying A_sR, we can expect this efficiency rule to be similar at different A_sRs as well. This is particularly true for P4 of Figures 38 and 39 from Chapter 4 where the curve force/power used is relatively constant across all A_sR. That being said, from the same figures, we see that P3 and P5

curves are not as constant with A_sR and suggest optimal efficiency around $1<A_sR<2$; thus, when considering the highest possible efficiency where increases in factors of 3-4 are necessary, the engineer should keep in mind this region depending on what the cantilever parameters A_r, ϕ and Re are expected to be.

Wing Design Considerations

When making a monolithically fabricated, piezoelectric cantilever to work as a wing, I found that some subtle issues need to be considered which put constraints on the wing shape and design. These points are listed as follows:

Arcing – To actuate the wing, it must have a conductive layer above and below the piezoelectric base, connected to the power source. The electric field between these layers is responsible for elongating and contracting the piezoelectric material, which generates the flapping motion. For small wing weight, high initial curvature ϕ, high flapping amplitudes per applied voltage, and high resonant frequency, it is generally favorable to have relatively thin wings (about 1000x less than the length) [3]. One issue with making the wing too thin is that the actuating voltage and distance between the electrodes may be near the 'Paschen minimum' where the onset of arcing begins. Aside from creating a direct short across the wing faces, an enormous amount of heat would be generated in the plasma, which further ionizes local molecules, and would likely cause instant damage to the wing. In air at STP, this minimum occurs at a distance of about 7.5um at a minimum voltage of about 330V [28]. Avoiding the Paschen minima can be done by making the separation thinner or thicker but the tradeoff would be a lower resonance frequency or a lower initial bending angle, respectively. This would place the operational wing in a non-ideal part of the C_{DN} plots as discussed earlier. A potential solution is presented in my wing redesign in section 5.1.9 inspired by the results from the first wing test, where we will see that failure from arcing can be disastrous.

Surface Smoothness – As we saw in Chapter 3, skin drag was the main effect responsible for the decrease in C_{DN} with decreasing Re. This effect is negligible at high Re but it is possible that a rough surface can increase the amount of skin drag overall and lower the C_{DN} values in the low Re regime. Thus, it is important to keep the surface of the wings smooth and uniform.

Edge Curving – During the fabrication steps which we will see later, the transverse stress induced in the x-y plane of a cantilever will, expectedly, induce strain in these directions as well. For wings of high A_sR, the strain in the x direction (the wing chord) is small and curvature will be approximately 1 dimensional along the y-axis. For wings of low A_sR, the strain in the x-direction may become significant, resulting in curled corners of the wing. Simulations show that this curling does exist but is very small even for high stresses and small thicknesses (see Figure 46a-c for simulated 3D stress on cantilever of $A_sR=1$). The effect of this curl on resonance and fluidic interaction is a point for further investigation.

Stability in Flight – Imperfections in the wing fabrication technique may lead to different resonant frequencies for each wing and thus, force imbalances, that need to be stabilized by appropriately controlling the voltage and frequency of each wing separately. As in typical airfoils, designs with low A_sRs will have a lower moment of rotational inertia and thus will benefit from higher angular acceleration (faster roll); however, the stability in the tendency to roll will be more difficult to control for this reason as well. Faster on-board processors and careful PID tuning will be required for wings of lower A_sRs.

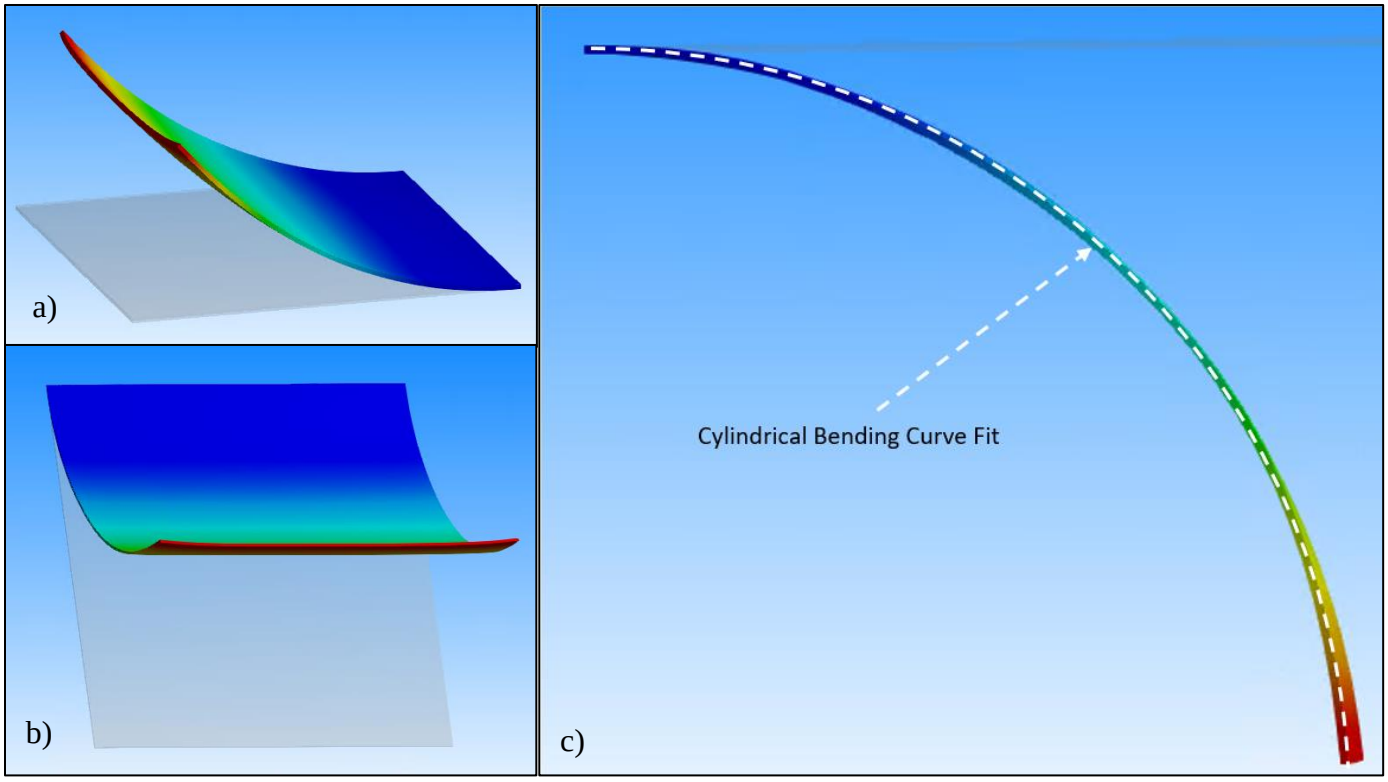


Figure 46: a) & b) Multiple views of a cantilever with uniform surface stress fixed at the base. c) side view of the bent cantilever with a circular curvature fitted, showing deviation from the 1D expectation.

Fabrication

5.1.3 Material Choice

In Minnick's work, a stack structure with x-cut quartz was used as the piezoelectric base. A coat of the photoresist SU-8 covered this base which created uniform stress along the surface after being exposed and cured. Although this resulted in an initially highly curved wing, SU-8 had the issue of changing its mechanical properties based on the surrounding environment; thus, the amount of curvature and warpage greatly changed during experiments and so this turned out to be an impractical material for testing and functionality.

A better material was needed with a high Youngs modulus for higher resonance frequency that could also be deposited onto the piezoelectric crystal in a way that was simple and had built-in intrinsic stress in the layer. In collaboration with Dr. Minnick, we found that Molybdenum was a suitable choice as its Youngs modulus is high and can be sputtered as a compressive or tensile

film. After we characterized the stress dependence of the film on parameters such as thickness, initial chamber pressure, sputtering pressure and power, I was able to create working molybdenum-quartz wings. The layers of the wings structure are presented in the following table in order of the layers on the wing, from top to bottom.

Table 6: Wing layers and thickness in top-to-bottom order for the new wing design.

Material	**Thickness (μm)**
Au	0.15
Mo (Tensile)	2.0
Quartz	6.8
Mo (Compressive)	0.1
Au	0.1

5.1.4 Shadow Mask Preparations

There are at least 3 shadow masks required for fabricating the old wing design from [3]. Originally, they were made for fabricating a wing with $A_sR=8$, while the new masks were made to fabricate wider wings. I chose the dimensions of the new masks to maximize the area used of a 1" wafer to obtain high curvature (which is length based) and lower A_sR to get into the higher C_{DN} regions. The 3 masks are described as follows:

1. Thinning Mask – used for thinning the area where the wings will be developed. I could not find a company to provide a 1" quartz wafer below 10um-thick, and so this thinning mask is required for thinning thicker masks later in the Reactive Ion Etching machine (RIE).

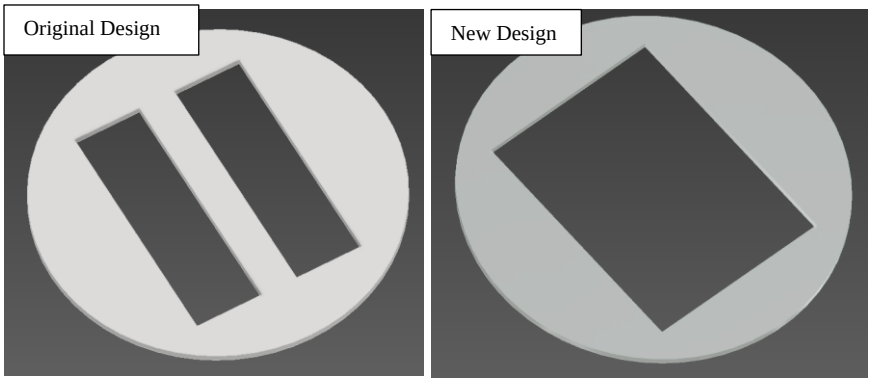


Figure 47: Original and updated release masks made (to be made from SiO_2)

2. Sputtering Mask – used for sputtering the conductive (Mo and Au) layers onto the quartz in the wing shape and for contacts. Original mask used in Minnick's work is shown in Figure 48a while the two masks I developed for wider wings is shown in Figure 48b. The first mask of the new designs is used to sputter on the top while the other is for the underside of the wing. Originally one mask was used for both sides, but I made a second mask of smaller sputtering area for the purpose of avoiding arcing at high voltages around the perimeter.

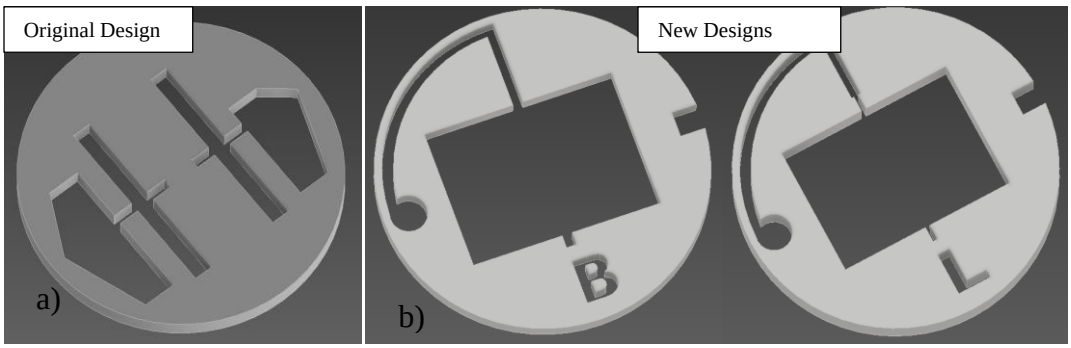


Figure 48: Sputtering shadow masks. Note that the original design (left) was used for both top and bottom surfaces, while the updated design has two masks of slightly different dimensions to avoid arcing and allow higher applied voltages across the wing. A notch has been included for alignment aid and the contacts are led away from the bridge for wires to not interfere with the visibility of the wings when looking in the x-direction (front view).

3. Release Mask – this mask covers sputtered area so that the bare quartz surrounding the wings can be etched away and thus, ‘released’. These masks along with the original designs can be seen in Figure 49.

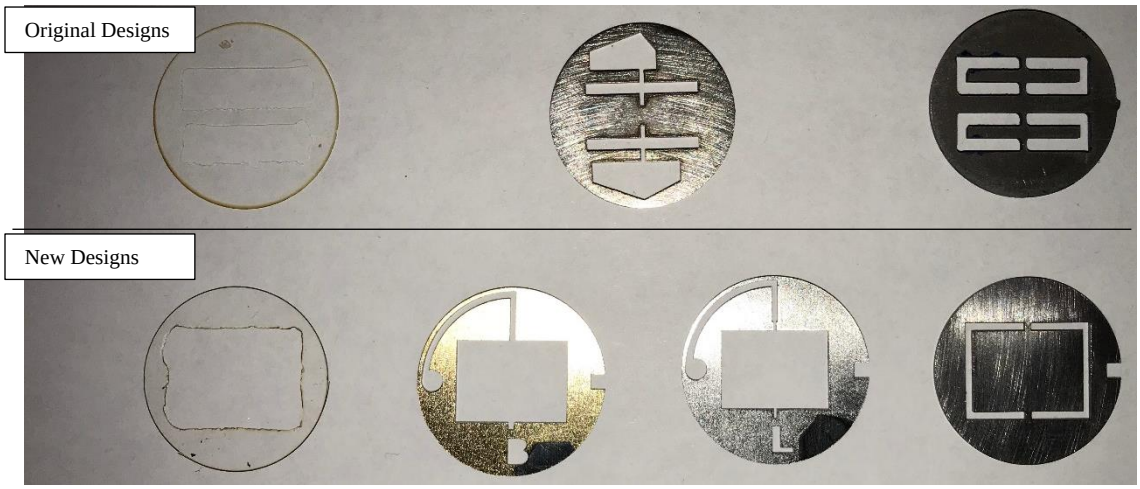


Figure 49: Final mask creations including Minnick’s original designs (top) and my lower A_sR wing designs (bottom). From left to right, the masks shown are for: thinning, sputtering and releasing.

Note that the thinning mask must be made from SiO_2 to not contaminate the quartz sample and cause pitting [29]. Unfortunately, most of the time the SiO_2 masks were created by careful hand milling which usually gave imprecise cuts (as seen in Figure 49 above in the transparent SiO_2 wafers). Precise laser ablation would solve this problem, but this is difficult due to the material’s transparency to visible light. My attempts at this using an available 50W pulsed CO_2 laser were unsuccessful due to long pulse times and large spot sizes leading to high local heat generation and cracking (shown in Figure 50). For the future, UV femtosecond lasers should be considered since they are more suitable for precision SiO_2 cuts due to the extremely short pulses which result in a precisely localized heat distribution [30].

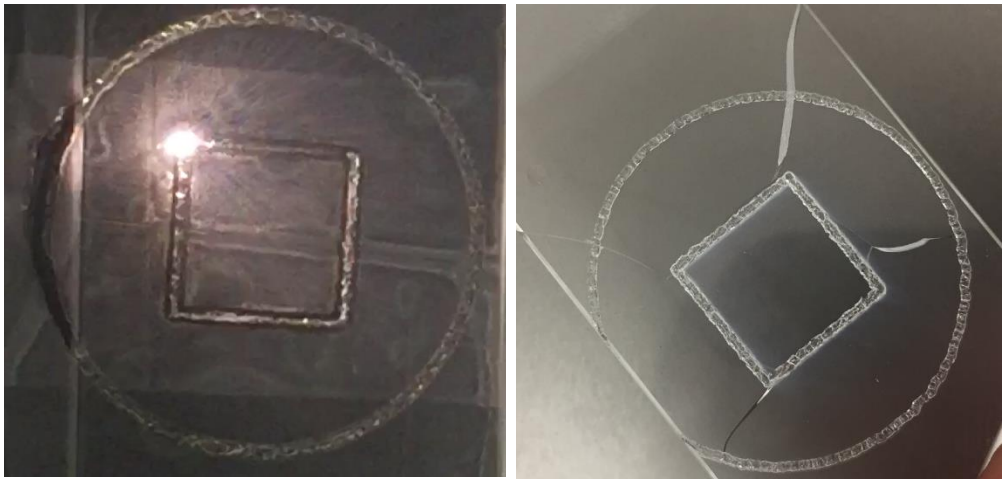


Figure 50: One of many attempts to cut out a shadow mask from an SiO_2 wafer of thicknesses varying from 0.4-1mm using a 50W CO2 laser. Left: Ablation during lasing. Right: Cracks appearing after ablating less than 0.1mm into the slide.

5.1.5 Measuring Thickness

As mentioned earlier, precise layer thickness is crucial for fabricated wings with selected properties. In Minnick's work, an optical profilometer was used to find interference fringes on the top and bottom of the quartz wing, but this method was dependent on the researcher's ability to differentiate between interference patterns from different interference sources. Furthermore, equipment failure lead to abandonment of this technique. Later attempts using an Alpha-Step stylus profilometer, micrometers, and weight scales gave inconsistent or error-prone results (see Figure 51 for the results of an RIE etch with etching time calibrated using a measured thickness value off by 3um from 100um). Ultimately, I created a more automated method based on FTIR spectra for IR radiation passing through the quartz. This has the same benefit of being precise and non-contact as the optical profilometer, but the results are more consistent between measurements due to eliminating the human factor. (See Appendix for details on this refined technique and the MATLAB code for post processing the spectra.)

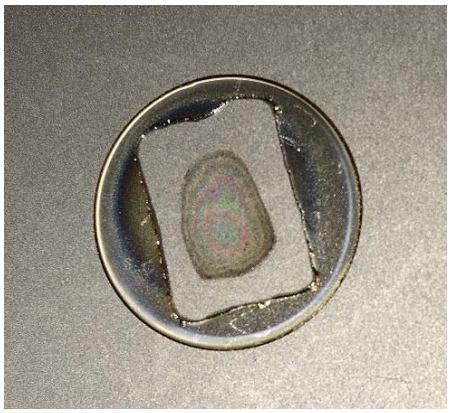


Figure 51: Result of a 3µm inaccuracy in the measured initial thickness. The oval shaped dark patch in the middle is a very thin film of SiO_2 (<500nm) left from the etching procedure. What should be there is a 7um thick area that is still attached to the rest of the wafer.

5.1.6 Procedure

The fabrication procedure is similar to that shown in [3] with the major difference being the coatings and deposition steps. It is outlined as follows:

1. The piezoelectric substrate (x-cut quartz) is thinned to the predetermined thickness (7um). The SiO_2 mask is used during the thinning process to selectively thin the wing region and keep the rest of the wafer thick for later handling. During the thinning, the sample thickness can be successively measured using the FTIR technique until the desired thickness is reached.
2. Both sides of the substrate are then sputtered with molybdenum and gold using a DC Magneton Sputtering machine, using the steel masks shown in Figure 48. The resultant Mo film can be in compressive or tensile stress, depending on a wide range of parameters. In my trials, using A_r as the chamber medium, it was found that highest tensile stress generally occurred at higher powers, lower initial vacuum pressures, and at a specific sputtering pressure. The sputtering pressure ultimately determined if the deposited layer was tensile or compressive, as commonly reported [31]–[33]. Exact values found for achieving maximum tensile and compressive stresses are shown in the following table.

Table 7: Optimal parameters for sputtering Mo on quartz using DC Magneton Sputtering machine. Note that for tensile stress, lower initial vacuum pressures produced higher stress, but due to outgassing and pump speed, the limit in our machine was about 6e-6Torr which was reached after pumping down for 24 hours. The large error in the working pressure is due to the pump changing internal pressure sensors in that regime, causing uncertainty.

	Tensile	**Compressive**
Initial Vacuum Pressure	6e-6Torr (or lower)	1e-4Torr
Working Pressure	7.21+/-3 e-3Torr	3.6e-3Torr
DC Power	250W (maxed)	250W (maxed)

3. The sample is then placed into the RIE with the release shadow mask to thin the area near the border of the wing shape and 'cut out' the wing shape. If the alignment has been done correctly and all the quartz around the boarder has been etched, the wing is released and can be mounted on a PCB for testing. Detailed RIE specifications can be found in [3], [29].

Molybdenum-Quartz Wings Test 1: $A_sR=8$

5.1.7 Results

Initially, to test the practicality of a Molybdenum-Quartz-based stack structure, the first wing shape was made to have the same A_sR as that used by Minnick with the SU-8 test. Figure 52 shows the final wing pair developed on a 1" quartz wafer while Figure 53 shows the wings flapping in the force measuring torsion bar. Note that high curvature was not achieved due to the difficulty in precisely controlling the sputtering parameters during sputtering. Since few simulations were done for $\phi<35°$ and forces were expected to be smaller anyway, the wing with $\phi=37°$ was primarily investigated.

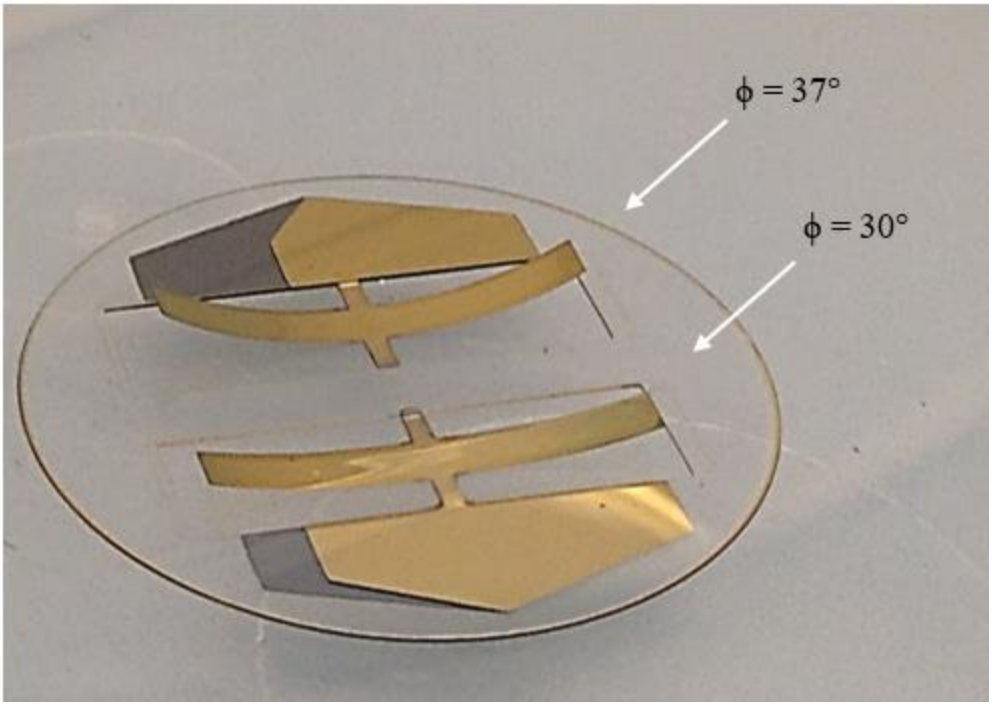


Figure 52: First successfully developed Molybdenum-Quartz wings fabricated monolithically using 1" x-cut quartz substrate.

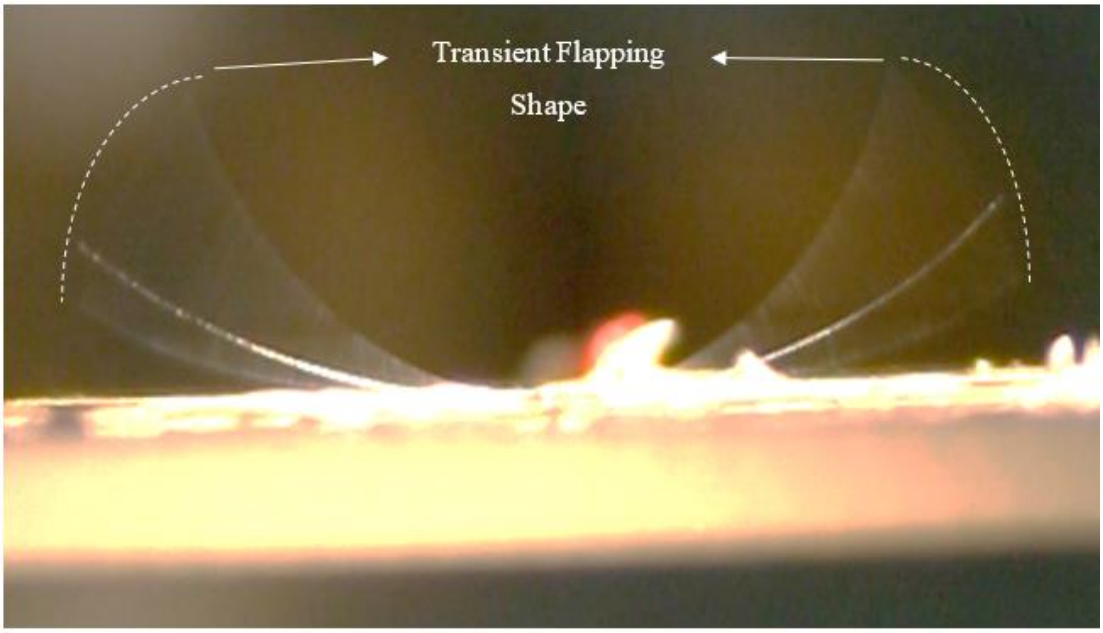


Figure 53: Side view of flapping wing pair mounted on the PCB with 190V applied at resonance (148Hz). The stationary wing within the flapping blur is the deactivated second wing pair with less initial curvature (ϕ=30º).

This wing was initially tested in air at STP, and later in SF_6 to operate the wings at a different Re. The force results in air are shown in the following plot which span applied voltages from 50-190V in intervals of 10V (force results in SF_6 are not shown due to unreliable data from the torsion bar). The overlapping data points represent data from two frequency sweeps for increasing and decreasing frequency. This was done for confirmation that the wings were not being operated in a non-linear regime where resonance hysteresis occurs. Each data point is the result of averaging the force on the torsion bar between 20-60 times to eliminate noise.

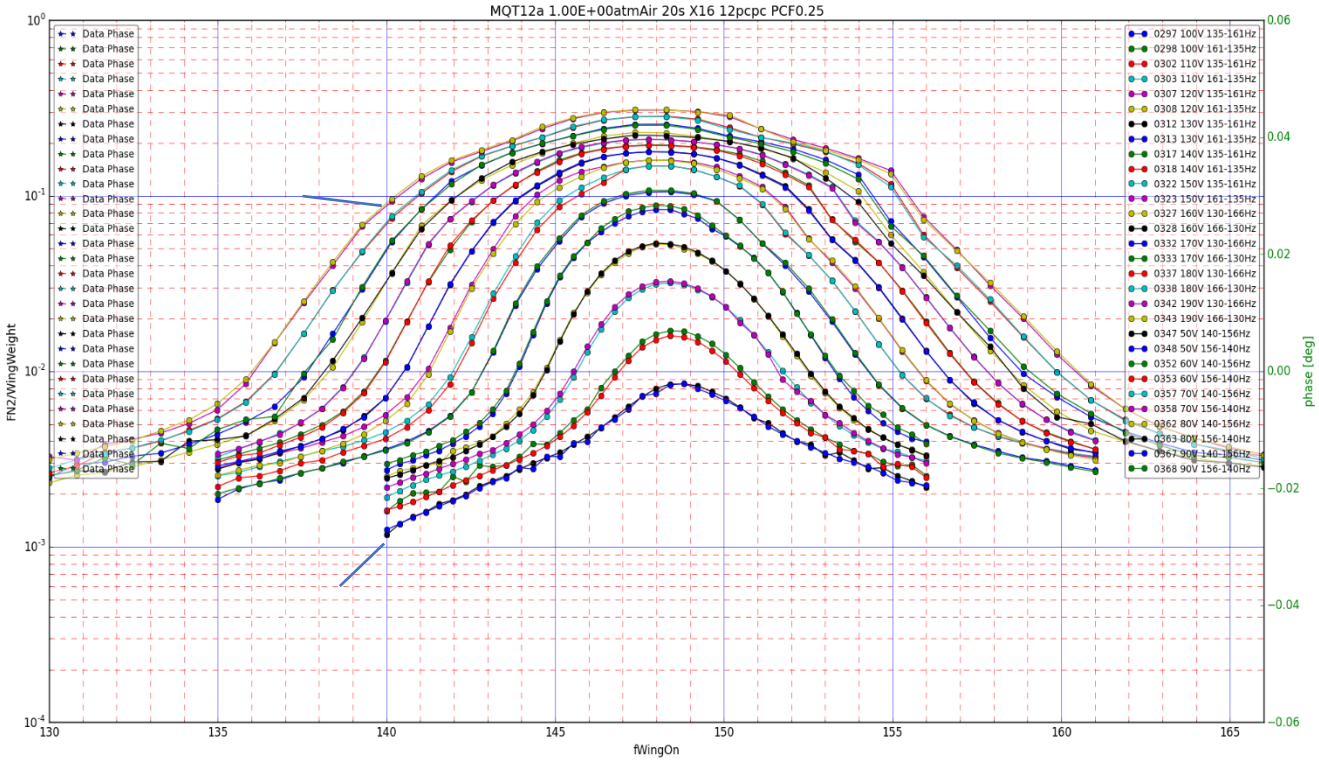


Figure 54: Force/wing weight vs. wing frequency [Hz] for sweeps done at varying voltages in air at STP.

In Figure 54, we see that maximum force increases with increased voltage until about 100V where the increase in force/applied voltage is less. Higher voltages were not applied due to risk of arcing and because the force appeared to approach a limit.

Unsure of the reason for the apparent force saturation, I investigated the frequency response of the flapping amplitude A_r at different voltages to see if perhaps this amplitude is not increasing with voltage. Note that for a linear system voltage should be related to A_r by

$$A_r = D\sqrt{\frac{V}{C_{DP}}} \tag{5.7}$$

where for a given wing flapping at any A_r and Re,

$$D \propto \sqrt{\frac{\rho_{air}}{\rho_{medium}}} \ . \tag{5.8}$$

A plot of A_r vs. frequency at varying voltages is shown in Figure 55. The values were found by visually measuring the amplitude of the flapping 37° wing pair from pictures like Figure 53.

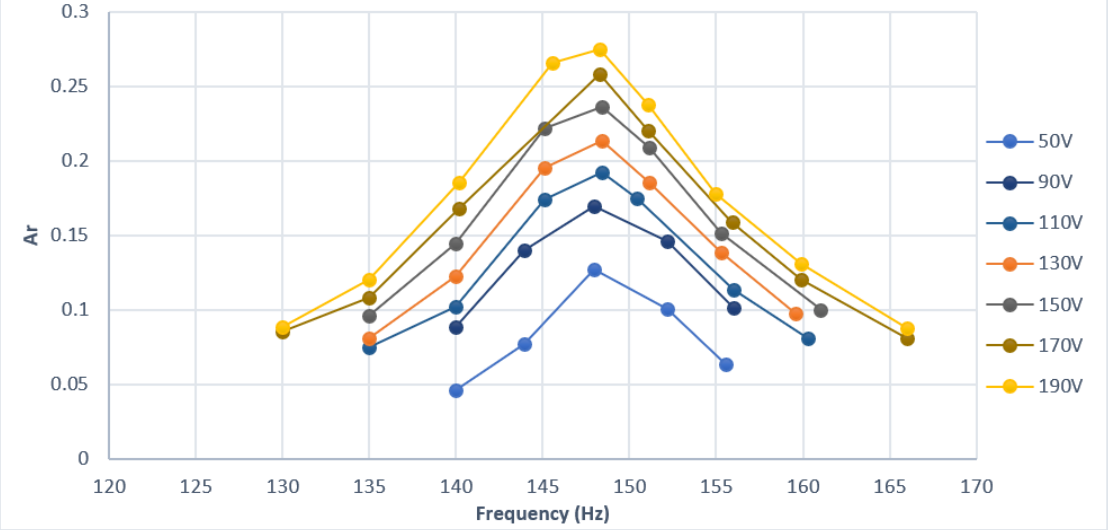


Figure 55: Frequency response of the wing flapping amplitude Ar.

Fitting a cubic spline weighted near the peaks, I approximated the resonant frequency and interpolated the corresponding A_r. A plot of the wings' A_r vs. voltage at maximum A_r (i.e., at resonance) in different mediums is shown below. All curves were found to be well approximated by the function $Ar = aV^{0.5}$ which is consistent with equation (5.7) for constant C_{DP} (note that C_{DP} is indeed roughly constant in this regime).

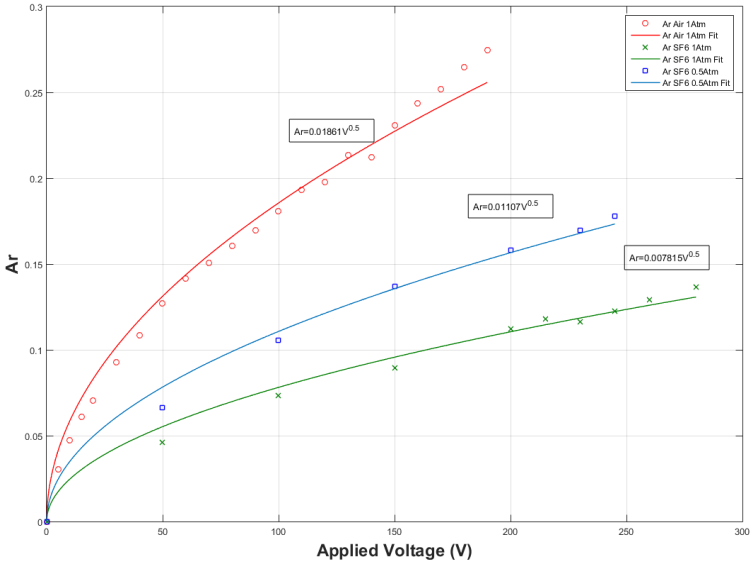


Figure 56: Flapping amplitude (Ar) at resonance Vs. applied voltage in: air at STP, SF_6 at 1 Atm, and SF_6 at 0.5Atm.

Following these tests, the torsion bar chamber was brought to a vacuum. The reason was to verify that the energy loss was mostly to the surrounding fluid (rather than internal mechanical losses), and to reach higher A_r at low voltages. Indeed, the Q in vacuum was roughly 20 times larger than in air and at voltages above 70V, the wings seemed to enter a highly non-linear regime where the wings' A_r rapidly increased to above 1.01 as seen in Figure 57. Furthermore, the A_r exhibited hysteresis as the resonant frequency depended on whether resonance was reached by increasing or decreasing frequency. The possibility of non-linear resonant effects was mentioned by Minnick in [3] but never actually seen since the wings did not retain shape under vacuum due to the SU-8 layer outgassing.

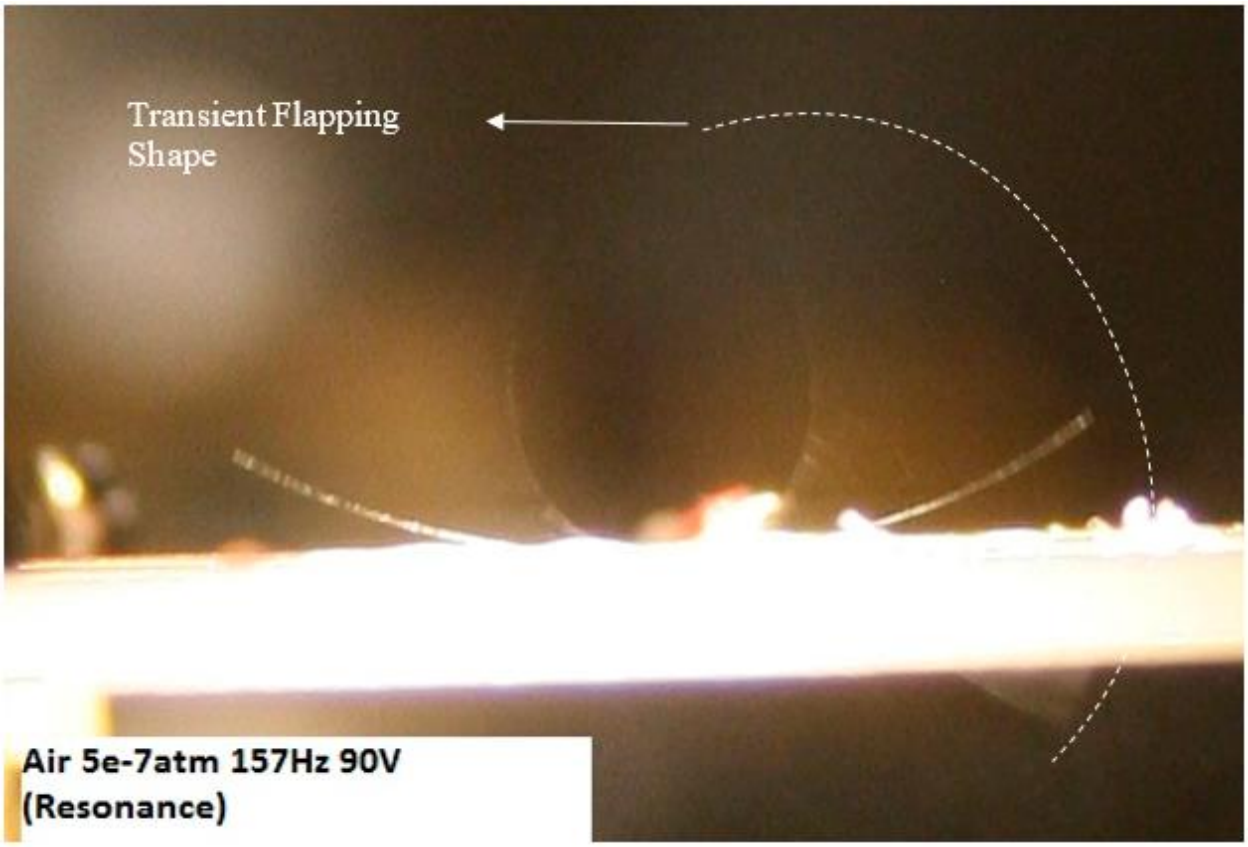


Figure 57: Wings at resonance show very large A_r when wing actuated at 90V, 157Hz in vacuum.

5.1.8 Discussion

Originally, the force saturation with increasing voltage seen in Figure 54 was discouraging. Based on the trend, it appeared that without going to voltages above 700V, these wings would never achieve the minimum force to lift themselves. Initial guesses were that non-linear fluidic damping was causing the wings to flap at lower A_r, but Figure 55 and Figure 56 show remarkable consistency in theory and that A_r was increasing expectedly with applied voltage. In fact, if we look at the constants for the best fits in Figure 56, we can see that their relative values between gas media are also consistent. For example, looking at equation (5.7), we can see that

$$A_r = a\sqrt{V} \tag{5.9}$$

where

$$a = \frac{D}{\sqrt{C_{DP}}} \tag{5.10}$$

Assuming that C_{DP} varies similarly in these different gasses, using (5.10) and (5.8) we get

$$a_{medium} = a_{air}\sqrt{\frac{\rho_{air}}{\rho_{medium}}} \tag{5.11}$$

The fit values of a_{medium} and the expected values based on equation (5.11) are shown in the following table:

Table 8: Comparison of values of D/$\sqrt{C_{DP}}$ based off fit to experimental values and relative densities.

Medium	a_{medium} **(fit to experiment)**	a_{medium} **(expected from density change)**
Air	0.019+/-0.001	-
SF6 @ 1ATM	0.0078+/-0.0004	0.0086 +/-0.0003
SF6 @ 0.5ATM	0.0111 +/-0.008	0.0122+/-0.0005

Based on Table 8, we have further verification that the wings are behaving expectedly in terms of A_r across different media.

Since the wings are working expectedly and considering that the force is proportional to Ar^2 (from equations (1.7)-(1.10)) , then the maximum force should have increased linearly with voltage. The frequency, μ, and ρ at a given medium also stayed the same so neither of these factors explain why the force data only marginally increased at higher voltages.

The exact reason becomes more apparent when noting where on the C_{DN} plot the air tests were done. Figure 58 shows this along with the SF_6 tests where the points from left to right respectively correspond to those of Figure 56. It turns out that as the voltage increased at each sweep along with A_r, the wings passed along the 'island' region talked about earlier in Chapter 3. Coincidentally, the higher voltage sweeps occurred where C_{DN} decreases with increasing A_r and Re, just above the island region. Specifically, we can see that from voltages 5-100V (first 12 points from the left for Air), the C_{DN} increases from 0.04 to 0.10, and then decreases past 120V.

Looking at Figure 54, we see that this is around where the increase in voltage showed minimal increase in force. Thus, the solution was simply to continue to increase the applied voltage to increase A_r which would place the wing much past the island region and into that of increasing C_{DN} where larger force/applied voltage is expected (albeit this might have been risky territory based on the possibility of arcing at higher voltages, as we shall see shortly). Note that this effect was not detected in the previous discussion because the C_{DP} in this region is relatively constant and has no such 'island region'.

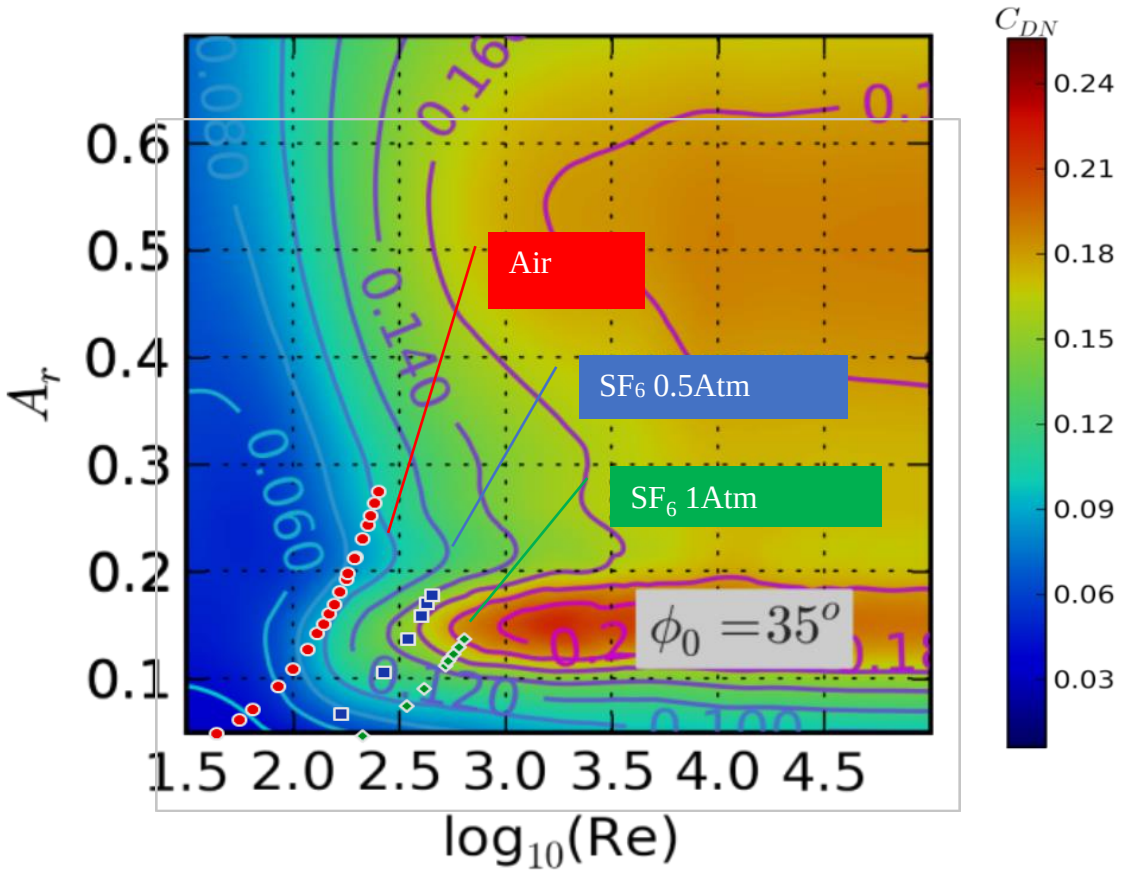


Figure 58: Overlay of Re and A_r at which the wings were operating that resulted in the force peaks. Each point corresponds to the applied voltage, where points plotted from left to right correspond to increasing applied voltage.

The vacuum results also show that the wings are durable against high stresses at large A_r, including the 0.8mm-wide bridge the supports both wings. Another point in favor of their durability is that throughout changing gasses and pressures, the wing ϕ never changed nor did the

wings deform like those made from SU-8. This justifies the use of the Molybdenum layers as a superior material to SU-8 in terms of stability to the environment. Ultimately, after about 0.8 billion flapping cycles (~1500 hours of flapping), a programing error resulted in a high applied voltage of an estimated 500-1000V on the wings. Afterwards, they became inoperable due to a short being created at the contact sides and a connection loss near the bridge. Looking at Figure 59, we can see that arcing likely occurred near the tips of the wings and the bridge, where the gold melted off due to the high heat from the plasma. This led to the modified arc-resistant wing design attempted for the $A_sR=1.3$ wings described in the next subsection.

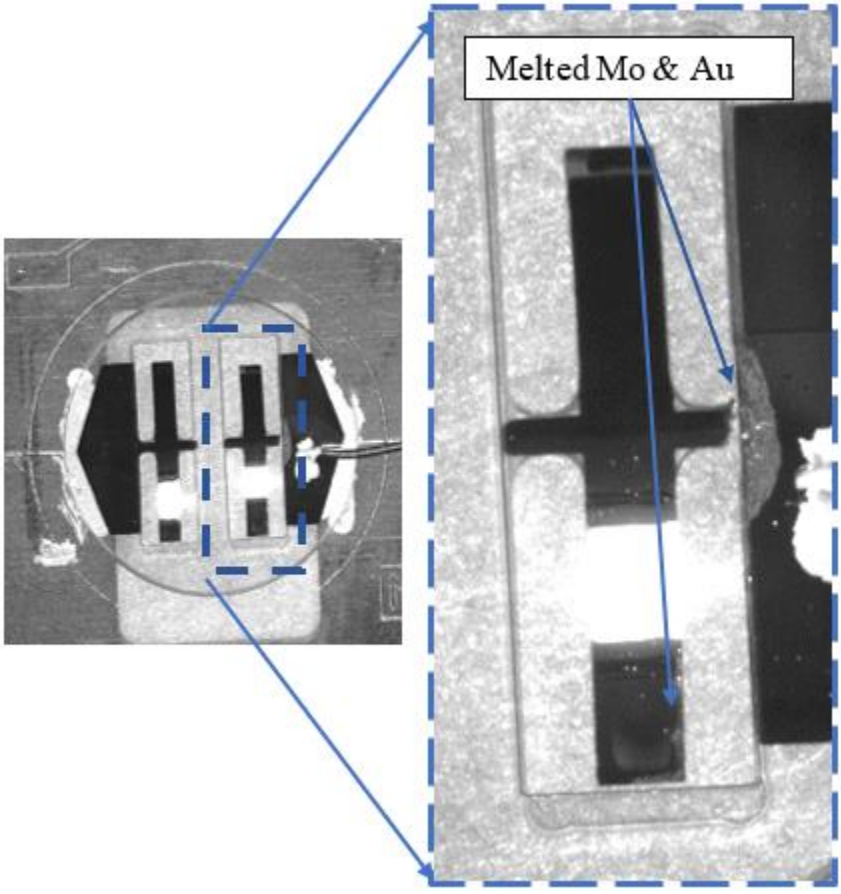


Figure 59: Post-mortem of the Molybdenum-Quartz wings after receiving a high applied voltage that caused arcing in the regions shown. Aside from this failure, the wings showed no difference in all other characteristic throughout their long hours of use.

Molybdenum-Quartz Wings Test 2: A_sR=1.3

5.1.9 Results & Discussion

Following the mask designs for the wider wings, initial attempts in fabrication failed due to changes in the sputtering machine. Nevertheless, the failures do show important results that should be considered in the fabrication process. In this section, I will present the failed fabrication attempts, discuss the reasons and make a prediction of the expected characteristics of such a wing once fabricated successfully.

About 2 years following the fabrication of the A_sR=8 wings, the DC Magneton sputtering machine had its gas flow sensors replaced and the vacuum pump was working slower (this was quantitatively confirmed after characterizing steady-state chamber pressures with known gas flows). Another issue included bad calibration between the two chamber pressure gauges (used separately for high and low pressures) at the point where they switched (around 7e-3mTor); coincidentally, this was the steady-state working pressure region for sputtering that was meant to develop the highest intrinsic tensile stress. Due to the sensitivity of the intrinsic film stress to the working pressure, these issues made it difficult to recreate the previous sputtering environment.

Using the new pump speed characterization, I estimated what changes had to be made to the Argon flow rate in order to work at the same sputtering pressure as in the old tests. The new rates used were 8, 9, and 11sccm. Initially, a 0.5um sputtering test was done for each rate using a Molybdenum target and a Silicon substrate. Using a profilometer to measure the curvature of the substrate and corresponding stress from the Stoney equation [34], the 9sccm results showed the highest tensile stress of about 850MPa+/-100MPa. This stress is more than twice as large as largest of the older Mo-Si tests for similar thicknesses and an order of magnitude larger than the intrinsic stress that caused the curvature in the A_sR=8 wings from the previous section. Proceeding with the normal fabrication steps as explained previously, 2μm of Mo and 0.15μm of Au were sputtered on a 9um quartz substrate without opening the sputtering chamber in between. A surprising result is shown in Figure 60 where we can see the shattered sample and mask in a sample holder. Since the middle of the sample holder was sputtered gold, this suggests that the substrate broke during deposition, and not in response to me opening the chamber.

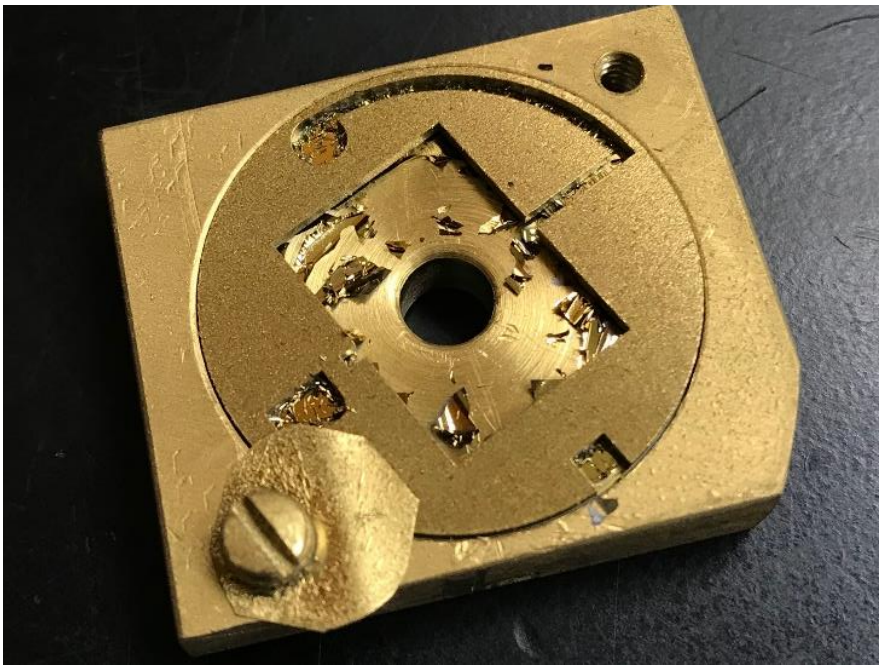


Figure 60: Shattered sample after Mo and Au deposition. The rest of the sample is covered by the shadow mask and both are held in place by a sample holder.

To examine the exact reason for this delamination (which to date had never occurred before), I reran the tests on bare Si with the same parameters but with 2um of Mo thickness instead of 0.5um, and no Au. Upon backfilling the chamber with room temperature nitrogen and opening it, I noticed rapid delamination of the Mo films, bursting into cylindrical flakes as seen in Figure 61 with radii of curvature of about 1.3+-0.3mm.

Figure 61: Result of sputtering Molybdenum on a Silicon wafer using the new 'optimal' sputtering parameters. The delamination continued to occur after taking the sample out of the chamber.

The delamination described here is a common occurrence in tensile thin film deposition where the intrinsic stress of the deposited film surpasses the "fracture resistance" at the material-substrate interface [35], [36]. The curling then naturally occurs if a stress gradient is present throughout the thickness of the film [37]. Similar flakes are shown by the authors in [38] that sputtered titanium nitride thin films with high intrinsic stresses built in.

Unfortunately, due to the delamination, I could not use the profilometer to test the built-in stress but a common formula for deflections for thin cantilevers with a residual stress gradient can be

found in [39]–[41]. In the limit where circular curvature can be approximated by quadratic curvature and assuming that the stress gradient is constant, the mean intrinsic bending stress calculated from the radius of curvature of the flakes is approximately 1.05+/- 0.4GPa. This is fairly larger than the highest measured stress of 800MPa done a couple years earlier, but lower than the 1.7GPa stress found by other authors where flaking did not occur [42], [43].

Repeating these tests and changing the Argon flow rates by ±1sccm still showed similar flaking upon exposure to room temperature nitrogen and air, but with stresses from gradients measuring only 0.5±0.2MPa. This, along with the spontaneous flaking when exposed to lower temperature environments suggests that the delamination is due to high intrinsic stresses, stress gradients and thermal stress from thermal strain mismatch between Mo and Si when cooled down after being sputtering at high temperatures. It is also known that the adhesion stability of thin films tends to decrease with thickness [44]; thus, the solution to this issue is simply to deposit lower thickness films using the same parameters, or work at much different flow rates to fabricate lower built-in stress.

By using the formulas derived by Minnick [3] based off Timoshenko's beam curvature theory for a multimorph stack structure, we can calculate the expected bending angle based on an expected layer stress (see Chapter 2 in [3] for details on beam statics and resonance). At 0.5um of deposited Mo, we already know that we can achieve around 850MPa of tensile stress without delamination; thus, by assuming these parameters for the tensile Mo layer of the stack structure shown in Table 6 and all other thickness/stresses equal, the fabricated wings would have a ϕ of 131.6°! Of course, this design would quickly become impractical at high A_r due to the two wings colliding; that being said, the benefit in being able to deposit a layer with so much intrinsic stress is that it allows us to increase the thicknesses of certain layers to reach higher resonant frequencies and hence, higher Re. For example, by using this tensile Mo layer and increasing the quartz thickness from 6.8um to 13.6um (by two times), the expected ϕ drops to 47° (close to the 37° in the old wings), while the resonant frequency ω increases by 1.64 times. Ignoring the change in C_{DN} with ϕ and considering constant A_r, this alone would increase the force by

$$\frac{F}{F_0} = \frac{\mathrm{Re}^2}{\mathrm{Re}_0^2} = \frac{\omega^2}{\omega_0^2} 1.6385^2 = 2.68 \tag{5.12}$$

If we now consider that the wing's A_sR=1.3, using equation (5.5), the adjusted C_{DN} would be

$$C_{DN} = C_{DN_0}\left(\frac{1.3}{8}\right)^{-3/2} \approx C_{DN_0}15.3 \tag{5.13}$$

Combining the two factors from (5.12) and (5.13) means that the expected force should be about 41 times larger for A_sR=1.3 than the A_sR=8 wings from the previous section. Now considering that these wings are about 6.2 times wider and thicker due to the new Quartz layer, they are expected to weigh about 8.6 times more. Thus, the force/weight is expected to be 4.7 times larger than the A_sR=8 wings. Given that the same maximum A_r=0.27 could be achieved as done for A_sR=8 wings (shown in Figure 58), this would mean a lifting force of at least 1.5 times the wings' weight – enough to lift themselves and a load of about 5mg. It is important to note that the forces from Figure 54 do not take into account interference from the PCB on which the wing is mounted. Minnick notes that simulations have shown that the interference of the wing flapping near the PCB can result in a decrease in measured force as much as 70%, which leads to two important possibilities:

1. The successfully fabricated wings of A_sR=8 may have actually achieved sufficient thrust to lift 1.1 times their weight instead of just 0.34 as seen in Figure 54– a milestone so far for these wing types.
2. The larger wings of A_sR=1.3 might produce a force 5 times their weight - enough to lift an extra 25mg which is more than twice the maximum lifting capacity of an average house fly [45]!

These are very promising results and suggest that this wing design might already be good enough to be implemented for tethered flight.

6 Conclusion

At the end of Minnick's work on using curved, resonating, piezoelectric cantilevers as a method of thrust, he found that such devices could be fabricated and do give a net force, yet little was still understood about where the force comes from and if other wing shapes may work better. Poor materials also led to inconsistent results and impractical wings which failed to produce high forces. I began this thesis with the intent to carry on this work by filling in these gaps in knowledge and making wings that are suitable to be incorporated in MAVs.

I found that the wing forces are generated from unsteady effects that strongly relate to the induced lagged flow during the wing's oscillation. By carefully examining the transient forces on the wing, I found that the net positive thrust comes from timely interactions with the unsteady flow which depends heavily on the transient flapping shape of the wing. After describing the force phenomena for a wing of an aspect ratio of 8, I investigated the changes in force and power for wings of different aspect ratios, to see if there are more optimal shapes that can considered when building functional NAV wings. A surprisingly consistent relationship was discovered which I incorporated into the force and power drag equations, allowing for a much broader use of the results from previous work. I updated the fabrication technique to create more robust wings and managed to create the first functioning Molybdenum-Quartz wings of this type. These wings proved to be better than earlier models and generated the highest forces to date for resonant curved cantilevers. Simulations showed that lower aspect ratios would work even better,

and although my attempts to fabricate such wings have so far been unsuccessful, I showed what changes need to be made in future attempts for them to work.

The next steps would be to fabricate and test the wider wings while continuing simulations in different parameter spaces. The aspect ratio dependence at other wing parameters should be confirmed and it is still unclear why the surprisingly consistent relationship of aspect ratio breaks down at lower values. In any case, even for the small range of parameters simulated, the relationship is promising, and lower aspect ratios should be explored physically where the fabrication method described in Chapter 5 is a good starting point. There are also more ways to optimize these wings than just changing the aspect ratio; for example, one might extend the simple 1-degree of freedom flapping motion of these wings to 2-degree designs (to mimic the more efficient insect flight patterns) by changing the layer properties in the x direction through appropriately etching holes or varying thicknesses. The planar fabrication method also allows for one to deviate from the rectangular-based wings and try other shapes that may capture and expel fluid in a more efficient way.

Although there is still much to learn about the potential of resonant curved cantilever wings, the results and furthered groundwork presented in this thesis show that the original wings can indeed be significantly improved and have serious potential as NAV flight mechanisms. We now know how unusually they work and even though they do not use the same flight mechanics as nature's fliers, we are one step closer to creating our own tiny flying machines and bringing these wings to life.

7 BIBLIOGRAPHY

[1] J. Huber and J. Noyes, “A new genus and species of fairyfly, Tinkerbella nana (Hymenoptera, Mymaridae), with comments on its sister genus Kikiki, and discussion on small size limits in arthropods,” *J. Hymenopt. Res.*, vol. 32, pp. 17–44, Apr. 2013.

[2] T. Weis-Fogh, “Flapping Flight and Power in Birds and Insects, Conventional and Novel Mechanisms,” in *Swimming and Flying in Nature: Volume 2*, T. Y.-T. Wu, C. J. Brokaw, and C. Brennen, Eds. Boston, MA: Springer US, 1975, pp. 729–762.

[3] M. D. Minnick, “Resonant Curved Piezoelectric Cantilever Fluid Diode Wings for Mass-Producible Flying Microrobots,” Thesis, 2014.

[4] R. Å. Norberg, “Hovering Flight of the Dragonfly Aeschna Juncea L., Kinematics and Aerodynamics,” in *Swimming and Flying in Nature: Volume 2*, T. Y.-T. Wu, C. J. Brokaw, and C. Brennen, Eds. Boston, MA: Springer US, 1975, pp. 763–781.

[5] C. P. Ellington, “The Aerodynamics of Flapping Animal Flight,” *Integr. Comp. Biol.*, vol. 24, no. 1, pp. 95–105, Feb. 1984.

[6] T. Weis-Fogh, “Quick Estimates of Flight Fitness in Hovering Animals, Including Novel Mechanisms for Lift Production,” *J. Exp. Biol.*, vol. 59, no. 1, pp. 169–230, Aug. 1973.

[7] C. P. Ellington, “The aerodynamics of hovering insect flight. IV. Aerodynamic mechanisms,” *Phil Trans R Soc Lond B*, vol. 305, no. 1122, pp. 79–113, Feb. 1984.

[8] P. C. Trizila, “Aerodynamics of Low Reynolds Number Rigid Flapping Wing Under Hover and Freestream Conditions,” p. 183, 2011.

[9] A. G. Bennett, R. C. Obye, and P. M. Jeglum, “Ornithopter Aerodynamic Experiments,” in *Swimming and Flying in Nature: Volume 2*, T. Y.-T. Wu, C. J. Brokaw, and C. Brennen, Eds. Boston, MA: Springer US, 1975, pp. 985–1000.

[10] C.-S. Lin, C. Hwu, and W.-B. Young, "The thrust and lift of an ornithopter's membrane wings with simple flapping motion," *Aerosp. Sci. Technol.*, vol. 10, no. 2, pp. 111–119, Mar. 2006.

[11] J.-S. Lee and J.-H. Han, "Indoor Flight Testing and Controller Design of Bioinspired Ornithopter," in *Frontiers of Intelligent Autonomous Systems*, vol. 466, S. Lee, K.-J. Yoon, and J. Lee, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 277–286.

[12] M. Karpelson, G.-Y. Wei, and R. J. Wood, "A review of actuation and power electronics options for flapping-wing robotic insects," in *2008 IEEE International Conference on Robotics and Automation*, 2008, pp. 779–786.

[13] R. Wood, R. Nagpal, and G.-Y. Wei, "Flight of the Robobees," *Sci. Am.*, vol. 308, no. 3, pp. 60–65, Mar. 2013.

[14] K. Naeli and O. Brand, "Dimensional considerations in achieving large quality factors for resonant silicon cantilevers in air," *J. Appl. Phys.*, vol. 105, no. 1, p. 014908, Jan. 2009.

[15] S. Kirstein, M. Mertesdorf, and M. Schönhoff, "The influence of a viscous fluid on the vibration dynamics of scanning near-field optical microscopy fiber probes and atomic force microscopy cantilevers," *J. Appl. Phys.*, vol. 84, no. 4, pp. 1782–1790, Aug. 1998.

[16] D. B. Ingham, B. Yan, and B. R. Morton, "The fluid flow induced by large-amplitude oscillations of a cascade," *Comput. Fluids*, vol. 21, no. 3, pp. 305–321, Jul. 1992.

[17] E. Woolam, "Drag coefficients for flat square plates oscillating normal to their planes in air," *NASA CR-66544 March*, 1968.

[18] F. Manar and A. R. Jones, "Transient Response of a Single Degree-of-Freedom Wing at High Angle-of-Attack," *AIAA J.*, vol. 55, no. 11, pp. 3681–3692, 2017.

[19] D. Floreano, Ed., *Flying insects and robots*. Heidelberg ; New York: Springer, 2009.

[20] P. Trizila, C. Kang, M. Visbal, and W. Shyy, "A Surrogate Model Approach in 2-D Versus 3-D Flapping Wing Aerodynamic Analysis," in *12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Victoria, British Columbia, Canada, 2008.

[21] T. Theodorsen, "General Theory of Aerodynamic Instability and the Mechanism of Flutter," Jan. 1949.

[22] R. L. Bisplinghoff, H. Ashley, and R. L. Halfman, *Aeroelasticity*, Dover ed. New York: Dover Publications, 1996.

[23] C. J. Pennycuick, *Modelling the Flying Bird*. Elsevier, 2008.

[24] J. Young, J. C. S. Lai, and C. Germain, "Simulation and Parameter Variation of Flapping-Wing Motion Based on Dragonfly Hovering," *AIAA J.*, vol. 46, no. 4, pp. 918–924, Apr. 2008.

[25] S. P. Sane, "The aerodynamics of insect flight," *J. Exp. Biol.*, vol. 206, no. 23, pp. 4191–4208, Dec. 2003.

[26] M. Sun and J. Tang, "Unsteady aerodynamic force generation by a model fruit fly wing in flapping motion," *J. Exp. Biol.*, vol. 205, no. 1, pp. 55–70, Jan. 2002.

[27] "Smithsonian miscellaneous collections," p. 89, 1962.

[28] M. A. Lieberman and A. J. Lichtenberg, *Principles of Plasma Discharges and Materials Processing*, 2 edition. Hoboken, N.J: Wiley-Interscience, 2005.

[29] M. D. Minnick, G. A. Devenyi, and R. N. Kleiman, "Optimum reactive ion etching of x-cut quartz using SF6 and Ar," *J. Micromechanics Microengineering*, vol. 23, p. 117002, Nov. 2013.

[30] C. Li, T. Chen, J. Si, F. Chen, X. Shi, and X. Hou, "Fabrication of three-dimensional microchannels inside silicon using a femtosecond laser," *J. Micromechanics Microengineering*, vol. 19, no. 12, p. 125007, 2009.

[31] M. Andritschky and V. Teixeira, "Residual stress and adhesion of molybdenum coatings produced by magnetron sputtering," *Vacuum*, vol. 43, no. 5–7, pp. 455–458, May 1992.

[32] K. Kuwahara, T. Sumomogi, and M. Kondo, "Internal stress in aluminium oxide, titanium carbide and copper films obtained by planar magnetron sputtering," *Thin Solid Films*, vol. 78, no. 1, pp. 41–48, Mar. 1981.

[33] J. Ray and T. K. Chaudhuri, "Growth and structural properties of Molybdenum thin films deposited by DC sputtering," *J. Optoelectron. Adv. Mater.*, vol. 17, no. 5–6, pp. 634–639, 2015.

[34] G. G. Stoney, "The tension of metallic films deposited by electrolysis," *Proc R Soc Lond A*, vol. 82, no. 553, pp. 172–175, May 1909.

[35] M. D. Thouless and H. M. Jensen, "The effect of residual stresses on adhesion measurements," *J. Adhes. Sci. Technol.*, vol. 8, no. 6, pp. 579–586, Jan. 1994.

[36] D. M. Mattox and R. E. Cuthrell, "Residual Stress, Fracture, and Adhesion in Sputter-Deposited Molybdenum Films," *MRS Proc.*, vol. 119, Jan. 1988.

[37] K. L. Mittal and W.-S. Lei, *Laser Technology: Applications in Adhesion and Related Areas*. John Wiley & Sons, 2018.

[38] R. Machunze and G. C. A. M. Janssen, "Stress gradients in titanium nitride thin films," *Surf. Coat. Technol.*, vol. 203, no. 5–7, pp. 550–553, Dec. 2008.

[39] F. Ericson, S. Greek, J. Söderkvist, and J.-Å. Schweitz, "High-sensitivity surface micromachined structures for internal stress and stress gradient evaluation," *J. Micromechanics Microengineering*, vol. 7, no. 1, pp. 30–36, Mar. 1997.

[40] C.-W. Baek, Y.-K. Kim, Y. Ahn, and Y.-H. Kim, "Measurement of the mechanical properties of electroplated gold thin films using micromachined beam structures," *Sens. Actuators Phys.*, vol. 117, no. 1, pp. 17–27, Jan. 2005.

[41] Y. Kilinc, U. Unal, and B. E. Alaca, "Residual stress gradients in electroplated nickel thin films," *Microelectron. Eng.*, vol. 134, pp. 60–67, Feb. 2015.

[42] Y. . Shen, “Effect of deposition conditions on mechanical stresses and microstructure of sputter-deposited molybdenum and reactively sputter-deposited molybdenum nitride films,” *Mater. Sci. Eng. A*, vol. 359, no. 1–2, pp. 158–167, Oct. 2003.

[43] T. J. Vink, M. a. J. Somers, J. L. C. Daams, and A. G. Dirks, “Stress, strain, and microstructure of sputter-deposited Mo thin films,” *J. Appl. Phys.*, vol. 70, no. 8, pp. 4301–4308, Oct. 1991.

[44] J. W. Hutchinson, “Stresses and Failure Modes in Thin Films and Multilayers 1,” 1996.

[45] G. L. Cox, F. C. Lewis, and E. E. Glynn, “The Number and Varieties of Bacteria carried by the Common House-fly in Sanitary and Insanitary City Areas,” *Epidemiol. Infect.*, vol. 12, no. 3, pp. 290–319, Oct. 1912.

8 APPENDICES

Measuring Substrate Thicknesses with FTIR

Here I describe how one can accurately measure the thickness of a thin translucent sample (quartz, for example) and provide the MATLAB code at the end.

Using the transmission mode *or* reflection in the FTIR, one should see interference patterns on the resulting spectrum due to the interference between the incident ray passing through the sample, and the secondary ray that reflects inside the sample, and comes out at the same point.

This is illustrated in the following diagram:

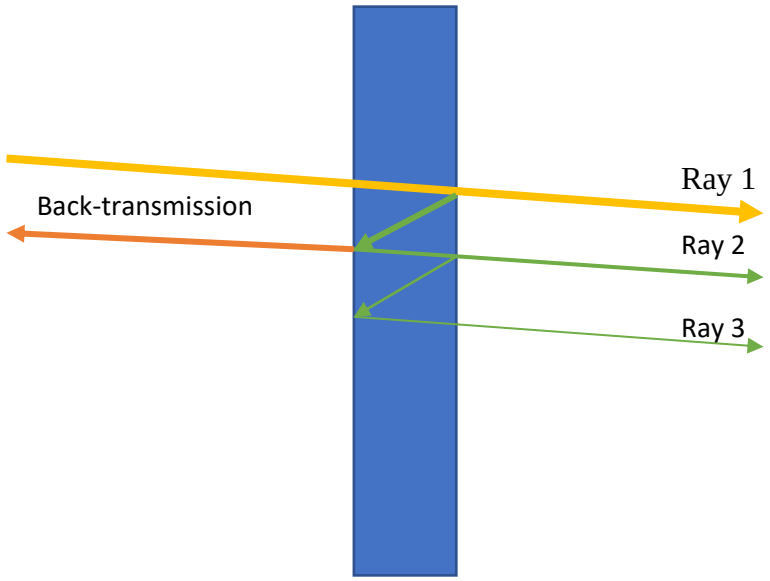


Figure 62: Diagram showing the path length of the incident light as it passes through and is reflected internally in the sample, before leaving again. They yellow ray represents the original incident light, while the green rays represent the internally reflected rays following the path that leads them out of the sample in the same direction. Note that in the actual setup, the Rays are at normal incidence, but are off-axis here for visual clarity; for this reason, the effects of refractance are not shown.

The internal reflecting pattern continues within the crystal with further attenuation, and more rays reflect and transmit through the crystal at each boundary of the incident ray. For the transmission mode of the FTIR, we do not consider the back-transmission of the ray (shown in read) as that component is not well captured. We also can ignore ray paths 3 and higher as they will be more attenuated (as shown by the array thickness) and will have perpetually smaller effects on the final interference pattern. For the following calculations, we consider exiting rays 1 and 2.

Two coincident, coherent rays at a point can be shown to have irradiance given by

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos(k(s_2 - s_1) + \Phi_2 - \Phi_1) \qquad \text{(A1)}$$

Where I is the irradiance, k is the wavenumber $\frac{2\pi}{\lambda}$, λ is the wavelength, s is the beam path length, Φ is the initial phase, and the indices correspond one of the two interfering waves. Since initially, there was only one beam, the initial phases are equal and $\Phi_2 - \Phi_1 = 0$. The difference in beam path lengths is simply the added path length of Ray 2 within the sample, $2N(\lambda)t$, where $N(\lambda)$ is the wave-dependent index of refraction and t is the thickness of the sample. Also note that this works under the assumption of perfect transparent material (i.e., no complex index of refraction), and only in transmission mode – here, we don't have to account for 180 degree phase shift in the calculation due to reflections because n1>n2, but this can be included in the final formula by adding the shift in if working in reflection mode (later, in equation 6)**.** Note that that the factor of two is present to account for the ray travelling twice the sample thickness. The equation thus simplifies to

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\left(\frac{2\pi}{\lambda}2N(\lambda)t\right) \qquad \text{A2}$$

The following is a typical FTIR spectrum for a 100um quartz sample.

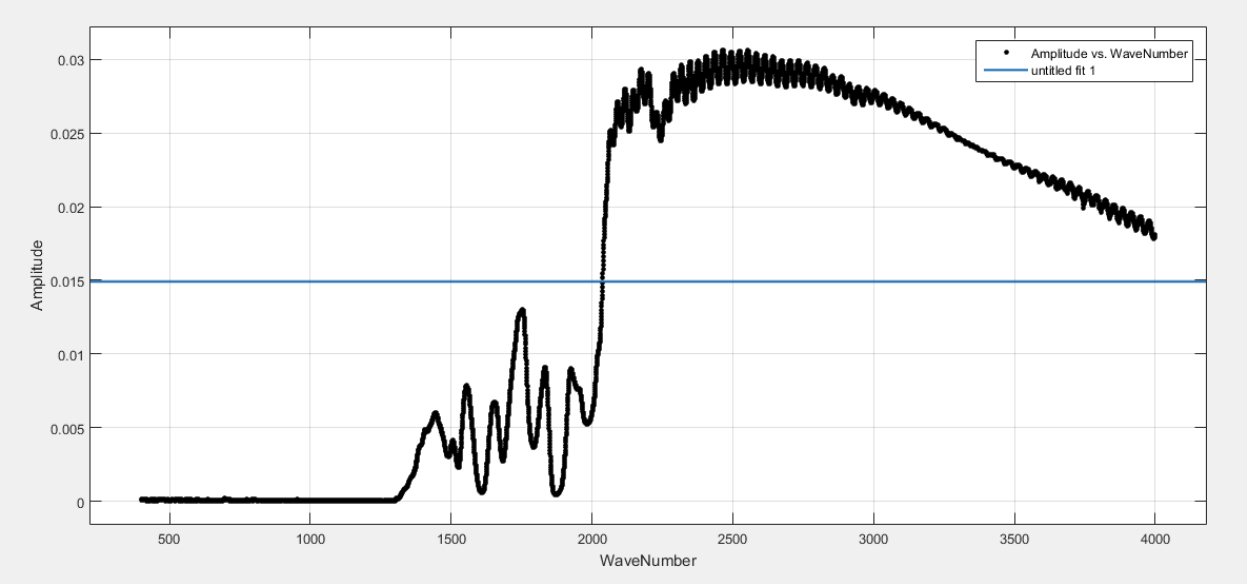


Figure 63: FTIR Spectrum for 100um quartz in transmission mode.

In Figure 63, we can see that small interference patterns are captured following the general transmission curve. We will select data at noticeable patterns for further processing to get 'cleaner' results.

Note that equation A2 does not account for the transmission curve of the light through quartz. To remedy this, we can multiply equation 2 by a transmission curve function $Tc(\lambda)$ that fits the model:

$$A = Tc(\lambda)\left(I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\left(\frac{2\pi}{\lambda}2N(\lambda)t\right)\right)$$
$$= Tc(\lambda)(I_1 + I_2) + Tc(\lambda)2\sqrt{I_1 I_2}\cos\left(\frac{2\pi}{\lambda}2N(\lambda)t\right) \quad \text{A3}$$

Where A is not the amplitude seen in the above figure. Ideally, to perform a Fourier Transform on A, we would only include the cosine term.

Now we can use a Fourier Transform on equation 3, but it would be difficult to determine which part of the result corresponds to the part inside the cosine term and, and to differentiate that part from noise in the transformed domain. Instead, we should eliminate this $Tc(\lambda)$ and any DC (constant) terms.

A simple way to do this is to take a moving average where the number of points selected for the average is larger than the points contained within the length of consecutive interference peaks. If this is done on a curve following equation 2, using a sufficiently large averaging period, the result of the average would be

$$\langle I\rangle = \langle I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\left(\frac{2\pi}{\lambda}2N(\lambda)t\right)\rangle$$
$$= I_1 + I_2 + \langle 2\sqrt{I_1 I_2}\cos\left(\frac{2\pi}{\lambda}2N(\lambda)t\right)\rangle$$
$$= I_1 + I_2 + 2\sqrt{I_1 I_2}\,\langle\cos\left(\frac{2\pi}{\lambda}2N(\lambda)t\right)\rangle$$
$$= I_1 + I_2$$

If the chosen averaging period is relatively small to the fluctuations in the curve of $Tc(\lambda)$, the resultant transmission part of curve is relatively constant, and such a moving average on equation A3 would follow as

$$Tc(\lambda)(I_1 + I_2) \qquad \text{A4}$$

A smoothing spline fit does a similar average but with higher curve continuity. Such a fit is presented in the following figure:

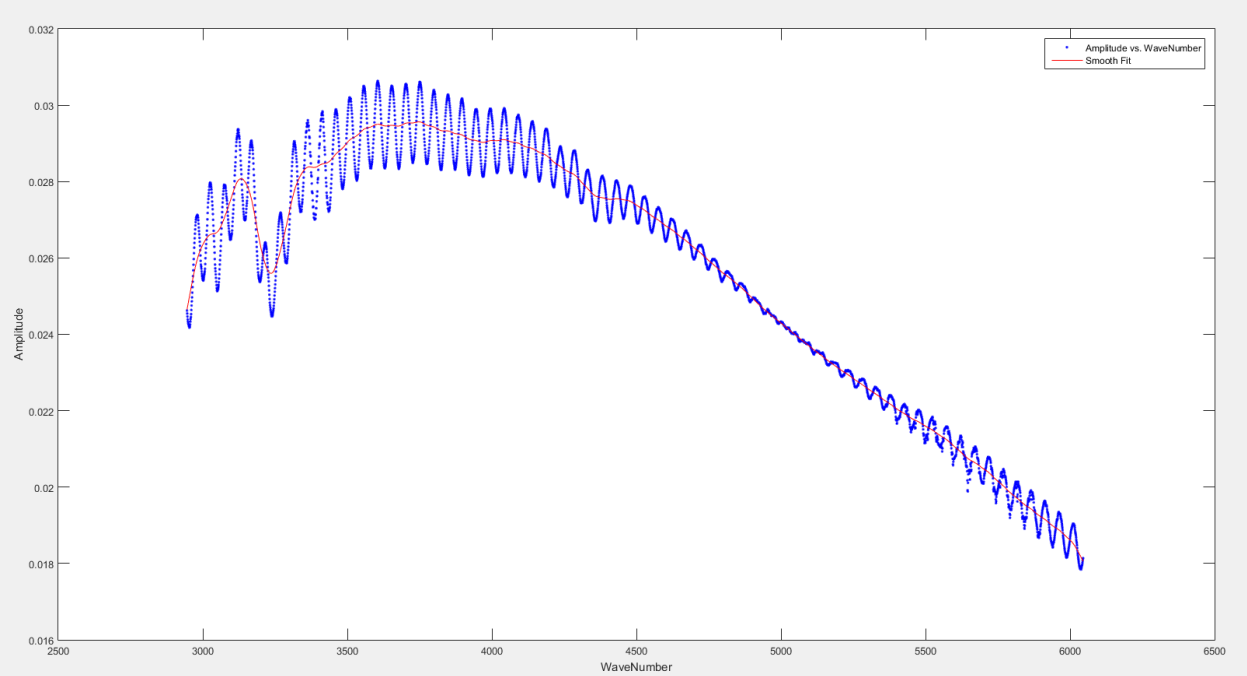


Figure 64: Smoothing spline fit (blue) of the FTIR curve about the range of noticeable peaks.

Now with this fit, we are able to eliminate the first term in equation 3 by simply dividing (3) by (4) which should give

$$A_2 = 1 + \frac{2\sqrt{I_1 I_2}}{(I_1 + I_2)} \cos\left(\frac{2\pi}{\lambda} 2N(\lambda)t\right) \qquad \text{A5}$$

Subtracting 1 from the above gives only the cosine term with a scaled amplitude.

$$A_3 = \frac{2\sqrt{I_1 I_2}}{(I_1 + I_2)} \cos\left(\frac{2\pi}{\lambda} 2N(\lambda)t\right) \qquad \text{A6}$$

To account for the wavelength dependence of the refractive index, we can plot $A_3\ vs\ \frac{N(\lambda)}{\lambda}$. Such a plot of the above curve following these adjustments is shown below.

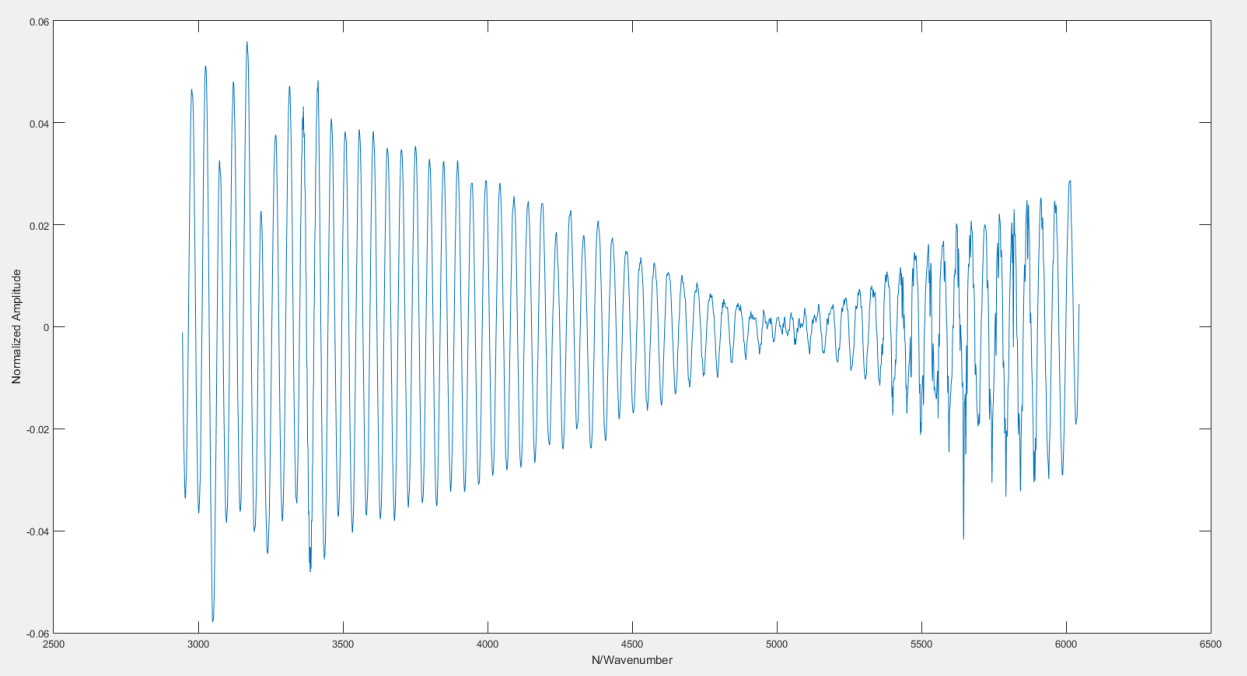


Figure 65: Plot of the normalized interference pattern, or A_3 from equation A6.

Here we can see that the curve is now sinusoidal, but appears to have a beat, which suggests a tertiary interference with another wave. Assuming that this tertiary interference comes from a linear combination of another ray, taking the Fourier transform should still show the main peak which is proportional to t in A6. This is shown in the following figure:

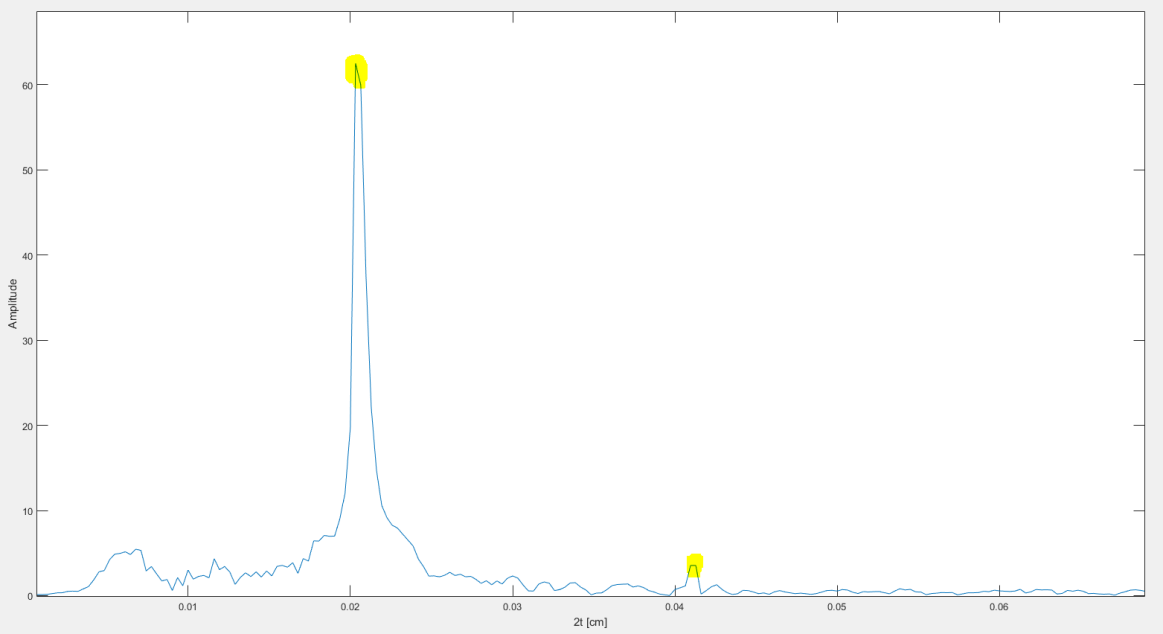


Figure 66: Result of Fast Fourier Transform on the data from the previous plot. Note that the x-axis values are proportional to twice the thickness, t, as expected from equation 6.

Note that two narrow peaks can be seen (highlighted in yellow). One corresponding to Ray 1-Ray 2 interference, and the second at *exactly* twice the dimension on the x axis. This is most likely due to the Ray 1-Ray 3 Interference where the thickness that Ray 3 travels is twice that of Ray 2. Since the peak location is proportional to *t*, we expect this second peak to be proportionally at twice the x-value.

Note quartz is birefringent in the x-cut plane, which is the cut used for the wings in this thesis. The birefringence is relatively small so interference from the ordinary and extraordinary ray would not be noticed on this scale but should be kept in mind that the resultant peak of the FFT is a linear combination of both. The effect is quantified below:

Assuming a linear combination of two interference patterns caused by the two rays, normalized to an amplitude of 1, the resultant linear sum will be a sum of two cosines

$$f(\lambda) = \cos\left(\frac{2\pi}{\lambda} 2N_{ex}(\lambda)t\right) + \cos\left(\frac{2\pi}{\lambda} 2N_{or}(\lambda)t\right) \qquad \text{A7}$$

Where N_{ex} and N_{or} are the indices of refraction for the extraordinary and ordinary rays respectively, as a function of the wavelength. If we now group the λ dependences of the first term such that $x = \frac{N_{ex}(\lambda)}{\lambda}$,

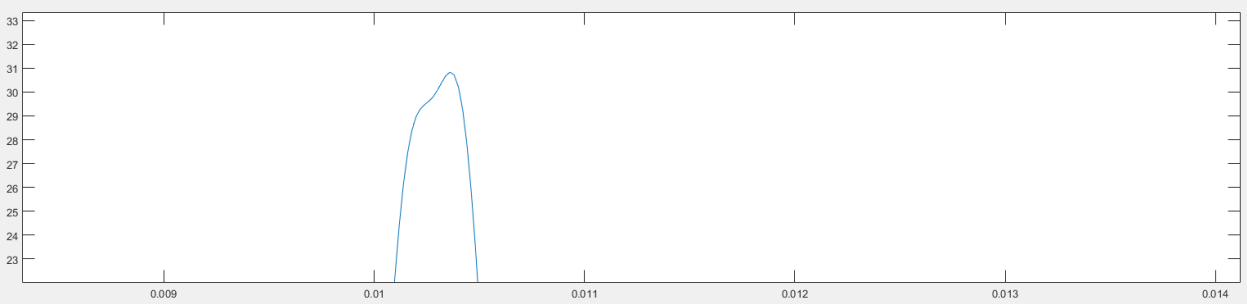
Then the above can be rewritten as

$$f(x) = \cos(4\pi x t) + \cos\left(4\pi \frac{N_{or}(\lambda)}{N_{ex(\lambda)}} t\right) \qquad \text{A8}$$

In general, $N_{ex}(\lambda) \approx 1.009\, N_{or}(\lambda)$ and thus we can see that a Fourier transform will gives us two peaks, one corresponding to the thickness *t* and the other at thickness *t/1.009*.

The refractive index chosen for this process is that of the extraordinary ray, as it consistently transmits slightly better than the ordinary ray and the peak can be assumed to belong to that ray.

This can actually be seen at a close up of the FFT as in the following figure.



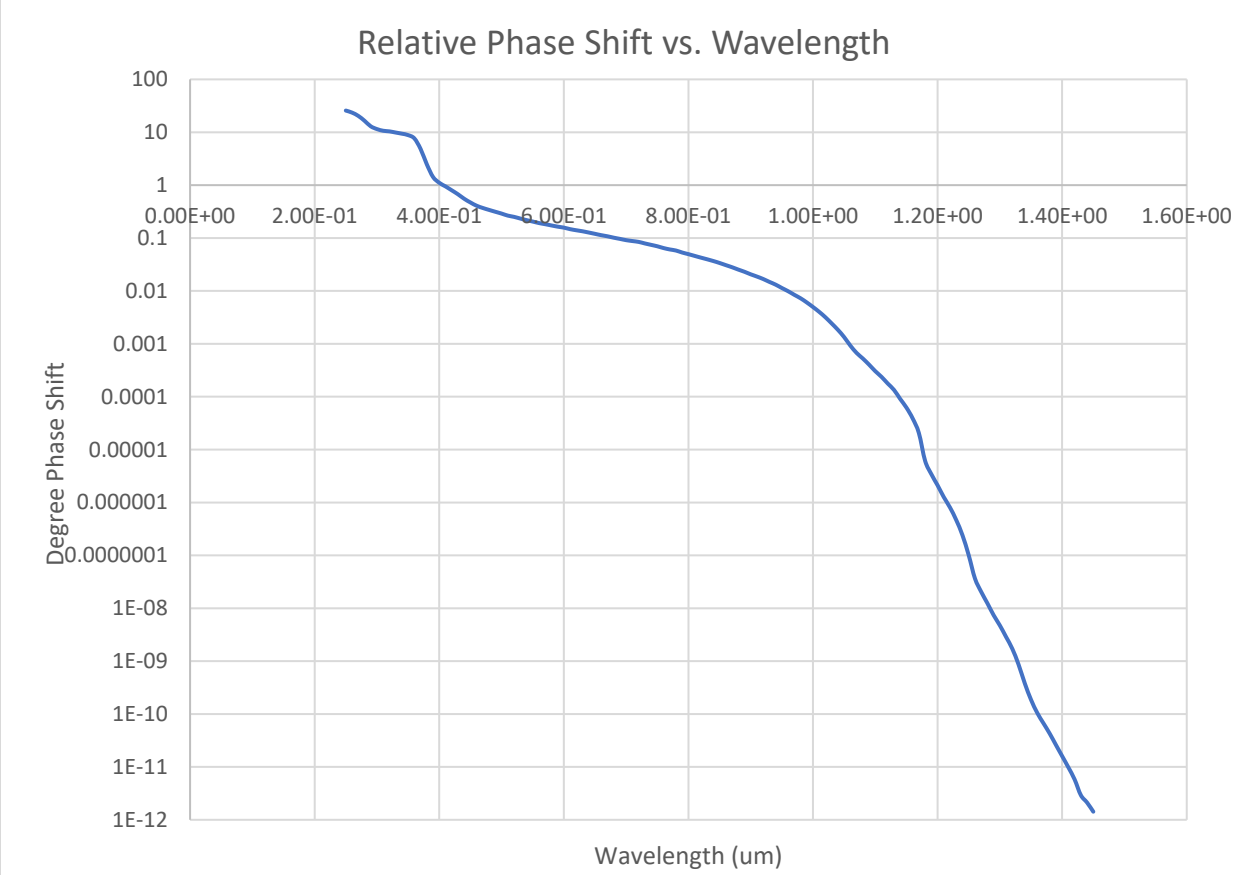
Complex Index of Refraction

To make this program more universal and have the ability to accurately measure thicknesses in semiconductors, we need to account for complex index of refraction and its effect on the phase change of the secondary ray (that gets reflected twice). For this ray at reflection, n1 (silicon for example) > n2 (vacuum or air).

For reflectance where the initial medium n1>n2, the phase shift of the wave relative to the unshifted is then:

$$\phi = \arctan\left(\frac{2n_2 k}{n_1^2 + k^2 - n_2^2}\right) \qquad \text{A9}$$

Although for something like Silicon, the FTIR wavelength range that we use (where good peaks are seen) is between 1.1 to 1.37um. The effect on the phase shift can be seen in the following plot with relation to wavelength:



So, we can see that in our working range, we are much below multiples of Pi (which I use as a bench mark as this is the required parameter for when simple counting the peaks and the distances between them). To avoid mistakes from the phase shift, one should simply choose the most transmitted part of the spectrum. If required however, using the above formula for Phi, one would have to adjust the formula to shift every point to account for the degree shift of the one reflecting arm.

We can see this by simply using a trigonometric identity:

$$\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) \qquad \text{A10}$$

In our case, β is essentially zero as even for thin films, the term in the cosine of A6 is on the order of 1. This eliminates the right term and makes cos(β) about 1. Otherwise, the adjustment necessary to the program would be to divide the solution A7 by G(λ) = cos(β (λ)) at each necessary wavelength. Taking the Fourier Transform after that *should* give a clean peak at cos(α) that trumps the result of the FFT of -sin(α)sin(β)/(G(λ)) .

The MATLAB code using these mentioned principles for calculating thickness from a transmission spectrum is presented in the following Appendix.

MATLAB Code for Finding Material Thickness Given FTIR Spectra

```
%%%%%%%%%%%%%%README%%%%%%%%%%%%%%
%This is a user-friendly script written to calculate the thickness of a
%sample given it's interference spectrum (amplitude vs wavenumber). It
%adjusts the pattern to account for the index of refraction,
%eliminates non oscillitory data (irrelevant to interference pattern), and
%then perform an FFT to determine the sample thickness.
%%%%%%%%%%%%%%%%%%IMPORTANT NOTES%%%%%%%%%%%
%The refraction indices here are given for the extraordinary and ordinary
%rays of crystal quartz. For the wavelengths covered by FTIR up to 60um,
%the transmission of the extraordinary ray is higher than the ordinary, and
%so the prevailing peak will corresond to that; hence, I recommend using
%the extraordinary index of refraction.
%You will often see two peaks close to each other, from left to right. This
%is likely attributed to the ordinary and extraordinary ray, in that order.
%Usually the extraordinary ray shoes up more prominently, so use that.
%More often you will see a second smaller peak at twice the x-value (thickness). It has been
%determined that this smaller peak is in fact a secondary internal
%reflection which can be used to verify the result. See documentation for
%more details: "On Sample Thickness Calculations"
%Note that this works for transmission mode of FTIR
%Author: Andrey Goussev 6/14/2018

clear all;
close all;

%Coefficients for calculating refractive index of x cut quartz
%or other material of choice (ordinary and extraordinary rays)
ao = 1.28604141; aex = 1.28851804;
bo = 1.07044083; bex = 1.09509924;
co = 1.00585997*10^-2; cex = 1.02101864*10^-2;
do = 1.10202242; dex = 1.15662475;
```

```
fo = 100; fex = 100;


[WaveNumber, Amplitude] = textread('C:\Users\Andrey\OneDrive - McMaster University\Research\Thickness Calculations through FTIR\repeatability_background.45.dpt');

%%%%%%%%%%%%%%%%%%
%Select Bounds
plot(Amplitude);  %plot Amplitude vs index to select the bounds
title('Select bounds in comman window');
Bounds(1) = 0;
while Bounds(1) == 0 %error handling since can't index at 0
   Bounds = input('Please select bounds based off plot with minimum bound at 1. E.g. "1:4000", "100:3500", etc": ');
   if Bounds(1) ==  0;
     fprintf('Bounds must start at an integer >0\n')
   end
end
close

WaveNumber1 = WaveNumber(Bounds,1);
Amplitude1 = Amplitude(Bounds,1);
Wavelength= (1./WaveNumber1)*10.^4;

%%%%%%%%%%%%%%%%%%
%Select which ray
%Remember to run/add to path the RefrIndex.m
which_Ray = input('Extraordinary [ex], Ordinary [or] ray or Other [other]? (suggest ex for quartz)  ','s')
if strcmp(which_Ray,'or')== 1;
   n = RefrIndex(ao,bo,co,do,fo,Wavelength);
elseif strcmp(which_Ray,'ex')== 1;
   n = RefrIndex(aex,bex,cex,dex,fex,Wavelength);
elseif strcmp(which_Ray,'other')==1;
   %Reads a file containing the values for wavelength and index of refraction
   %in order "wavelength [nm], index of refraction" with the first row omitted
   which_file = input('Input CSV file name in working folder with extension like "myfile.csv": ','s');
   m = csvread(strcat('C:\Users\Andrey\OneDrive - McMaster University\Research\Thickness Calculations through FTIR\',which_file),1,0);
   w = m(:,1)/1000; %reads and converts to [um] wavelength
   n1 = m(:,2);
   Refr_Index_Fit = fit(w,n1,'cubicinterp');
   n = feval(Refr_Index_Fit,Wavelength);
else
   fprintf('Input not valid; Defaulting to extraordinary ray');
   n = RefrIndex(aex,bex,cex,dex,fex,Wavelength);
end
```

```
WaveNumber_Adjusted = WaveNumber1.*n;
[xData, yData] = prepareCurveData( WaveNumber_Adjusted, Amplitude1 );

%%%%%%%%%%%%%%%%%%
%Set up fit type and options.
refit = 0;
opts = fitoptions( 'Method', 'SmoothingSpline' );
opts.Normalize = 'on';
opts.SmoothingParam = 0.9995803772631482;
while refit ~= 1
    ft = fittype( 'smoothingspline' );
   %Change this depending on the density of the peaks if you see that normalization is not working well
    %opts.SmoothingParam =0.9525741268224331
    % Fit model to data.
    [fitresult, gof] = fit( xData, yData, ft, opts );

    % Plot fit with data.
    figure( 'Name', 'untitled fit 1' );
    h = plot( fitresult, xData, yData );
    legend( h, 'Amplitude vs. WaveNumber', 'Smooth Fit', 'Location', 'NorthEast' );
    % Label axes
    xlabel WaveNumber
    ylabel Amplitude
    grid on
    isgoodfit = input('Does the fit look good? It should follow the general trend of the curve, not the small sinusoids.\n Yes [y],
Smoother [s], Rougher [r], Manual [m]: ','s');
    close
    if strcmp(isgoodfit,'y')== 1;
        refit = 1;
    elseif strcmp(isgoodfit,'s')== 1;
        opts.SmoothingParam = opts.SmoothingParam^2
    elseif strcmp(isgoodfit,'r')== 1;
        opts.SmoothingParam = (opts.SmoothingParam+1)/2
    elseif strcmp(isgoodfit,'m')== 1;
        opts.SmoothingParam = input('Input Smoothing Parameter 0-1: ');
    else
        fprintf('Input not valid; try again');
    end
end

%%%%%%%%%%%%
%Normalize
fit_eval = fitresult(WaveNumber_Adjusted);
Normalized = Amplitude1./fit_eval;
```

```
Normalized_Centered = Normalized-1;

%%%%%%%%%%
%Equally Space Data for FFT
Normalized_Centered_Fit = fit(WaveNumber_Adjusted,Normalized_Centered,'cubicinterp');

Lin_WaveNumber_Adjusted = linspace(min(WaveNumber_Adjusted),
max(WaveNumber_Adjusted),length(WaveNumber_Adjusted)); %create equally spaced query points
Lin_Spaced_Normalized_Centered = feval(Normalized_Centered_Fit,Lin_WaveNumber_Adjusted); %evaluates the previous cubic fit
at equally spaced points Lin_WaveNumber_Adjusted
Sampling_Period = abs(Lin_WaveNumber_Adjusted(2)-Lin_WaveNumber_Adjusted(1)); %for calculating effective Sampling
Frequency
L = length(Lin_WaveNumber_Adjusted);

%%%%%%%%%%%%%
%%%%%%%%%%%%%%
%Evaluate the FFT and thickness of sample. Makes sure user is happy with
%result :)

happy = 0;
while happy == 0
   trails = input('Input order number of padded zeros to add\nExample: 0-10+ for higher precision (not necessarily
accuracy)\nRecommend between 5-10:  ')
   L_Sample = L*2^trails;
   min_wave = min(Lin_WaveNumber_Adjusted);
   technique = input('Use hanning window or not [y/n]? Hanning gives less spectral leakage than 0 padding,\n and is generally good
for single spectral component detection, and detecting small signals next to large\n although worse for detecting adjacent large
peaks.:  ', 's')
   if strcmp(technique,'y');
     %%%For using the hann method
       Y=fft(hanning(length(Lin_Spaced_Normalized_Centered)).*Lin_Spaced_Normalized_Centered,L_Sample);
   elseif strcmp(technique,'n');
       Y = fft(Lin_Spaced_Normalized_Centered, L_Sample);
   end
   %%%
   P2 = abs(Y);
   P1 = P2(1:(L_Sample)/2+1);  %Cuts half of the spectrum off
    P1(2:end-1) = 2*P1(2:end-1);   %Multiplies spectrum by two except 1st (to keep properly normalized spectrum since we got rid of
half)
   Fs = 1/Sampling_Period;  %Equivalent to a 'Sampling Frequency' for normal sinusoid
   f = Fs*(0:((L_Sample)/2))/(L_Sample);
   f = f*10000/2; %Adjust axis to represent scale in [um]
%     plot(WaveNumber_Adjusted, Normalized_Centered)
   figure
   s(1) = subplot(2,1,1);
```

```
    plot(WaveNumber_Adjusted, Normalized_Centered); %plots subplot of normalized amplitude
    s(2) = subplot(2,1,2);
    plot(f,P1); %Plot FFT
    title(s(1),'Zero-Centered Spectrum')
    xlabel(s(1),'Wavenumber [cm^-1]')
    title(s(2),'Fourier Transform Domain')
    xlabel(s(2),'Thickness [um]')
    [psor,lsor] = findpeaks(P1,f,'SortStr','descend');
    fprintf('1st peak %s at %s [um]\n',psor(1),lsor(1));
    fprintf('2nd peak %s at %s [um]\n',psor(2), lsor(2));
    fprintf('Average %s at %s [um]\n',(psor(2)+psor(1))/2, (lsor(2)+lsor(1))/2);
    happy_quest = input('Are you happy (with results) [y/n]:  ','s');
    if strcmp(happy_quest,'y')== 1;
        happy = 1;
    elseif strcmp(happy_quest,'n')== 1;
        close all;
        continue
    end
end
 fprintf('Done!\n')
%%%%%%%%%%%%%
```