ANALYSIS OF THE NONLINEAR VIBRATIONS OF ELECTROSTATICALLY ACTUATED MICRO-CANTILEVERS IN HARMONIC DETECTION OF RESONANCE (HDR)

Jonathan Taylor
Clemson University, jdtaylo@clemson.edu

Follow this and additional works at: https://tigerprints.clemson.edu/all_theses
Part of the Physics Commons

Recommended Citation
https://tigerprints.clemson.edu/all_theses/458

This Thesis is brought to you for free and open access by the Theses at TigerPrints. It has been accepted for inclusion in All Theses by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.
ANALYSIS OF THE NONLINEAR VIBRATIONS OF ELECTROSTATICALLY
ACTUATED MICRO-CANTILEVERS IN HARMONIC DETECTION OF
RESONANCE (HDR)

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Physics

by
Jonathan Davis (JD) Taylor
August 2008

Accepted by:
Dr. Apparao Rao, Committee Chair
Dr. Malcolm Skove
Dr. Chad Sosolik
ABSTRACT

Micro- and nano-cantilevers have the potential to revolutionize physical, chemical, and biological sensing; however, an accurate and scalable detection method is required. In this work, a fully electrical actuation and detection scheme is presented, known as the Harmonic Detection of Resonance (HDR), in which harmonic components of the electrical current are measured to determine the cantilever’s resonance frequency. These harmonics exist as a result of nonlinearities in the system, principally in the electrostatic actuation force. In order to better understand this rich harmonic structure, a theoretical investigation of the micro-cantilever is undertaken. Both a lumped parameter model and a more accurate continuum model are used to derive the governing nonlinear equations of motion (EOM) of the cantilever. Various approximate solution methods applicable to nonlinear equations are then discussed including numerical integration, perturbation, and averaging. The method of harmonic balance is then used to obtain steady state solutions of the micro-cantilever EOM. Low-order closed-form harmonic balance solutions are derived which explain many of the important features of the HDR results, such as the presence of parasitic capacitance in the first harmonic and super-harmonic resonance peaks in higher harmonics. Finally, higher-order computer generated harmonic balance solutions are presented which show good agreement with the experimental HDR results, validating both the modeling and the solution methods used.
DEDICATION

This work is dedicated to my father, Donald Taylor, whom I shall never forget.
ACKNOWLEDGMENTS

My utmost gratitude is extended to Dr. Apparao Rao and Dr. Malcolm Skove for their continued patience and support over the many years of this project.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE PAGE</td>
<td>i</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td><strong>I. INTRODUCTION AND LITERATURE REVIEW</strong></td>
<td>1</td>
</tr>
<tr>
<td>Preliminary Remarks and Chapter Summary</td>
<td>1</td>
</tr>
<tr>
<td>Micro- and Nano-Cantilever Sensing</td>
<td>3</td>
</tr>
<tr>
<td>Transduction Mechanisms</td>
<td>6</td>
</tr>
<tr>
<td>Detection Methods</td>
<td>13</td>
</tr>
<tr>
<td>Criticisms of Micro-Cantilever Sensors</td>
<td>15</td>
</tr>
<tr>
<td><strong>II. HARMONIC DETECTION OF RESONANCE (HDR)</strong></td>
<td>16</td>
</tr>
<tr>
<td>Description of HDR</td>
<td>16</td>
</tr>
<tr>
<td>Mechanical vs. Electrical Responses</td>
<td>17</td>
</tr>
<tr>
<td>HDR Experimental Apparatus</td>
<td>18</td>
</tr>
<tr>
<td>Lock-In Amplifiers</td>
<td>20</td>
</tr>
<tr>
<td>Resonance</td>
<td>21</td>
</tr>
<tr>
<td>Experimental Frequency Responses</td>
<td>23</td>
</tr>
<tr>
<td>Polar Plots of Resonance</td>
<td>27</td>
</tr>
<tr>
<td><strong>III. REVIEW OF LINEAR SYSTEMS</strong></td>
<td>29</td>
</tr>
<tr>
<td>Relevance of Linearized Models</td>
<td>29</td>
</tr>
<tr>
<td>Universal Linear Oscillator</td>
<td>30</td>
</tr>
<tr>
<td>Simple Harmonic Oscillator (SHO)</td>
<td>32</td>
</tr>
<tr>
<td>Quality Factor</td>
<td>41</td>
</tr>
<tr>
<td>Forced Damped Harmonic Oscillator</td>
<td>41</td>
</tr>
<tr>
<td>Green’s Functions</td>
<td>47</td>
</tr>
<tr>
<td>Euler-Bernoulli Beam Theory</td>
<td>48</td>
</tr>
</tbody>
</table>
Table of Contents (Continued)

<table>
<thead>
<tr>
<th>Mode vs. Harmonics</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modes vs. Harmonics</td>
<td>56</td>
</tr>
</tbody>
</table>

IV. SYSTEM MODEL AND EQUATIONS OF MOTION ........................................... 59

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>59</td>
</tr>
<tr>
<td>Micro-Cantilever System Variables</td>
<td>59</td>
</tr>
<tr>
<td>Lumped Parameter Model</td>
<td>60</td>
</tr>
<tr>
<td>Non-Dimensionalized EOM</td>
<td>66</td>
</tr>
<tr>
<td>Hamilton’s Principle</td>
<td>69</td>
</tr>
<tr>
<td>Energy Expressions</td>
<td>70</td>
</tr>
<tr>
<td>Variational Equations of Motion</td>
<td>74</td>
</tr>
<tr>
<td>Modal Expansion and Discrete EOMs</td>
<td>76</td>
</tr>
<tr>
<td>Uncoupled Non-Dimensionalized Modal EOMs</td>
<td>79</td>
</tr>
<tr>
<td>Validity of the Parallel Plate Model</td>
<td>81</td>
</tr>
</tbody>
</table>

V. APPROXIMATE SOLUTION METHODS TO NONLINEAR PROBLEMS ......................... 84

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>84</td>
</tr>
<tr>
<td>The General Nonlinear Problem</td>
<td>86</td>
</tr>
<tr>
<td>The Duffing Oscillator</td>
<td>87</td>
</tr>
<tr>
<td>Hardening and Softening Springs</td>
<td>88</td>
</tr>
<tr>
<td>Numerical Integration</td>
<td>90</td>
</tr>
<tr>
<td>Numerical Solutions to the Duffing Oscillator</td>
<td>93</td>
</tr>
<tr>
<td>Period Doubling, Bifurcations, and the Onset of Chaos</td>
<td>102</td>
</tr>
<tr>
<td>Perturbation Methods</td>
<td>103</td>
</tr>
<tr>
<td>Averaging Methods</td>
<td>112</td>
</tr>
</tbody>
</table>

VI. ANALYTICAL AND NUMERICAL RESULTS .................................................... 115

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>115</td>
</tr>
<tr>
<td>Closed-Form Harmonic Balance Solutions</td>
<td>116</td>
</tr>
<tr>
<td>Electrical Response Calculations</td>
<td>126</td>
</tr>
<tr>
<td>Higher-Order Computational Harmonic Balance Solutions</td>
<td>129</td>
</tr>
<tr>
<td>Nonlinearities, Duffing, and Chaos in a Micro-Cantilever</td>
<td>138</td>
</tr>
</tbody>
</table>

APPENDIX

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harmonic Balance Program in MAPLE™</td>
<td>141</td>
</tr>
</tbody>
</table>

REFERENCES .................................................................................................. 149
<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Roots of frequency equation for fixed-free beam</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>1.1</td>
<td>5</td>
</tr>
<tr>
<td>1.2</td>
<td>7</td>
</tr>
<tr>
<td>2.1</td>
<td>18</td>
</tr>
<tr>
<td>2.2</td>
<td>22</td>
</tr>
<tr>
<td>2.3</td>
<td>25</td>
</tr>
<tr>
<td>2.4</td>
<td>26</td>
</tr>
<tr>
<td>2.5</td>
<td>28</td>
</tr>
<tr>
<td>3.1</td>
<td>32</td>
</tr>
<tr>
<td>3.2</td>
<td>40</td>
</tr>
<tr>
<td>3.3</td>
<td>44</td>
</tr>
<tr>
<td>3.4</td>
<td>45</td>
</tr>
<tr>
<td>3.5</td>
<td>46</td>
</tr>
<tr>
<td>3.6</td>
<td>49</td>
</tr>
<tr>
<td>3.7</td>
<td>49</td>
</tr>
<tr>
<td>3.8</td>
<td>55</td>
</tr>
<tr>
<td>3.9</td>
<td>56</td>
</tr>
<tr>
<td>4.1</td>
<td>60</td>
</tr>
<tr>
<td>4.2</td>
<td>61</td>
</tr>
</tbody>
</table>
List of Figures (Continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3</td>
<td>Comparison of lumped parameter model, modal model, and experiment</td>
<td>62</td>
</tr>
<tr>
<td>4.4</td>
<td>Parallel-plate and parallel-cylinder capacitance schematics</td>
<td>81</td>
</tr>
<tr>
<td>4.5</td>
<td>Parallel-plate and parallel-cylinder capacitances as functions of gap separation</td>
<td>83</td>
</tr>
<tr>
<td>5.1</td>
<td>The nonlinear restoring force for a Duffing oscillator for hardening, linear, and softening springs</td>
<td>89</td>
</tr>
<tr>
<td>5.2</td>
<td>The symmetric potential of a Duffing oscillator</td>
<td>95</td>
</tr>
<tr>
<td>5.3</td>
<td>The time response and phase space portraits of the Duffing oscillator</td>
<td>96</td>
</tr>
<tr>
<td>5.4</td>
<td>The numerically determined potential, time response, phase portrait and frequency response curve of a Duffing oscillator with weak nonlinearity and forcing</td>
<td>98</td>
</tr>
<tr>
<td>5.5</td>
<td>Poincaré map of the weakly nonlinear Duffing oscillator</td>
<td>100</td>
</tr>
<tr>
<td>5.6</td>
<td>Numerically determined frequency response curves for the forced damped Duffing oscillator</td>
<td>101</td>
</tr>
<tr>
<td>5.7</td>
<td>Numerically determined phase portrait and Poincaré map illustrating the onset of chaos in a forced damped Duffing oscillator</td>
<td>103</td>
</tr>
<tr>
<td>5.8</td>
<td>Perturbation based analytic solution to the frequency-response curve of a hardening undamped forced Duffing oscillator</td>
<td>111</td>
</tr>
<tr>
<td>6.1</td>
<td>Mechanical first harmonics from closed form harmonic balance calculations</td>
<td>124</td>
</tr>
<tr>
<td>6.2</td>
<td>Mechanical second harmonics from closed form harmonic balance calculations</td>
<td>125</td>
</tr>
<tr>
<td>6.3</td>
<td>Electrical first harmonics from closed form harmonic balance showing parasitic</td>
<td>127</td>
</tr>
</tbody>
</table>
List of Figures (Continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.4</td>
<td>Electrical second harmonics from closed form harmonic</td>
<td>128</td>
</tr>
<tr>
<td>6.5</td>
<td>Computed mechanical frequency</td>
<td>131</td>
</tr>
<tr>
<td>6.6</td>
<td>Computed electrical (current) frequency response</td>
<td>132</td>
</tr>
<tr>
<td>6.7</td>
<td>Comparison of experimental data and harmonic balance results</td>
<td>134</td>
</tr>
<tr>
<td>6.8</td>
<td>Polar plot showing first harmonic of current signal</td>
<td>136</td>
</tr>
<tr>
<td>6.9</td>
<td>Polar plot showing second and third harmonic of current signal</td>
<td>137</td>
</tr>
<tr>
<td>6.10</td>
<td>Computed response of system with large nonlinearities</td>
<td>139</td>
</tr>
<tr>
<td>6.11</td>
<td>Experimental polar plots for the electrical response illustrating Duffing jumps</td>
<td>140</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

1. PRELIMINARY REMARKS AND CHAPTER SUMMARY

In 1984, Binnig, Quate, and Gerber described the first atomic force microscope (AFM) which utilized a micro-scale cantilever to measure the surface topography of a sample to atomic precision [1]. The small size and relatively high natural frequency of the micro-cantilever made it especially sensitive to minute surface forces. In the years that followed, micro-cantilevers demonstrated their versatility in a wide variety of different applications, including sensing, actuation, and power generation. Micro-cantilevers are currently one of the most prolific and well studied components of micro-electro-mechanical systems (MEMS). Furthermore, because there are advantages to reducing the size of the devices even further, there has been intense research in recent years into nano-scale cantilevers, the most common of which are based on carbon nano-tubes (CNTs)

In order to effectively design MEMS and NEMS devices, it is important to have effective theoretical models predicting their behavior. Unfortunately the system models often become quite complex and cumbersome. At these scales, systems tend to become more nonlinear and they begin to depend more on various “coupled-domains”, e.g. mechanics, electromagnetism, and chemistry. For instance, the mechanical motion of the micro-cantilever considered in this study is governed by an inherently nonlinear electric force.
The aim of this work is to provide a more complete analysis of the vibrations of a micro- or nano-cantilever subject to a nonlinear electrostatic actuation force. Simplistic lumped parameter models and more accurate modal models are presented. Various numerical and analytical solution methods are then introduced and used to determine the response of the system.

The preceding analysis is then used to explain some interesting experimental results obtained in technique known as the Harmonic Detection of Resonance (HDR) [2]. HDR is an electrostatic actuation and detection method relevant to micro- and nano-cantilever based sensing. The experimental results exhibit many features that are unique to nonlinear systems, such as significant harmonic contributions and multiple resonances.

Remarkably, in addition to their practical utility, micro-cantilevers are also fascinating from a theoretical point of view. Micro-cantilevers exhibit a transition from periodic to chaotic motion that is extremely interesting to those in the field of nonlinear dynamical systems. Furthermore, the ability to delicately control the onset of nonlinearities by varying the voltages and gap distances makes the electrostatically actuated micro-cantilever especially suitable to theoretical analysis and experimental verification.

This work places emphasis on the fundamental principles underlying the problem. To this end, two chapters have been included which are strictly reviews, Chapter III covers the linear models relevant to the micro-cantilever system: the linear harmonic oscillator and the Euler-Bernoulli beam. Chapter V surveys the approximate solution methods
available for solving nonlinear problems. Both chapters are immensely valuable in familiarizing oneself with the basic terminology and also in gaining insight into the wide range of possible responses of an electrostatically actuated micro-cantilever.

The other chapters address more specifically the electrostatically actuated micro-cantilever. Chapter II describes HDR and presents relevant experimental results. Chapter IV derives the equations of motion using a general variational energy approach, and, finally, Chapter VI presents the analytical and numerical results obtained for the micro-cantilever system.

2. MICRO- AND NANO-CANTILEVER SENSING

Micro- and nano-cantilevers have the potential to revolutionize physical, chemical, and biological sensing. Their exceptionally small size allows for unprecedented sensitivities, improved dynamic performance and reliability, and low power consumption. In particular, micro-cantilevers are easily integrated into standard high-volume silicon manufacturing processes making them relatively inexpensive and mass-producible [3].

They are sensitive to a variety of environmental parameters including temperature, pressure, humidity, and infrared radiation [4], and they can be made to respond selectively to specific chemical and biological species by means of functionalized surface treatments [5]. These features make micro- and nano-cantilevers ideal candidates for a wide variety of sensing applications and attractive alternatives to traditional sensing technologies.
However, in order to fully realize cantilever based sensing in micro- and nano-electro-
mechanical systems (MEMS and NEMS), an accurate and scalable detection method is
required. This detection method must be capable of measuring changes in the dynamic
response that result from changes in environmental parameters. A fully electrical
(actuation and detection) scheme is presented in this work known as the Harmonic
Detection of Resonance (HDR) that meets these requirements and provides several
unique advantages not present in other detection techniques.

Micro-cantilevers are arguably the simplest micro-scale mechanical structures and
certainly one of the most versatile. They may be considered as basic building blocks for
more complicated micro-systems [6]. The singly–clamped cantilever (“diving-board”),
geometry will respond to much less force than other micro-structures and is thus better
suited for sensing applications in which deflections must be measured; however, doubly-
clamped (“bridge”) structures have on many occasions been used and are in general
easier to manufacture [7].

Typical micro-cantilevers, such as those used throughout this article, are singly-clamped
structures made of single crystal silicon using lithographic and surface micromachining
processes. Several geometries are available depending on the application. A few are
shown in Figure 1.1.
Nano-cantilevers, however, are most commonly nano-tubes or nano-wires grown using chemical vapor deposition (CVD) [8]. In addition to sensing, micro- and nano-cantilevers have been used as both actuators [9] and micro-generators [10].

In general, as the size of a sensor is reduced, sensitivity improves. It is therefore advantageous to investigate nano-scale sensors. Many measured environmental parameters depend on the surface to volume ratio which is roughly a thousand times greater for nano-scale than for micro-scale structures. Also, intrinsic damping is generally less in nano-structures because of fewer defects.

A full understanding of micro-cantilever based sensing is necessary in order to extend this technology to the nano-scale. At the nano-scale the sensitivity to environmental parameters is usually much improved; however, the signal to background noise is generally much lower and the number of detection schemes applicable to the nano-scale is limited. For these reasons, this work focuses primarily on micro-cantilevers; however, all the results may be applied to nano-scale structures as long as the continuum approximations are still valid.
3. TRANSDUCTION MECHANISMS

There are several basic transduction mechanisms applicable to micro- and nano-cantilever based systems. Transduce means to convert one type of energy or physical attribute to another for various purposes. These can broadly be classified as adsorbed / absorbed mass, induced stress, changes in pressure / damping, piezo effects, and applied mechanical and electromagnetic forces. A useful graphical summary of these transduction mechanisms by Baller [11] is provided in Figure 1.2.

Systems may utilize these mechanisms for sensing (sensed parameter → mechanical response), detection (mechanical response → output signal), or actuation (input signal → mechanical response).

Detection involves converting either the static deflection or the dynamic response of the cantilever into a useful output signal, usually electrical in nature. The dynamic response may include shifts in natural resonance frequency, changes in vibrational amplitude and phase, or changes in quality factor (Q-factor).

3.1 Adsorbed and Absorbed Mass

Substances which adsorb on the surface or absorb into the bulk of a micro-cantilever increase its effective mass and decrease its resonance frequency. The sensitivity, $S$, to changes in mass of a micro-cantilever sensor with resonance frequency $\omega_0$ and effective mass $m$ is given in (1.1) assuming constant bending stiffness. From this expression it is
obvious that structures with low effective mass and relatively high natural frequencies, e.g. micro- and nano-cantilevers, are best suited to mass sensing because they offer higher sensitivities. Recently, adsorbed molecular masses as low as a few zepto-grams ($10^{-21}$ g) have been observed using nano-cantilevers [12].

$$S = \frac{\Delta \omega}{\Delta m} = \frac{d (k/m)^{1/2}}{dm} = \frac{\omega_0}{2m}$$

(1.1)

**Figure 1.2:** Overview of transduction methods for micro-cantilever based sensors [11].
3.2 Induced Stresses

Several mechanisms exist that can induce differential stresses in asymmetrically coated micro-cantilevers which may result in either static deflection (bending) of the cantilever or shifts in its resonance frequency. The effects of induced surface stresses are especially significant in micro- and nano-structures due to their large surface area to volume ratio.

3.2.1 Molecular Adsorption / Interfacial Chemical Reactions

Cantilevers that have been coated on one side with a thin chemically selective receptor layer will bend as molecules adsorb on the surface. This adsorption may be of the low energy Van der Waals type (physisorption) or the higher energy covalent type (chemisorption). Spontaneous molecular adsorption causes a reduction of interfacial free energy and surface stress and a concomitant expansion of the material [6]. The resulting stress gradient causes a static deformation of the cantilever. The adsorption process also tends to stiffen the cantilever, thereby increasing its resonance frequency, as opposed to mass loading which lowers the frequency. A wide variety of highly selective chemical sensors have been developed utilizing these phenomena [5].

3.2.2 Analyte Induced Expansion

Cantilevers coated with a relatively thick analyte permeable receptor layer may bend due to analyte induced swelling. Molecules may absorb into the bulk of the coating thereby changing either the internal stress or pressure depending on whether the coating is solid
or gel-like [4]. This effect has been employed to measure humidity using polymeric hydrogel coatings [13].

3.2.3 Thermally Induced Stresses and Calorimetry

Thermally induced stresses arise due to unequal coefficients of thermal expansion in layered cantilevers. Typically the cantilevers are coated with a thin metallic layer, *e.g.* gold. This mechanism is commonly referred to as the “bimetallic effect” and is frequently employed in home thermostats. The heat producing the thermal stresses may arise from several sources including embedded resistors [5] or IR radiation. Also, micro-cantilevers can be used as micro-calorimeters to detect the heat produced during molecular adsorption or during subsequent associated exothermal reactions [4]. Typically thermal actuation requires significantly more power than other transduction methods.

3.2.4 Optical Radiation

Electromagnetic radiation of a variety of wavelengths from approximately 1 nm to 1 μm (UV – visible – near IR) gives rise to mechanical strains in micro-cantilevers, though the greatest deflections occur at IR frequencies. This can be attributed to both radiative heating and the generation of photo-induced free charge carriers. For silicon, these effects act in opposite directions [14].

9
3.3 Pressure and Damping

Changes in the ambient pressure may affect the damping experienced by a microcantilever. Pressure changes affect the vibrational amplitude and Q-factor of the micro-cantilever resonator and thus must be measured dynamically. In gases, there are three basic pressure regimes: intrinsic, molecular, and viscous. Depending on the pressure regime, various physical parameters may be determined, such as defect density of the oscillator [15], molecular mass and gas composition [16], or viscosity of the surrounding medium [17, 18]. Variations in pressure due to acoustic waves can also be detected [19].

3.4 Piezoelectric and Piezoresistive

Piezoelectric materials will generate a mechanical strain when subjected to an applied electric field. This phenomenon is extensively applied to actuate micro-cantilevers notably in atomic force microscopes. Silicon is not intrinsically piezo-electric; therefore, a piezoelectric layer, \textit{e.g.} lead zirconium titanate (PZT), must be deposited in post-processing. This leads to a more complicated and costly production process as compared to capacitive designs. Piezoelectric materials also generate a voltage when mechanically strained. Therefore, piezoelectric coated micro-cantilevers may be used to detect both static and dynamic deflections arising from any of the other transduction mechanisms.

Piezoresistance is the change in resistivity of a material with applied stress. It is commonly used to detect the deflection of micro-cantilevers. Silicon is intrinsically piezoresistive and this property can be enhanced by doping; thus piezoresistive detection
is highly compatible with standard CMOS processes. Piezoresistive elements are typically placed at the base of the cantilever where the stresses from bending are greatest and are usually arranged in a Wheatstone bridge configuration in order to negate common mode effects such as thermal variations [5].

3.5 Applied Forces

A variety of externally applied forces are capable of generating static and dynamic deflections in micro- and nano-cantilevers. These forces may be either mechanical (applied directly or through inertial loading) or electromagnetic in origin. MEMS and NEMS based mechanical sensors offer the potential for improved sensitivities, lower power consumption, and wider bandwidths than conventional resistance strain gauges [3].

3.5.1 Mechanical Force and Torque

Micro- and nano-cantilevers are extremely sensitive to mechanically applied forces. Pico-newton forces are routinely measured and even higher sensitivities have been achieved by cooling the cantilevers to milli-kelvin temperatures [20]. This mechanism forms the basis of contact mode atomic force microscopy (AFM), in which the surface topography of a sample is determined by scanning a micro-cantilever in the \(xy\)-plane and detecting its static deflection in the \(z\)-direction in response to surface contours [1]. Resonant strain gauges have also been developed using doubly-clamped micro-beams in which externally applied tensile strains stiffen the beam and increase its resonant frequency [21].
3.5.2 *Gravitational and Inertial Forces*

Micro-mechanical accelerometers and gyroscopes, which measure linear and angular acceleration respectively, are some of the most widespread MEMS devices. They are used in applications ranging from airbag release systems to military inertial guidance [3]. These sensors function by measuring the static or dynamic deflection of a proof mass attached to a compliant support such as a micro-cantilever. The cantilever deflects in order to counteract the inertial loading of the proof mass due to the base acceleration. Accelerometers have also been developed based on resonant silicon structures which are more immune to environmental noise and better suited to sensing dynamic accelerations [22].

3.5.3 *Electrostatic Forces*

The electrostatic transduction mechanism is based on Coulomb’s law from which it follows that two oppositely charged elements will experience an attractive force. This mechanism is very common because it can be used for both actuation and detection and is quite straightforward to fabricate. If the elements can be modeled as a parallel plate capacitor, the electrostatic force, $F_E$, is given by (1.2).

$$F_E = \frac{\varepsilon AV^2}{2d^2}$$

(1.2)

where $\varepsilon$ is the permittivity of the medium separating the electrodes, $A$ is the plate area, $V$ is the applied voltage, and $d$ is the separation distance. The electrostatic force is a nonlinear function of the separation distance and voltage. This nonlinearity is essential to
the HDR method; however, it is not always desirable, as in the case of non-contact mode
AFM, in which a micro-cantilever is vibrated above the surface of a sample and shifts in
its resonance due to variations in the electrostatic force are measured. In these cases,
feedback mechanisms are usually employed to keep the response sufficiently linear.

3.5.4 Magnetic Forces

A current carrying element placed in a magnetic field experiences a Lorentz force in a
direction perpendicular to both the current and magnetic field. This mechanism is the
basis for magnetic force microscopy (MFM) and scanning hall probe microscopy
(SHPM) [3]. Also, a magnetic micro-actuator has been developed that utilizes an
electroplated permalloy that possesses a high magnetic permeability [23]; however,
because there are a limited number of magnetic materials compatible with current micro-
manufacturing processes and only planar coils are possible, it is very difficult to generate
magnetic fields on chip, and thus magnetic transduction’s applicability to MEMS and
NEMS has been somewhat limited.

4. DETECTION METHODS

Several detection schemes have been proposed to measure the static and/or dynamic
response of micro- and nano-cantilevers. The most common is laser reflectometry, in
which a low power laser is reflected off the cantilever and measured with a position
sensitive photo-detector. This method is employed successfully in almost all AFMs.
Other detection methods include various forms of microscopy (optical, scanning electron,
or transmission electron), piezo-resistive or piezo-electric, interferometry, and diffraction methods.

However, all of these detection schemes require complicated electronics that take up significant space and power and are not possible to integrate on a single chip. They are consequently not scalable to the micro- and nano-scale. For these purposes, the standard detection methods mentioned above are not suitable.

4.1 Electrostatic Actuation and Capacitive Detection

Electrostatic actuation and capacitive detection presents an alternative that does meet these scaling criteria. In this method, a potential difference is applied between a conductive micro-cantilever and counter-electrode resulting in an attractive electro(quasi)static force. In response to this force, the cantilever deflects, and the capacitance of the arrangement varies, causing charge to move on and off the cantilever. If this charge or current can be measured, the mechanical vibration of the micro-cantilever can be deduced. This is the basis of the harmonic detection of resonance method (HDR) and the topic of the next chapter.

However, electrostatic actuation and capacitive detection have traditionally proven difficult to implement. This difficulty can in large part be attributed to a parasitic signal that obscures the dynamic signal from the cantilever. This parasitic signal includes both the static capacitance of the micro-cantilever and counter-electrode and all the stray capacitance of nearby circuit elements. Several methods have been proposed to enhance
the dynamic capacitance or lower the parasitic capacitance of the system, including single
electron transistors [24], and comb drives [25, 26]. Also since the ratio of dynamic to
parasitic signal depends on the ratio of cantilever deflection to total gap distance, some
groups have attempted to minimize the parasitic effects by manufacturing cantilevers that
are extremely close to the counter-electrodes. Each of these solutions increases the
complexity, cost of production, and potential for malfunction.

5. CRITICISMS OF MICRO-CANTILEVER SENSORS

There are some reservations about micro-cantilever based sensing. One author wrote,
“Attempts to achieve a more rigorous understanding of the stresses that cause
[microcantilever] bending would involve models from very distinct scientific areas, such
as molecular modeling, surface science, colloidal chemistry, and mechanical engineering.
Naturally, researchers question the consistency of these models and their appropriateness
for evaluating deflections” [27].

Micro-electro-mechanical systems are indeed complex and difficult to model using
traditional approaches. This is especially evident in this work for the coupling of the
electrostatic and mechanical domains. However, as with all new advances, the
simulation methods will eventually be optimized and with more and more successes, the
doubts will eventually be assuaged.
1. DESCRIPTION OF HDR

A capacitive detection method has been developed known as HDR that avoids parasitic capacitance without significantly increasing the complexity of the device. The electrical signal from a micro-cantilever when driven by a nearby counter-electrode has a “rich harmonic structure” [2], which can be attributed to nonlinearities in the electrostatic force and mixing of the mechanical and electrical signals.

The higher harmonic components of the signals, at integer multiples of the driving frequency, do not suffer from significant parasitic effects. Consequently, by measuring the dynamic response of micro-cantilevers at these harmonic frequencies, significantly higher signal to background ratios (SBR) and Q-factors can be obtained, resulting in much improved sensitivity in HDR based sensing devices.

HDR presents several advantages over other detection schemes. It is an entirely electrical actuation and detection scheme, and consequently, it is directly scalable to micro- and nano-devices with straightforward integration into standard micro-lithographic processes. This allows for portable HDR based sensing devices. HDR is also extremely simple. It requires no complicated components, such as lasers, magnets or piezo-electric elements to actuate or measure the cantilever, thereby reducing cost and potential for failure. HDR
does require circuitry to detect the higher harmonics, but this should be possible to realize on a single chip. Finally, the gap distances in HDR can be relatively large increasing available working distances and voltages and facilitating alignment.

Also HDR operates effectively under ambient conditions; therefore it does not require any expensive equipment to regulate temperature or pressure, such as vacuum systems. Thus, HDR devices can be made small and portable.

2. MECHANICAL VS. ELECTRICAL RESPONSES

There are distinct differences between the mechanical and electrical responses of an electrostatically actuated micro-cantilever. The mechanical response of a cantilever is simply its physical deflection as seen under a microscope, or measured using AFM based laser reflectometry. Like all oscillators, the mechanical response depends on both the amplitude and frequency of the applied force. In this chapter, the mechanical response is characterized by the tip deflection of the cantilever, $z(t)$.

As the cantilever deflects, charge moves on and off due to the variable capacitance. This current, which we call the electrical response, is measured in HDR. The current depends on the capacitance of the system which has the mechanical deflection as a parameter. In general, the electrical response exhibits more features than the mechanical response, such as parasitic capacitance and super-harmonic resonances.
3. HDR EXPERIMENTAL APPARATUS

In order to examine the differences between the mechanical and electrical responses and understand the unique advantages of HDR as a sensing technology, we describe an experiment in which the mechanical motion of a cantilever and the electrical response are measured simultaneously using standard AFM based laser reflectometry and the HDR technique respectively.

**Figure 2.1:** A schematic of the AFM and HDR experiment.
In this experiment, a cantilever is manipulated over an optical dark-field microscope. In most cases, this allows for simple positioning of the cantilever near the counter-electrode without the need for time consuming lithographic processes. The cantilever is placed parallel to and within 1–10 µm from the counter-electrode depending on its dimensions. An electrostatic force is generated by applying an ac voltage, $V_{ac}$, with a dc offset, $V_{dc}$. This experiment was performed under ambient conditions, demonstrating that HDR does not require any elaborate apparatus to control temperature or pressure.

The HDR system consists of an A250 pre-amplifier, a voltage oscillator, a dc power supply, and a lock-in amplifier, Figure 2.1. It is useful to employ a Faraday cage which surrounds the metal contacts. This minimizes crosstalk between the metal contacts which hold the cantilever and the counter electrode and helps to increase the signal to background ratio (SBR). These noise considerations are crucial when working at the nanoscale. The lock-in amplifier detects the output of the A250, which is proportional to the current, at a harmonic (integer multiple) of the oscillator driving frequency, $\Omega$. As will be discussed later, harmonics of the applied ac voltage are essential to the HDR method.

The micro-cantilever assembly is then placed inside an atomic force microscope (Veeco CPII). The mechanical deflection of the cantilever can then be directly measured using the laser and photo-sensitive position detectors of the AFM. As in HDR, a lock-in is used in this case to separate out the harmonics of the mechanical signal from the photo-diode.
4. LOCK-IN AMPLIFIERS

It is worthwhile to briefly describe the operation of lock-in amplifiers since they are such an integral component of the HDR detection system. Lock-in amplifiers are electronic instruments capable of extracting extremely small signals of known frequency from otherwise noisy signals. For this reason they are ideally suited to measuring the higher harmonic components of the cantilever electrical response which can be many orders of magnitude smaller than the noisy first harmonic.

Lock-in amplifiers operate based on heterodyne detection principles, in which a reference frequency signal is mixed (multiplied) with the input signal, $V_0 \cos(\Omega t)$, (2.1). The result is a signal with a dc component that is proportional to the amplitude of the input signal, $V_0$, at the reference frequency and a component at twice the reference frequency. The dc component is isolated using a low-pass filter (integrator) with a time constant chosen such that the $2\omega$ signal is strongly attenuated [28].

$$V_{max} = V_m(t) \cdot V_{ref}(t) = V_0 \cos(\Omega t) \cos(\Omega t) = \frac{1}{2} V_0 [1 + \cos(2\Omega t)] \quad (2.1)$$

The outputs of typical lock-in amplifiers are the in phase and quadrature (90° out of phase) components of the input signal at the reference frequency, from which the overall amplitude and phase shift of the measured signal can be determined. Most lock-in amplifiers multiply by a square wave reference signal, which includes many higher harmonics. However, in HDR a “digital” lock-in is used (Stanford Research Systems Model SR830) which multiples by a pure sinusoid thereby providing more accurate harmonic measurements.
The standard steady-state mechanical response (amplitude and phase) of a single degree of freedom (SDOF) forced damped harmonic oscillator is given (2.2) [29]. At the resonance frequency, \( \omega_0 \), the amplitude peaks and the phase difference changes by \( 180^\circ \).

\[
A = \frac{F_0 / m}{\sqrt{\left(\omega_0^2 - \Omega^2\right)^2 + \left(2\gamma\omega_0\Omega\right)^2}}
\]

\[
\phi = \tan^{-1}\left(\frac{2\gamma\omega_0\Omega}{\omega_0^2 - \Omega^2}\right)
\]

where \( A \) and \( \phi \) are the amplitude and phase shift of the steady-state displacement. \( F_0 \) and \( \Omega \) are the magnitude and angular frequency of the applied force, and \( m, \omega_0, \) and \( \gamma \) are the mass, natural resonance frequency, and dimensionless damping ratio of the oscillator respectively.

The micro- and nano-cantilevers discussed in this chapter exhibit more complicated resonance phenomena than the SDOF oscillator described above. Specifically, the cantilevers possess several resonance peaks due to both higher modes of vibration and higher harmonics in the nonlinear electrostatic driving force. The origin and nature of these peaks will be discussed thoroughly in the analysis section of this chapter.

In general, each of the several peaks observed in both the mechanical and electrical cantilever responses resemble the SDOF resonance behavior governed by (2.2), with the notable exception of the first harmonic of the current signal. As an example, the primary
Figure 2.2: The amplitude and phase measured by our HDR system in the second harmonic of the current signal at $\omega_0$ for a silicon cantilever 110 $\mu$m long, 35 $\mu$m wide, and 2 $\mu$m thick. The inset shows the downshift in resonance frequency as the ac voltage is varied from 3V (bottom) to 5 V (top) in 0.5 V increments. This shift is caused by the decrease in the effective spring constant as the ac voltage is increased. The black dashed line is a guide to the eye.
resonance peak, near \( \omega_0 \), as observed in the second harmonic of the electrical (current) signal is presented in Figure 2.2. This figure demonstrates that the resonance frequency and quality factor (Q-factor) of a micro- or nano-cantilever can be accurately determined by examining only its electrical response. Furthermore the inset of Figure 2.2 shows that shifts in the resonance frequency can be observed electrically. In this case the shift is due to changes in the spring constant due to the applied voltage, though shifts in resonance due to other environmental parameters can also be measured for various sensing applications.

6. EXPERIMENTAL FREQUENCY RESPONSES

The mechanical and electrical responses exhibit a variety of resonance peaks, which are evident when the amplitudes of the mechanical (AFM) or electrical (HDR) signals are plotted over a wide range of frequencies. These spectra are useful in differentiating between the mechanical and electrical responses and in examining the unique advantages of HDR.

Typical mechanical and electrical spectra from the experiments described above are presented in Figure 2.3 and Figure 2.4. Both the mechanical and the electrical signals contain several significant harmonic components. Each harmonic of the mechanical signal exhibits a single dominant resonance peak at a driving frequency of, \( \omega_0 / n \), where \( n \) is the order of the harmonic. These are known as super-harmonic resonances [30]. The resonance peak of the first harmonic at \( \omega_0 \), is well defined with a high signal to background ratio (SBR). This is why laser based detection of mechanical resonance has
proven so successful in AFM applications. Parasitic capacitance does not affect mechanical responses. We can also see, in the second harmonic, a small peak driven at $\omega_0$. This is due to nonlinearities in the force on the cantilever, as will be examined analytically later.

The electrical (current) response is noticeably different. The amplitude of the first harmonic increases nearly linearly with applied frequency until it approaches $\omega_0$. This parasitic capacitance, linear in frequency, obscures the true resonant signal and dramatically reduces the SBR. This explains why capacitive detection proved so difficult before the advent of HDR.

The electrical response exhibits several resonance peaks in each harmonic. These peaks are at the super-harmonic frequencies $\omega_0/n$ for $n = 1$ to the number of the harmonic measured, with one peak in the first harmonic, two in the second, and so on. It will be shown that the number and location of these peaks, as well as the mechanical ones can be explained by considering the nonlinear force exerted on the cantilever and the beating of the signals from the induced charge and the cantilever motion.
Figure 2.3: Mechanical response spectrum of a silicon microcantilever measured using AFM based laser reflectometry. The largest peak is in the first harmonic at the primary resonance, $\omega_0 \sim 15.7$ kHz. A super-harmonic resonance is visible in the second harmonic at $\omega_0/2$ and in the third harmonic at $\omega_0/n$. 
Figure 2.4: Electrical (current) response spectrum of a silicon microcantilever measured using HDR. A parasitic capacitance exists in the first harmonic that increases linearly with frequency and obscures the resonance at $\omega \sim 21$ kHz. The higher harmonics do not exhibit significant parasitic effects. Note the first harmonic has a different scale than the higher harmonics.
7. POLAR PLOTS OF RESONANCE

The nature of the resonance peaks can often better be understood by examining their polar representations in which amplitude is plotted versus phase with driving frequency as the parameter. In the HDR polar plots, overlapping curves occur for each resonance peak (primary and super-harmonic) existing in the harmonic spectrum. For instance the second harmonics of the electrical signals for both a micro- and nano-cantilever are presented in Figure 2.5. The resonance frequency may be determined from the polar graph by noting where phase changes most rapidly, e.g. the top of the larger circle in Figure 2.5. Polar plots are often output directly from lock-in amplifiers.

In some cases the polar representation shows that the resonance is no longer circular, but rather is closely approximated by a class of curves known as limaçons. It will later be shown that this results from highly nonlinear systems.
Figure 2.5: A polar plot of a silicon micro-cantilever (300 x 35 x 2 μm$^3$) (diamonds) and a MWNT (triangles). The frequency is a parameter, with the beginning and ending frequencies indicated. The plot for the silicon microcantilever, for which the amplitude is 20 times the indicated scale, illustrates the circle that a resonance displays on a polar plot. The much smaller signal from the MWNT shows the effect of a background signal of the same order of magnitude as the resonant signal. The double-headed arrow indicates the background signal amplitude and phase. One can see that the MWNT resonance shows a nearly complete circle on the polar plot with an offset due to the background signal.
1. RELEVANCE OF LINEARIZED MODELS

Linear models of physical systems are extensively used in science and engineering despite the fact that most real systems are inherently nonlinear. In many cases the effects of the nonlinearities are small, and the system response can accurately be predicted by a linear model. Linear models are invaluable tools because of their simplicity and the availability of analytical solutions.

When the nonlinearities are larger, entirely new phenomena may emerge that do not exist in principle in linear systems. In these cases, the models must be modified to account for the nonlinear effects. Approximate solution methods must be employed since analytical solutions are generally no longer available, cf. Chapter V. However, it is still essential to have a good understanding of the linear models because often the solutions to the nonlinear problems are built up from sequences of linear solutions, as in the various perturbation techniques.

For these reasons, the following chapter is devoted to reviewing two linear models which are fundamental to the response of the electrostatically actuated micro-cantilever: the harmonic oscillator and the Euler-Bernoulli beam.
2. **UNIVERSAL LINEAR OSCILLATOR**

In real physical systems, there is always some degree of nonlinearity. However, in most cases a linear approximation is valid. We will show that for small deflections around a point of stable equilibrium, any potential can be closely approximated by a linear harmonic oscillator \([31]\).

The general form of the Taylor series expansion of the Lagrangian, \(L\), to second order about a point of equilibrium, \(q_{eq} = 0\), is given below

\[
L = A + Bq + C\dot{q} + Dq^2 + Eq\ddot{q} + Fq^2
\]  \(3.1\)

where \(A, B, C, D, E,\) and \(F\) are constants that can be found from the derivatives of the Lagrangian evaluated at the equilibrium position and velocity.

We may identify the point at which the potential is a minimum as a stable equilibrium and a maximum as an unstable equilibrium. For holonomic constraints, the kinetic energy is only a function of velocity, therefore the condition of equilibrium implies that \(\partial L / \partial \dot{q} = 0\). From this, it follows that the coefficient \(B = \partial L / \partial \dot{q} \bigg|_{q_{eq}, \dot{q}_{eq}} = 0\).

Lagrange’s equations for each of the \(N\) generalized coordinates, \(q_k\), are provide in \(3.2\).

\[
\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} = 0 \quad k = 1, \ldots, N
\]  \(3.2\)
Applying this to general Lagrangian in (3.1) gives

\[ \frac{d}{dt} (C + Eq + 2Fq) = 2Dq + E\dot{q} \]  
(3.3)

Therefore the equation of motion (EOM) is

\[ \ddot{q} - \frac{D}{F} q = 0 \]  
(3.4)

We then define the natural frequency to be, \( \omega_n^2 \equiv -D/F \). By scaling time as \( \tau \equiv \omega_n t \), we obtain the EOM of a general simple harmonic oscillator.

\[ \ddot{q} + q = 0 \]  
(3.5)

where the dot indicates differentiation with respect to \( \tau \). This EOM is identical to that given by an equivalent approximate Lagrangian.

\[ L_{approx} = \frac{1}{2} \left( \dot{q}^2 - q^2 \right) \]  
(3.6)

The first term in (3.6) is associated with the kinetic energy and the second with the (negative) potential energy. A quadratic potential results in a force that is linear in the displacement as seen in (3.7). Since this force returns the system toward the equilibrium position, it is said to be a restoring force and is indicative of a stable equilibrium.

\[ F = -\nabla V = -\frac{d}{dq} \left( \frac{q^2}{2} \right) = -x \]  
(3.7)

Note that if the sign in front of \( q^2 \) was instead positive, then the force would tend to push the system away from \( x=0 \), corresponding to an unstable equilibrium.
Therefore, we see that as long as the displacements away from equilibrium are small, then the second order approximation to the Lagrangian is valid. That is near equilibrium, almost all holonomic systems behave like simple harmonic oscillators.

3. SIMPLE HARMONIC OSCILLATOR (SHO)

3.1 Mass-Spring-Damper (MSD) Model

The standard mechanical example of a simple harmonic oscillator (SHO) consists of a mass $m$ attached to a rigid support via a linear spring, with stiffness $k$, as shown in Figure 3.1. There may also be some viscous damping, $b$ (proportional to velocity), shown in the figure as a “dash-pot” and an externally applied force, shown as $F(t)$.

![Figure 3.1: Mass-Spring-Damper (MSD) system](image)

As was shown in the previous section, many systems, not just mechanical ones, behave like an SHO, a notable example being the series RLC (resistance-inductance-capacitance) circuit in electronics.
3.2 Non-Dimensionalized Equation of Motion (EOM)

An application of Newton’s second law \((F=ma)\) to the MSD system results in the following governing differential equation of motion.

\[
mx'' + bx' + kx = F(t)
\]  (3.8)

This equation is classified as a linear, second-order, inhomogeneous differential equation. Linear equations depend on only the first power of the dependent variable and its derivatives \((x, x', x'')\). Nonlinear systems, which we consider at length in Chapter V, depend on higher powers of the dependent variables.

Equation (3.8) is second-order because the highest time derivative is \(x''\), and inhomogeneous because the right-hand side is not equal to zero. Homogeneous systems (right hand side = 0) lead to so called free oscillations because there is no forcing term.

Homogeneous linear differential equations are especially simple to solve, and have the important property of super-position. Super-position entails that any solution can be multiplied by an arbitrary constant or added to another solution, and the result is still a solution. Super-position does not apply in general to nonlinear systems.

Now we simplify the SHO equation of motion by dividing through by \(m\) and defining the natural (angular) frequency, \(\omega_0\), and dimensionless damping ratio, \(\gamma\), as
\[ \omega_0 \equiv \sqrt{\frac{k}{m}} \]  
\[ \gamma \equiv \frac{b}{2\sqrt{km}} \]  

The EOM therefore becomes
\[ \ddot{x} + 2\gamma \omega_0 \dot{x} + \omega_0^2 x = \frac{F(t)}{m} \]  
(3.10)

Now defining the quality factor, \( Q \), as
\[ Q \equiv \frac{1}{2\gamma} \]  
(3.11)

The EOM takes the form
\[ \ddot{x} + \frac{\omega_0}{Q} \dot{x} + \omega_0^2 x = \frac{F(t)}{m} \]  
(3.12)

We now non-dimensionalize the EOM by scaling the variables according to
\[ \tau \equiv \omega_0 t \]
\[ q \equiv \frac{x}{x_0} \]  
(3.13)

Substituting these into (3.12) gives
\[ \omega_0^2 x_0 \frac{d^2 q}{d\tau^2} + \frac{\omega_0^2 x_0}{Q} \frac{dq}{d\tau} + \omega_0^2 x_0 q = \frac{F(t)}{m} \]  
(3.14)

Now dividing through by \( \omega_0^2 x_0 \) gives
\[ \ddot{q} + \frac{1}{Q} \dot{q} + q = \frac{F}{\omega_0^2 x_0 m} \cos(\Omega \tau) \]  
(3.15)
where we take the dot to indicate differentiation with respect to the scaled time, \( \tau \). The forcing is also assumed to be harmonic, i.e. \( F(t) = F \cos(\omega t) \), where the dimensionless driving frequency, \( \Omega \), is the ratio of the angular driving and natural frequencies, \( \Omega \equiv \omega / \omega_0 \).

Finally, we define the length scale as in (3.16). The length scale may be interpreted as the displacement of the mass under a constant force of magnitude \( F \).

\[
x_0 \equiv \frac{F}{\omega_0^2 m} = \frac{F}{k} \tag{3.16}
\]

The general non-dimensional EOM is therefore

\[
\ddot{q} + \frac{1}{Q} \dot{q} + q = \cos(\Omega \tau) \tag{3.17}
\]

3.3 Free Undamped SHO

The EOM for an undamped unforced SHO (3.18) is found by substituting \( Q = \infty \) and \( F = 0 \) into the general form (3.17). This EOM agrees perfectly with that derived in (3.5).

\[
\ddot{q} + q = 0 \tag{3.18}
\]

A general solution to (3.18) is given by

\[
q(\tau) = A \cos(\tau + \phi) \tag{3.19}
\]
where \( A \) is called the amplitude of motion, and \( \phi \) is the phase difference between the driving force and the displacement.

The theory of differential equations states that the number of initial conditions must equal the order of the differential equation. Since all mechanical systems are second-order with respect to time, two initial conditions are required: the initial position and velocity. Any general solution must therefore contain two arbitrary constants. In (3.19), these arbitrary constants are the amplitude and phase.

All subsequent motion is uniquely determined once the initial conditions are specified. This is true regardless of whether the equations are linear or nonlinear. For certain nonlinear systems, the response is extremely sensitive to, yet still completely determined by the initial conditions. This leads to the phenomena of deterministic chaos, in which it is impossible to predict the state of a system at a later time due to uncertainty in the initial conditions.

Using the trigonometric identity \( \cos(\theta + \phi) = \cos(\theta)\cos(\phi) - \sin(\theta)\sin(\phi) \), we may write the general solution in an equivalent form

\[
q(\tau) = A\cos(\tau + \phi) = A\cos(\phi)\cos(\tau) - A\sin(\phi)\sin(\tau) = a\cos(\tau) + b\sin(\tau) \tag{3.20}
\]

where the two arbitrary constants are now \( a \equiv A\cos(\phi) \) and \( b \equiv -A\sin(\phi) \). In this form it is easy to see that the initial conditions are satisfied by
\[ q(0) = a \quad \dot{q}(0) = b \] (3.21)

Also we can readily determine the overall amplitude, \( A \), and phase, \( \phi \), of the response from \( a \) and \( b \) using the vector identity, \( \cos^2(\theta) + \sin^2(\theta) = 1 \).

\[
\sqrt{a^2 + b^2} = \sqrt{A^2 \cos^2(\phi) + A^2 \sin^2(\phi)} = A \sqrt{\cos^2(\phi) + \sin^2(\phi)} = A
\]

\[
\tan^{-1}\left(\frac{b}{a}\right) = \tan^{-1}\left(\frac{A \sin(\phi)}{A \cos(\phi)}\right) = \tan^{-1}\left(\tan(\phi)\right) = \phi
\] (3.22)

We may verify that the trigonometric forms are valid general solutions to the EOM. Substituting (3.19) and (3.20) into (3.18) we have

\[
\ddot{q} + q = -A \cos(\tau + \phi) + A \cos(\tau + \phi) =
-\left( a \cos(\tau) + b \sin(\tau) \right) + \left( a \cos(\tau) + b \sin(\tau) \right) = 0
\] (3.23)

An alternative and generally more useful form of the solution can be formulated using complex numbers (3.24). The equivalence of the trigonometric and exponential forms is given by Euler’s formula, \( \exp(\imath \phi) = \cos(\phi) + \imath \sin(\phi) \). Since the physical displacement cannot be imaginary, only the real part of the complex solution must, in the end, be considered.

\[
q_\tau(\tau) = \tilde{A}_\tau \exp(\imath \tau)
\] (3.24)

where the amplitude, \( \tilde{A}_\tau \), is a complex valued constant. It may appear that there is only one arbitrary constant; however, information about both the initial position and initial velocity are contained in the complex amplitude.

This solution is verified by (3.25) where the imaginary unit satisfies, \( \imath^2 = -1 \).
\[ \ddot{q} + q = i^2 \tilde{\lambda} \exp(i\tau) + \tilde{\lambda} \exp(i\tau) = 0 \quad (3.25) \]

We note that any complex number can be written in either real and imaginary or polar form

\[ \tilde{\lambda} = \text{Re}[\tilde{\lambda}] + i \text{Im}[\tilde{\lambda}] = A \exp(i\phi) \quad (3.26) \]

where \( A \) and \( \phi \) are the real valued amplitude and phase of the complex valued amplitude \( \tilde{\lambda} \). These quantities are equivalent to those in (3.19).

Substituting into the complex solution and taking only the real part gives

\[ q(t) = \text{Re}[\tilde{\lambda} \exp(\tau)] = \text{Re}[\left(\text{Re}[\tilde{\lambda}] + i \text{Im}[\tilde{\lambda}]\right) \cdot (\cos(\tau) + i \sin(\tau))] \]
\[ = \text{Re}[\tilde{\lambda}] \cos(\tau) - \text{Im}[\tilde{\lambda}] \sin(\tau) \quad (3.27) \]

We immediately recognize that the real part represents the initial position and the imaginary part is the (negative) initial velocity.

### 3.3 Free Damped SHO

For a damped free oscillator, the EOM is given by

\[ \ddot{q} + \frac{1}{Q} \dot{q} + q = 0 \quad (3.28) \]

We assume a complex solution of the form

\[ q(\tau) = \tilde{\lambda} \exp(i\alpha \tau) \quad (3.29) \]

Substituting this solution yields
\[
\left( -\alpha^2 + \frac{\alpha}{Q} + 1 \right) = 0 \tag{3.30}
\]

The solutions of this equation correspond to three distinct cases: the under damped solution, the critically damped solution, and the over damped solution.

When \( Q > \frac{1}{2} \) the square root is positive and the solution takes the form of (3.31) with an exponentially decaying amplitude. This is known as the under damped or lightly damped case.

\[
q(\tau) = \tilde{A}_e \exp \left( \frac{-\tau}{2Q} \right) \exp (i\omega' \tau) \tag{3.31}
\]

where the damped natural frequency, \( \omega' \), is given in

\[
\omega' \equiv \sqrt{1 - \frac{1}{4Q^2}} \tag{3.32}
\]

The time response and phase space portraits of a damped harmonic oscillator are given in Figure 3.2. The amplitude is seen to decay exponentially with time as the response approaches the stable equilibrium position at \( q=0 \).
Figure 3.2: Time response and phase space portrait of a free damped harmonic oscillator with $Q = 10$ and $\omega_0 = 1$.

In all micro-cantilever applications considered in this report, the Q-factor is much greater than $\frac{1}{2}$. In fact Q-factors of several thousand have been obtained under vacuum. High Q-factors are essential to sensing applications in which the resonance frequency must be measured. Note that for high Q-factors, the undamped and damped natural frequencies are essentially identical.

The other two cases, critically damped ($Q = \frac{1}{2}$) and over damped ($Q < \frac{1}{2}$) will thus not be considered further, other than to note that the critically damped case is usually desirable when vibrations could have adverse effects.
4. QUALITY FACTOR

As shown in the previous section, the amplitude of an under damped harmonic oscillator decays exponentially. The total energy of the system is proportional to the square of the amplitude. Therefore, the time required for the power to fall off to $1/e$ of its original value is given by (3.33), assuming the damped and natural frequencies are the identical.

$$Power \propto \exp\left(\frac{-\tau}{Q}\right) = \frac{1}{e} \Rightarrow \tau = \omega_0 t = Q$$  \hspace{1cm} (3.33)

Therefore, the quality factor is related to the energy lost per cycle and the decay time of the oscillator. For example, a typical micro-cantilever has a natural frequency of $\omega_0 = 2\pi \cdot 20$ kHz and a $Q$-factor of, $Q = 1000$. Therefore the decay constant is approximately 8 ms. A lock-in integration time of 1 second is common, assuring that the micro-cantilever response has reached steady-state.

5. FORCED DAMPED HARMONIC OSCILLATOR

We now turn our attention to forced linear oscillators which contain an inhomogeneous driving term.

$$\ddot{q} + \frac{1}{Q} \dot{q} + q = \cos(\Omega \tau)$$  \hspace{1cm} (3.34)

The general solution to an inhomogeneous differential equation consists of the sum of a particular (steady-state) solution and a free (transient) solution. The particular solution is a solution to the inhomogeneous problem, and the transient solution to the homogeneous problem.
The steady-state solution should be periodic with the same frequency as the driving force. Therefore, we assume the particular solution takes the form

\[ q_{\text{steady-state}} = a \cos(\Omega \tau) + b \sin(\Omega \tau) \]  

(3.36)

Substituting this into the EOM (3.34) gives

\[
-\Omega^2 \left( a \cos(\Omega \tau) + b \sin(\Omega \tau) \right) + \frac{\Omega}{Q} \left( -a \sin(\Omega \tau) + b \cos(\Omega \tau) \right) + \left( a \cos(\Omega \tau) + b \sin(\Omega \tau) \right) = \cos(\Omega \tau) 
\]

(3.37)

The sines and cosines are linearly independent; therefore, their coefficients must independently equate to zero (3.38). This is the basis of a solution method known as the harmonic balance which is used extensively in solving nonlinear problems.

\[
a\left(1 - \Omega^2\right) + b \frac{\Omega}{Q} = 1 \\
b\left(1 - \Omega^2\right) - a \frac{\Omega}{Q} = 0 
\]

(3.38)

The overall amplitude and phase of the steady state response are given by

\[
A = \sqrt{a^2 + b^2} = \frac{1}{\left( (1 - \Omega^2)^2 + \left( \frac{\Omega}{Q} \right)^2 \right)^{1/2}} \\
\phi = \tan^{-1} \left( \frac{b}{a} \right) = \tan^{-1} \left( \frac{-\Omega}{Q \left(1 - \Omega^2\right)} \right) 
\]

(3.39)
We may now plot the steady state amplitude, Figure 3.3, and phase, Figure 3.4, as functions of the driving frequency. The amplitude is seen to peak at resonance near $\Omega = 1$. However, it should be noted that the resonance peak actually occurs when the derivative of amplitude with respect to frequency is zero (3.40), which is neither the undamped nor damped natural frequencies.

$$\Omega_{resonance} = \sqrt{1 - \frac{1}{2Q^2}} \quad (3.40)$$

In Figure 3.4 the phase is seen to decrease from 0 to $-\pi$ radians ($0^\circ$ to $-180^\circ$) reaching $-\pi/2$ ($-90^\circ$) at resonance. The slope of the phase is greatest at resonance, and greater for higher Q-factors. Resonances are usually determined by phase considerations.

It is very useful to plot the amplitude and phase in polar form as a function of driving frequency, Figure 3.5. The resonance of a linear oscillator forms a single closed circle on a polar plot which is traversed in a clock-wise direction with increasing driving frequency, $\Omega$. It is usually easier discern resonances in the presence of noise by examining the polar representations, and many lock-in amplifiers include this as a standard output format.
Figure 3.3: Steady-state amplitude response of a forced damped harmonic oscillator for two quality factors, $Q = 2$ and $Q = 10$. 
Figure 3.4: Steady-state phase response of a forced damped harmonic oscillator for two quality factors, $Q = 2$ and $Q = 10$. 
Figure 3.5: Polar plot of resonance of a linear damped driven harmonic oscillator. Amplitude and phase are plotted with the driving frequency as a parameter. Resonance appears as a circle on polar plots which is traversed in a clockwise direction with increasing driving frequency. The solid line represents damping at $Q = 2$ and the dashed line for $Q = 10$. 
6. GREEN’S FUNCTIONS

There is an especially general method known as Green’s Function for determining the response of a linear oscillator to an arbitrary external driving force. This method approximates the driving force by an infinite number of impulse functions and integrates to find the response.

The general Green’s function solution is

$$q(t) = \int_{-\infty}^{t} F(t')G(t-t')dt'$$  \hspace{1cm} (3.41)

Where Green’s function, $G(t-t')$, is the response of the system to an impulse force at time, $t'$. The Green’s function for a harmonic oscillator is

$$G(t-t') = \begin{cases} 0, & t \leq t' \\ \sin(t-t'), & t \geq t' \end{cases}$$  \hspace{1cm} (3.42)

As an example, we will consider the response of an undamped forced harmonic oscillator driven at its resonance frequency, $F(t') = \sin(t')$. Using trigonometric identities and integrating, we obtain the following expression for the response.

$$q(t) = \int_{-\infty}^{t} \sin(t')\sin(t-t')dt' = \frac{1}{2}\sin(t) - \frac{1}{2}t\cos(t)$$  \hspace{1cm} (3.43)

As with all inhomogeneous problems, we expect the solution to consist of a steady state and a transient part. In (3.43), the first term on the right is identified as the steady-state solution. The second term, which is the transient solution, is not actually transient in this case, but rather increases linearly and without bound with increasing time. Of course this
situation is not physically realizable since at some amplitude, all systems exhibit some level of damping or nonlinearity that would limit the response. These terms which we will see actually destroy the periodicity of the response are known as secular terms and are of utmost importance to the perturbation methods used in nonlinear analysis, *cf.* Chapter V.

7. EULER-BERNOULLI BEAM THEORY

Euler Bernoulli beam theory is a simplification of the general isotropic theory of elasticity. It is one of the few examples in which a continuous system with distributed parameters has an analytic solution available. In this section, we will derive the governing equation for the transverse vibration of a uniform thin beam in which the deflection is assumed to be a result of the bending moment effects only [32].

A schematic of the thin beam considered in this analysis is presented in Figure 3.6. Though the cantilever geometry is shown, *i.e.* fixed-free boundary conditions, the equation of motion that is derived applies equally for other boundary conditions, such as the fixed-fixed “bridge” arrangement.

Consider the differential element of the beam, located at *x*, presented in Figure 3.7. The shear forces and bending moments acting on the element are shown. As the beam vibrates, this differential element moves up and down vertically. This deflection varies over the length of the beam and also depends on time and is denoted by *z(x,t)*.
As the beam deflects, the element also rotates slightly. However, for small deflections, this rotation is insignificant. Therefore, it is reasonable to neglect the rotational inertia of the beam in deriving the equations of motion. A more thorough model accounting for both rotary inertia and shear effects has been developed and is known as the Timoshenko beam; however, the deviation of this model from the traditional Euler-Bernoulli beam is only significant for large depth to length ratios. Since the micro-cantilever has a very small length to depth ratio, the Euler-Bernoulli assumptions are sufficient.
Summing the moments on the differential element, and neglecting rotational inertia yields

\[ \sum M = M + Vdx - (M + \frac{\partial M}{\partial x} dx) = 0 \]  \hspace{1cm} (3.44)

Therefore the shear force is related to the moment by

\[ V = \frac{\partial M}{\partial x} \]  \hspace{1cm} (3.45)

Applying Newton’s second law to the forces in the vertical direction results in

\[ \sum F = V + \frac{\partial V}{\partial x} dx - V = \rho A dx \frac{\partial^2 z}{\partial t^2} \]  \hspace{1cm} (3.46)

or

\[ \frac{\partial V}{\partial x} dx = \rho A dx \frac{\partial^2 z}{\partial t^2} \]  \hspace{1cm} (3.47)

Combining (3.45) and (3.47), results in

\[ \frac{\partial^2 M}{\partial x^2} = \rho A \frac{\partial^2 z}{\partial t^2} \]  \hspace{1cm} (3.48)

The bending moment is related to the change in slope of the beam’s deflection curve by

\[ M = -EI \frac{\partial^2 z}{\partial x^2} \]  \hspace{1cm} (3.49)

where \( E \) is the elastic modulus of the beam, and \( I \) is the moment of inertia about the beam’s neutral axis. Substituting (3.49) into (3.48), yields the governing equation for the deflection of the beam

\[ EI \frac{\partial^4 z}{\partial x^4} + \rho A \frac{\partial^2 z}{\partial t^2} = 0 \]  \hspace{1cm} (3.50)
A solution to this equation can be sought by the method of separation of variables. The separated solution is given by

\[ z(x, t) = \psi(x) \zeta(t) \quad (3.51) \]

The equation of motion, (3.50), can be written in terms of the separated functions

\[ \frac{-c^2}{\psi(x)} \frac{d^4 \psi(x)}{dx^4} = \frac{1}{\zeta(t)} \frac{d^2 \zeta(t)}{dt^2} \quad (3.52) \]

where the wavelength is given by \( c = \sqrt{EI / \rho A} \).

For these functions of independent variables to be equal, they both must be equal to a constant which we choose to label as \( -\omega_n \). The equations of motion can then be written as two independent ordinary differential equations.

\[ \frac{d^2 \zeta(t)}{dt^2} + \omega_n^2 \zeta(t) = 0 \quad (3.53) \]

\[ \frac{d^4 \psi(x)}{dx^4} - \frac{\omega_n^2}{c^2} \psi(x) = 0 \quad (3.54) \]

The general solutions to these equations are given by (3.55) and (3.56). The spatial function, \( \psi_n(x) \), is known as the mode shape. The temporal function, \( \zeta_n(t) \), is the time response of the tip for a given mode.

\[ \zeta(t) = A \sin \omega_n t + B \cos \omega_n t \quad (3.55) \]

\[ \psi(x) = C_1 \sinh \beta x + C_2 \cosh \beta x + C_3 \sin \beta x + C_4 \cos \beta x \quad (3.56) \]

where the modal parameter \( \beta^4 = \frac{\omega_n^2}{c^2} \).  

51
The coefficients in the above equations are determined from the initial and boundary conditions of the system.

The beam is fixed at the left end; therefore, the first boundary condition is given by

\[ \psi(0) = 0 \]  
(3.57)

The slope of the deflection curve must be equal to zero at the wall therefore, leading to the second boundary equation.

\[ \psi'(0) = 0 \]  
(3.58)

Since there can be no bending moment at the free end of the beam, the third boundary and is given by

\[ \psi''(L) = 0 \]  
(3.59)

The final boundary condition results from the fact that there can be no shear force at the free end of the beam.

\[ \psi'''(L) = 0 \]  
(3.60)

The first three derivatives of the spatial function, \( \psi(x) \) are provided in (3.61)-(3.63)

\[ \psi'(x) = \beta(C_1 \cosh \beta x + C_2 \sinh \beta x + C_3 \cos \beta x - C_4 \sin \beta x) \]  
(3.61)

\[ \psi''(x) = \beta^2 (C_1 \sinh \beta x + C_2 \cosh \beta x - C_3 \sin \beta x - C_4 \cos \beta x) \]  
(3.62)

\[ \psi'''(x) = \beta^3 (C_1 \cosh \beta x + C_2 \sinh \beta x - C_3 \cos \beta x + C_4 \sin \beta x) \]  
(3.63)
Applying the four boundary conditions yields

\[\psi(0) = C_1 \sinh 0 + C_2 \cosh 0 + C_3 \sin 0 + C_4 \cos 0 = C_2 + C_4 = 0 \quad (3.64)\]

\[\psi'(0) = \beta (C_1 \cosh 0 + C_2 \sinh 0 + C_3 \cos 0 - C_4 \sin 0) = \beta (C_1 + C_3) = 0 \quad (3.65)\]

\[\psi''(L) = \beta^2 (C_1 \sinh \beta L + C_2 \cosh \beta L - C_3 \sin \beta L - C_4 \cos \beta L) = 0 \quad (3.66)\]

\[\psi'''(L) = \beta^3 (C_1 \cosh \beta L + C_2 \sinh \beta L - C_3 \cos \beta L + C_4 \sin \beta L) = 0 \quad (3.67)\]

Equations (3.66) and (3.67) may be simplified by substituting (3.64) and (3.65) leading to

\[C_1 (\sinh \beta L + \sin \beta L) + C_2 (\cosh \beta L + \cos \beta L) = 0 \quad (3.68)\]

\[C_1 (\cosh \beta L + \cos \beta L) + C_2 (\sinh \beta L - \sin \beta L) = 0 \quad (3.69)\]

A nontrivial solution to the above two equations exist only if the determinant of the coefficient matrix is zero.

\[
\begin{vmatrix}
\sinh \beta L + \sin \beta L & \cosh \beta L + \cos \beta L \\
\cosh \beta L + \cos \beta L & \sinh \beta L - \sin \beta L
\end{vmatrix}
= (\sinh \beta L + \sin \beta L)(\sinh \beta L - \sin \beta L) - (\cosh \beta L + \cos \beta L)(\cosh \beta L + \cos \beta L) = 0
\quad (3.70)
\]

Expanding (3.70) yields

\[ (\sinh^2 \beta L - \cosh^2 \beta L) - (\sin^2 \beta L + \cos^2 \beta L) - 2 \cos \beta L \cosh \beta L = 0 \quad (3.71)\]
Applying the trigonometric identities to (3.71) leads to the frequency equation (3.72).

The roots of this function correspond to the allowable natural frequencies of the system.

\[ \cos \beta L \cosh \beta L + 1 = 0 \quad (3.72) \]

The natural frequencies of the system can be calculated from

\[ \omega_n = (\beta_n L)^2 \sqrt{\frac{EI}{\rho AL^3}} \quad (3.73) \]

The \( \beta_n L \) values and natural frequencies for the first five modes of vibration are presented in Table 3.1.

**Table 3.1: Roots of frequency equation for fixed-free beam**

<table>
<thead>
<tr>
<th>Mode Number, ( n )</th>
<th>( \beta_n L )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.8751</td>
</tr>
<tr>
<td>2</td>
<td>4.6941</td>
</tr>
<tr>
<td>3</td>
<td>7.8548</td>
</tr>
<tr>
<td>4</td>
<td>10.9955</td>
</tr>
<tr>
<td>5</td>
<td>14.1372</td>
</tr>
</tbody>
</table>

The mode shapes for a fixed-free cantilever beam are given by

\[ \psi(x) = C_1 \left( (\sinh \beta_n x - \sin \beta_n x) - \left( \frac{\cosh \beta_n L + \cos \beta_n L}{\sinh \beta_n L - \sin \beta_n L} \right)(\cosh \beta_n x - \cos \beta_n x) \right) \quad (3.74) \]
The constant $C_I$ is arbitrary. To normalize the shape function of the first mode to -1 at $x=L$, $C_I = (1/2.72445)$. The normalized mode shapes corresponding to the first three modes are shown plotted in Figure 3.8.

![Figure 3.8: First Three Mode Shapes of Fixed-Free Cantilever Beam](image)

We can compare these theoretical free mode shapes with the actual mode shapes of a forced microcantilever as observed under an SEM microscope, Figure 3.9.
Figure 3.9: SEM images of a micro-cantilever vibrating at (a) the fundamental mode and (b) the second mode. The dimensions of this cantilever are $w = 800$ nm, $t = 2$ μm and $l = 40$ μm [33].

8. MODES VS. HARMONICS

Note the distinction between modes and harmonics. Often the term harmonic, which is defined as being an integer multiple of some fundamental frequency is confused with the modes of vibration. The confusion arises because for doubly clamped structures, e.g. violin strings, the frequencies of higher modes of vibration are all integer multiples of the
first mode frequency. Thus for doubly clamped systems, harmonic and modal frequencies are essentially interchangeable. For cantilevers, however, the frequencies of the higher modes are not integer multiples of the first, and thus harmonic and modal frequencies are not equivalent.

In this study, we are generally only concerned with the first mode of vibration ($m = 1$) for several reasons. The first mode has the greatest tip deflection which facilitates both actuation and detection. It will be shown in Chapter IV that the higher modal forcing functions are generally less because of symmetries in the mode shape. Also the lock-in amplifier used has a limited frequency range; therefore, harmonics of higher modes could generally not be measured. However, the second mode has been observed experimentally as shown in Figure 3.10.
Figure 3.10: Second harmonic electrical response of a microcantilever measured using the Harmonic Detection of Resonance (HDR) technique. Visible in this response is the primary resonance of the first mode (a) $\omega_{01}$ ~ 235 kHz, as well as the primary, $\omega_{02}$ ~ 1377 kHz = 5.86 $\omega_{01}$, and first super-harmonic resonances, $\omega_{02}/2$ ~ 688 kHz, of the second mode (b).
CHAPTER IV

SYSTEM MODEL AND EQUATIONS OF MOTION

1. INTRODUCTION

In this chapter, the system model for an electrostatically actuated microcantilever will be presented, and the validity of the “parallel-plate model” for the capacitance and electrostatic force will be specifically addressed. A variational energy approach will be used to derive the governing equations of motion. An assumed modal expansion will be used to reduce the continuous distributed parameter system to a discrete multiple degree of freedom (MDOF) system. It will be shown that the equation governing the temporal motion of each mode is that of a damped linear harmonic oscillator subject to a nonlinear forced excitation.

2. MICRO-CANTILEVER SYSTEM VARIABLES

A schematic of the microcantilever system is presented in Figure 4.1. The deflection of the cantilever is given by \( z(x,t) \). At the tip the deflection is \( \zeta(t) \). The gap distance at time \( t \) separating the micro-cantilever and counter-electrode at the tip is \( d(t) \). The nominal gap distance at zero applied voltage is \( d_0 \). Both an ac voltage and dc bias is applied between the cantilever and counter-electrode.
Figure 4.1: Schematic of the micro-cantilever system used in deriving the equations of motion

3. LUMPED PARAMETER MODEL

The lumped parameter model is the simplest model of the micro-cantilever system. It assumes that the cantilever behaves as single degree of freedom (SDOF) damped driven harmonic oscillator, i.e. a mass-spring-damper (MSD) system. The driving force is taken to be the electrostatic force on a parallel plate capacitor with a variable gap distance equal to the tip separation of the cantilever, d. A diagram of the lumped model is given in Figure 4.2.
Figure 4.2: Schematic of the lumped parameter model of electrostatically-actuated micro-cantilever.

The lumped parameter is quite often used to describe the motion of electrostatically actuated micro-cantilevers. It provides a reasonably accurate first approximation to the response. It will be shown in the remainder of this chapter that the equation governing the time response of each mode of vibration is that of a driven damped SDOF oscillator with a nonlinear electrostatic driving force. In this respect, the lumped parameter model is reasonable.

However, since the lumped model does not take into account beam bending or variations from the parallel plate capacitance model, it is not an especially accurate representation of the beam response. For instance, it has been shown that the lumped parameter model underestimates the pull-in voltage significantly [34], Figure 4.3.
3.1 Mass-Spring-Damper (MSD) Model

A single degree of freedom (SDOF) linear mass-spring-damper (MSD) model for the beam is assumed in the lumped parameter model. The response of a forced SDOF linear harmonic oscillator was discussed thoroughly in the previous chapter. The equation of motion for the forced linear oscillator is given by

\[
m\dddot{x}(t) + b\ddot{x}(t) + k\dot{x}(t) = F(t)
\]  

(4.1)

The mass, \(m\), damping, \(b\), stiffness, \(k\), and forcing, \(F(t)\), are not strictly defined in the lumped model. They will later be identified as the modal parameters and generalized force in the variational energy derivation of the EOM.
3.2 Parallel-Plate Capacitance Model

A parallel-plate capacitor consists of two flat conductors separated by a thin layer of dielectric or air. When a voltage difference is applied, equal and opposite charges accumulate on each plate. The relationship of charge and voltage for a capacitor is given by

$$Q = CV$$

(4.2)

where \( Q \) is the charge on the plates (\( Q \) and \(-Q\) respectively), \( C \) is the capacitance, and \( V \) is the voltage or potential difference.

For a parallel plate capacitor, the gap distance is assumed to be much smaller than the plate area. It can be shown using Gauss’s law and symmetry arguments that the electric field is then a constant in the region between the plates and negligible elsewhere [35]. This leads to an especially simple expression for the parallel-plate capacitance

$$C_{pp} = \frac{Q}{V} = \frac{\varepsilon A}{d} = \frac{\varepsilon A}{d_0 - \zeta(t)}$$

(4.3)

where \( V \) is the voltage, \( d = d_0 - \zeta(t) \) is the gap distance, and \( \varepsilon \) is the permittivity of the dielectric separating the plates of the capacitor.

The energy required to move an amount of charge, \( dq \), from one plate to the other is

$$dW = Vdq = \frac{q}{C} dq$$

(4.4)

Therefore, the total energy required to store an amount of charge, \( Q \), is
The electrostatic force which drives the cantilever motion is found by taking the positive derivative of the energy stored in the electrostatic field (4.5). Force is, in general, the negative derivative of energy with respect to distance. The positive sign here comes from the fact that the battery does work in moving charge off the cantilever in order to maintain a constant voltage [36].

\[
F_e(t) = \frac{d}{d\zeta} \left( \frac{1}{2} CV^2 \right) = \frac{1}{2} \frac{dC}{d\zeta} V(t)^2
\]  \hspace{1cm} (4.6)

A general capacitance can be expanded in a Taylor series about, \( \zeta = 0 \), (undeflected or nominal position).

\[
C(\zeta) = C_0 + C_1\zeta(t) + \frac{1}{2} C_2\zeta^2(t) + ... + \frac{1}{n!} C_n\zeta^n(t)
\]  \hspace{1cm} (4.7)

where the capacitive coefficients, \( C_n \), are given by

\[
C_n = \frac{\partial^n C}{\partial \zeta^n} \bigg|_{\zeta=0}
\]  \hspace{1cm} (4.8)

Therefore the electrostatic force for a general capacitor is given by

\[
F_e(t) \approx \frac{1}{2} \left( C_1 + C_2\zeta(t) + \frac{1}{2} C_3\zeta^2(t) + ... \right) V(t)^2
\]  \hspace{1cm} (4.9)

Now the capacitive coefficients for a parallel plate capacitor are
\[ C_{pp,0} = \frac{\varepsilon A}{d_0} \quad C_{pp,1} = \frac{\varepsilon A}{d_0^2} \]
\[ C_{pp,2} = \frac{2\varepsilon A}{d_0^3} \quad C_{pp,3} = \frac{6\varepsilon A}{d_0^4} \]  

(4.10)

Therefore the electrostatic force experienced by a parallel-plate capacitor is

\[ F_{pp,e}(t) \approx \frac{\varepsilon A}{2d_0^2} \left( 1 + 2 \left( \frac{\zeta(t)}{d_0} \right) + 3 \left( \frac{\zeta(t)}{d_0} \right)^2 + \ldots \right) V(t)^2 \]  

(4.11)

The lumped parameter EOM then may be simplified to

\[ m \ddot{\zeta}(t) + b \dot{\zeta}(t) + k \zeta(t) = \frac{\varepsilon A}{2d_0^2} \left( 1 + 2 \left( \frac{\zeta(t)}{d_0} \right) + 3 \left( \frac{\zeta(t)}{d_0} \right)^2 + \ldots \right) V(t)^2 \]  

(4.12)

The term in front of the parentheses on the right hand side is the electrostatic force on a parallel plate capacitor with a gap separation \( d_0 \). Since this term does not depend on the motion of the cantilever it is called the static capacitance. The beam deflection is always smaller than the nominal gap separation, \( z < d_0 \), under normal operating conditions; therefore the higher order (dynamic capacitance) terms are all much smaller than the static term. The static (parasitic) capacitance is the source of the difficulty for traditional capacitive detection methods, which is avoided in HDR.

3.3 Applied Voltage

The applied voltage consists of an ac term at the driving frequency, \( \Omega \), and a dc bias. The electrostatic force, proportional to the voltage squared, may be expanded using trigonometric identities (4.13). In this form, it is evident that the force consists of a constant term, which shifts the average deflection of the cantilever towards the counter-
electrode by a distance $\delta$, and two harmonic terms at the driving frequency, $\Omega$, and twice
the driving frequency, $2\Omega$.

$$
V^2(t) = \left(V_{dc} + V_{ac} \cos(\Omega t)\right)^2 = V_{dc}^2 + \frac{1}{2}V_{ac}^2 + 2V_{dc}V_{ac} \cos(\Omega t) + \frac{1}{2}V_{ac}^2 \cos(2\Omega t)
$$

(4.13)

From this we expect to see normal resonance phenomena when the applied frequency is equal to the natural frequency of the cantilever ($\Omega = \omega_0$) and additionally, due to the second harmonic term in the force, we expect a super-harmonic resonance peak when the driving frequency is half the natural frequency ($\Omega = \omega_0/2$).

4. NON-DIMENSIONALIZED EOM

Following a procedure analogous to that used in the last chapter for a linear harmonic oscillator, we seek to non-dimensionalized, and therefore generalize, the nonlinear EOM for the lumped parameter model with a general capacitance.

$$
m\ddot{\zeta}(t) + b\dot{\zeta}(t) + k\zeta(t) = \frac{1}{2}\left(C_1 + C_2\zeta(t) + \frac{1}{2}C_3\zeta^2(t) + \ldots\right)V^2(t)
$$

(4.14)

We begin by dividing through by $m$ and defining the natural frequency to be $\omega_0^2 \equiv k/m$, and the quality factor to be $Q \equiv \sqrt{km/b}$.

$$
\frac{d^2\zeta}{dt^2} + \frac{\omega_0}{Q} \frac{d\zeta}{dt} + \omega_0^2 \zeta = \frac{V^2(t)}{2m}\left(C_1 + C_2\zeta(t) + \frac{1}{2}C_3\zeta^2(t) + \ldots\right)
$$

(4.15)

Now scaling time and distance according to $\tau \equiv \omega_0 t$ and $q \equiv \zeta/d_0$. 66
\[
\omega_0^2 d_0 \frac{d^2 q}{d\tau^2} + \frac{\omega_0^2 d_0}{Q} \frac{dq}{d\tau} + \alpha_0^2 d_0 q = \frac{V^2(t)}{2m} \left( C_1 + C_2 d_0 q + \frac{1}{2} C_3 (d_0 q)^2 + \ldots \right) \tag{4.16}
\]

Now dividing through by \(\omega_0^2 d_0\) and factoring out \(C_1\).

\[
\frac{d^2 q}{d\tau^2} + \frac{1}{Q} \frac{dq}{d\tau} + q = \frac{C_1 V^2(t)}{2m\omega_0^2 d_0} \left( 1 + \left( \frac{C_2}{C_1} \right) d_0 q + \frac{1}{2} \left( \frac{C_3}{C_1} \right) (d_0 q)^2 + \ldots \right) \tag{4.17}
\]

We now define a small dimensionless parameter, \(\tilde{\varepsilon}\) (not to be confused with the permittivity).

\[
\tilde{\varepsilon} \equiv \frac{C_1 V^2}{2m\omega_0^2 d_0} \tag{4.18}
\]

where we have identified \(V^2\) with the constant terms in the voltage expansion (4.13)

\[
V^2 = V_{dc}^2 + \frac{1}{2} V_{ac}^2 \tag{4.19}
\]

We now note that the static electrostatic force (when \(\zeta = q = 0\)) is

\[
F_{e,\text{static}} = \frac{C_1 V^2}{2} \tag{4.20}
\]

And consequently, the static displacement of the mass-spring due to this force is

\[
\zeta_0 = \frac{F_{e,\text{static}}}{k} = \frac{C_1 V^2}{2m\omega_0^2} \tag{4.21}
\]

From this it is clear that \(\tilde{\varepsilon}\) must be less than 1 since the parameters of the system must be such that the static displacement, \(\zeta_0\), is less than the nominal gap distance, \(d_0\), otherwise crashing would occur. We also see from (4.22) that \(\tilde{\varepsilon}\) is dimensionless.
\[ \tilde{\varepsilon} = \frac{\varepsilon_0}{d_0} \]  

(4.22)

Assuming a parallel-plate capacitance and using typical system parameters with \( d_0 = 15 \mu m \), it can be shown that \( \tilde{\varepsilon}_{typical} \approx 0.0015 \).

The non-dimensionalized EOM can now be rewritten

\[ \ddot{q} + \frac{1}{Q} \dot{q} + q = \tilde{\varepsilon} f(\tau) \left( \tilde{\alpha} + \tilde{\beta} q + \tilde{\gamma} q^2 + ... \right) \]  

(4.23)

Where the dot indicates differentiation with respect to the scaled time, \( \tau \), and the dimensionless system constants \( \tilde{\alpha} \), \( \tilde{\beta} \), and \( \tilde{\gamma} \) are given by

\[ \tilde{\alpha} \equiv 1 \]
\[ \tilde{\beta} \equiv \frac{C_2 d_0}{C_1} \]
\[ \tilde{\gamma} \equiv \frac{C_3 d_0^2}{2C_1} \]  

(4.24)

For a parallel-capacitance the dimensionless system parameters are

\[ \tilde{\alpha} = 1 \]
\[ \tilde{\beta} = \left( \frac{2 \varepsilon A}{d_0^3} \right) \left( \frac{d_0^2}{\varepsilon A} \right) d_0 = 2 \]
\[ \tilde{\gamma} = \left( \frac{1}{2} \right) \left( \frac{6 \varepsilon A}{d_0^4} \right) \left( \frac{d_0^2}{\varepsilon A} \right) d_0^2 = 3 \]  

(4.25)

The normalized forcing function is given by
\[ f(\tau) = 1 + \Lambda_1 \cos(\tilde{\Omega} \tau) + \Lambda_2 \cos(2\tilde{\Omega} \tau) \]

\[ \Lambda_1 = \frac{2V_{dc} V_{ac}}{V_{dc}^2 + V_{ac}^2/2} \]

\[ \Lambda_2 = \frac{V_{ac}^2}{2V_{dc}^2 + V_{ac}^2} \]  \hspace{1cm} (4.26)

Where the dimensionless driving frequency is \( \tilde{\Omega} = \Omega/\omega_0 \).

A parallel plate lumped parameter model possesses the especially simple EOM, cf. (4.12)

\[ \ddot{q} + \frac{1}{Q} \dot{q} + q = \tilde{\varepsilon} f(\tau)(1 + 2q + 3q^2 + ...) \] \hspace{1cm} (4.27)

The dimensionless forms of the governing EOMs are convenient for the approximate solution methods presented in the following chapter. Specifically, as either the voltage is decreased or the gap distance is increased, the parameter \( \tilde{\varepsilon} \) goes to zero. This corresponds to the nonlinearities and all forcing being removed from the system.

5. HAMILTON’S PRINCIPLE

The Hamiltonian formulation of mechanics can be shown to be equivalent to both the Lagrangian and Newtonian formulations [37]; however, it is usually much easier to apply in general cases. Consequently, it is used extensively in quantum and field theories, e.g. Feynman path-integrals. Also, Hamilton’s principle provides a direct method of determining the correct EOMs for an electrostatically actuated microcantilever. This is the project of the remainder of this chapter.
Hamilton’s principle states that the trajectory followed by a system is that which minimizes (more precisely makes stationary) the time integral of the Lagrangian, otherwise known as the action (4.28).

\[ \delta S = \delta S = \int_{t_i}^{t_f} \delta L dt = \int_{t_i}^{t_f} \frac{\delta L}{\delta q} \delta q(t) dt = 0 \]  

(4.28)

Non-conservative forces, such as friction, present a difficulty in standard Hamiltonian formulations. Often, the appropriate frictional forces are simply added to the equation of motion once they have been found using the variational approach. This is the approach taken in the present derivation. However, we may instead add a dissipative energy term to the action integral which will lead to the same equations of motion, while retaining some generality [34].

\[ \delta \int_{t_i}^{t_f} (T - U + W_{nc}) dt = 0 \]  

(4.29)

6. ENERGY EXPRESSIONS

Consider the electrostatically actuated micro-cantilever shown in Figure 4.1. In order to apply Hamilton’s principle to determine the equations of motion, expressions for the kinetic, potential, and non-conservative energies must be obtained. In each case the total energy is found by summing (integrating) the energy contributions from all of the differential elements along the length of the beam.

In the following analysis, the same assumptions that were used to derive the Euler-Bernoulli beam equations in the previous chapter will again be applied. The cross-
sectional area, flexural rigidity, and mass density are assumed constants along the length of the beam. The beam is also assumed thin and deflections small, so that rotational and shear effects can be neglected, and only motion in the transverse, \( z \), direction is considered significant.

6.1 Kinetic Energy

The total kinetic energy of the beam is given by

\[
T = \frac{1}{2} \int_0^L \rho A \left( \frac{\partial z}{\partial t} \right)^2 dx
\]  

(4.30)

where \( \rho \) is the mass density, \( A \) is the cross-sectional area, and \( z \) is the transverse displacement of the micro-cantilever beam, which varies with both time and distance along the beam, \( z = z(x, t) \).

6.2 Mechanical Potential Energy

The mechanical strain energy of a differential element experiencing a bending moment, is given by (4.31). It is assumed that the normal stress varies linearly with distance from the neutral axis and the material obeys Hooke’s law.

\[
dU_m = \frac{EI}{2} M^2 dx
\]  

(4.31)

The bending moment in a beam is related to the beam curvature as shown in (3.49). Therefore, the total mechanical strain energy in the cantilever beam is
\[ U_m = \frac{1}{2} \int_0^L EI \left( \frac{\partial^2 z}{\partial x^2} \right)^2 dx \]  

(4.32)

Where \( E \) is the Young’s modulus (bulk modulus), and \( I \) is the area moment of inertia of the beam.

6.3 Electrostatic Potential Energy

The most general form for the energy stored in the electrostatic field, \( U_e \), is

\[ U_e = \frac{\varepsilon_0}{2} \iiint_{All\ Space} E^2 \, dx \, dy \, dz \]  

(4.33)

Where \( \varepsilon_0 \) is the permittivity of free space, and \( E \) is the electric field which each point in space may be determined once the electric potential, \( \varphi \), is known.

\[ \bar{E} = -\nabla \varphi \]  

(4.34)

The electric potential, in turn, is given by Poisson’s equation (in the Coulomb gauge).

\[ \Delta \varphi = -\frac{\rho_s}{\varepsilon_0} \]  

(4.35)

Thus, in order to determine the electrostatic energy of an arbitrary distribution of charge, (4.35) must be solved. There are many ways to do this. For simple geometries, Gauss’s law may be employed. For more complicated arrangements, a numerical integration may be performed.
For a general capacitor, the electrostatic potential energy is given by \( U_e = CV^2/2 \).

However, again, in order to determine the capacitance of a general arrangement, Poisson’s equation must be solved. Since this is a computationally intensive process, it is desirable to assume an approximate capacitance.

In this case we assume that capacitance of each differential area element can be considered separately and summed to find the total energy. Using the general Taylor series form of the capacitance, the electrostatic potential energy is given by

\[
U_e = \int_0^L \frac{1}{2} CV^2 dx = \frac{1}{2} \int_0^L \left( C_0 + C_1 \frac{z}{2} + \frac{1}{2} C_2 z^2 + \frac{1}{6} C_3 z^3 + \ldots \right) V^2 dx \tag{4.36}
\]

6.4 Non-Conservative Forces

There are many damping mechanisms that may affect the motion of a micro-cantilever. These have been discussed thoroughly elsewhere; however, they in general depend on the pressure regime (intrinsic, molecular, and viscous) and gap distance, e.g. squeeze-film damping. Often the damping is also a source of nonlinearity in a problem.

It is reasonable in our case to assume a viscous damping force (proportional to velocity). For non-conservative functions that only depend on the velocity we may define the Rayleigh dissipation function.

\[
F \equiv \frac{1}{2} \sum_{i,j} b_{ij} \dot{q}_i \dot{q}_j \tag{4.37}
\]
The generalized force in the Lagrangian formulation is then

$$\frac{\partial F}{\partial \dot{q}_i} = \sum_j c_j \dot{q}_j = -Q_i \quad (4.38)$$

Therefore Lagrange’s equations may be written as

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} + \frac{\partial F}{\partial \dot{q}_i} = 0 \quad (4.39)$$

This becomes the action functional in Hamilton’s principle (4.28).

$$\int_{t_0}^{t_f} \left[ \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} + \frac{\partial F}{\partial \dot{q}_i} \right] \delta q_i dt = 0 \quad (4.40)$$

We identify the last term as the virtual work due to the non-conservative frictional forces

$$\delta W_{nc} = \frac{\partial F}{\partial \dot{q}_i} \delta q_i = -b \dot{q}_i \delta q_i \quad (4.41)$$

7. VARIATIONAL EQUATIONS OF MOTION

Substituting the energy expression (4.30), (4.32), and (4.36), into the variational equation (4.29), and ignoring the non-conservative forces for now

$$\delta \frac{1}{2} \int_{t_0}^{t_f} \left[ \rho A \left( \frac{\partial z}{\partial t} \right)^2 + EI \left( \frac{\partial^2 z}{\partial x^2} \right)^2 - \frac{1}{2} \left[ C_0 + C_1 z + \frac{1}{2} C_2 z^2 + \frac{1}{6} C_3 z^3 + \ldots \right] V^2 \right] dx dt = 0 \quad (4.42)$$
Integrating by parts and applying variational constraints

\[ \int_{t_1}^{t_2} \int_{0}^{L} -\rho A \frac{\partial^2 z}{\partial t^2} - EI \frac{\partial^4 z}{\partial x^4} + \frac{1}{2} \left[ C_1 + C_2 z + \frac{1}{2} C_3 z^2 + \ldots \right] V^2 \delta z \, dx \, dt \]

\[ + \left[ \int_{t_1}^{t_2} \left[ \left( EI \frac{\partial^2 z}{\partial x^2} \right) \frac{\partial^2 \delta z}{\partial x^2} + \left( EI \frac{\partial^3 z}{\partial x^3} \right) \frac{\partial \delta z}{\partial x} \right] \right] dt - \int_{t_1}^{t_2} \left[ \left( EI \frac{\partial^3 z}{\partial x^3} \right) \frac{\partial \delta z}{\partial x} \right] dt = 0 \]  

(4.43)

Since the virtual displacements, \( \delta z \), are arbitrary, (4.43) can only be zero if each integrand independently equates to zero. The first integral yields the EOM of the beam from variational methods. This is the same as the EOM for a Euler-Bernoulli beam derived using Newtonian methods in the previous chapter, except there is now a nonlinear forcing term which is immediately identified as the electrostatic force (per unit length) on a general capacitor, cf. (4.9).

\[ \rho A \frac{\partial^2 z}{\partial t^2} + EI \frac{\partial^4 z}{\partial x^4} = \frac{1}{2} \left[ C_1 + C_2 z + \frac{1}{2} C_3 z^2 + \ldots \right] V^2(t) \]  

(4.44)

The second and third integrals in (4.43) yield the boundary conditions for the beam.

\[ \begin{align*}
\left( EI \frac{\partial^2 \delta w}{\partial x^2} \right) \delta \left( \frac{\partial w}{\partial x} \right) &= 0 \\
\left( EI \frac{\partial^3 \delta w}{\partial x^3} \right) \delta w &= 0 \\
\end{align*} \]

at \( x = 0 \) and \( x = L \)  

(4.45)

The first is the boundary condition requiring that either the slope or the moment on the beam is zero at the boundaries. The second is the condition that either the deflection or the shear force is zero at the boundaries. Of course, for a cantilever beam, the deflection and slope is zero at \( x = 0 \) and the moment and shear is zero at \( x = L \).
8. MODAL EXPANSION AND DISCRETE EOMS

A modal expansion is performed in order to transform the continuous distributed parameter system into a discrete multiple degree of freedom (MDOF) system. This method is often referred to as the “Assumed modes method” and falls under the general theory of the Galerkin Procedure.

We assume that the response can be expanded as a sum of the undamped unforced mode shapes. The validity of this expansion is guaranteed since the modes form an independent basis set which can be proved using Sturm-Liouville theory.

\[
z(x,t) = \sum_{m=1}^{M} \zeta_m(t)\psi_m(x)
\]  

(4.46)

Where \( \psi_m(x) \) are the free undamped mode shapes of the cantilever, and \( \zeta_m(t) \) is the modal participation factor which represents the time response of the mode at some arbitrary position, e.g. the end of the beam, \( x = L \).

Substituting the modal expansion (4.46) into the equation of motion (4.44).

\[
\sum_{m=1}^{M} \left[ \rho A \psi_m(x) \dddot{z}_m(t) + EI \psi_m'''(x)\zeta_m(t) \right] = \frac{V^2(t)}{2} \left( C_1 + C_2 \left( \psi_m \zeta_m \right) + \frac{1}{2} C_3 \left( \psi_m \zeta_m \right)^2 + \ldots \right)
\]  

(4.47)
The natural modes possess the following ortho-normality conditions

\[
\int_0^L \psi_i(x) \psi_j(x) \, dx = 0, \quad \text{for } i \neq j
\]

\[
\int_0^L \psi_i''(x) \psi_j''(x) \, dx = 0, \quad \text{for } i \neq j
\]

(4.48)

Multiplying (4.47) by \( \psi_i(x) \) and integrating over the length of the beam.

\[
\sum_{m=1}^M \left[ \dot{\zeta}_m(t) \int_0^L \rho A \psi_i \cdot \psi_m \, dx + \zeta_m(t) \int_0^L E l \psi_i \cdot \psi_m'''' \, dx \right] = \sum_{m=1}^M \frac{V^2(t)}{2} \left( C_1 \int_0^L \psi_i \cdot \psi_m \, dx + C_2 \zeta_m \int_0^L \psi_i \cdot \psi_m \, dx + \frac{1}{2} C_3 \zeta_m^2 \int_0^L \psi_i \cdot \psi_m^2 \, dx + \ldots \right)
\]

(4.49)

The following system nonlinear differential equations is equivalent to (4.49), with the addition of a viscous damping term. In this formulation, the continuous micro-cantilever has been reduced to a discrete MDOF system, with one degree of freedom for each mode of vibration. However, we will see that in general these modal equations may be coupled.

\[
[m] \{ \ddot{\zeta}(t) \} + [b] \{ \dot{\zeta}(t) \} + [k] \{ \zeta(t) \} = [F_0] + [F_1] \{ \zeta(t) \} + [F_2] \{ \zeta^2(t) \} + \ldots
\]

(4.50)

Since the first integral in (4.49) obeys the first ortho-normality condition (4.48), the mass matrix, \([m]\), and damping matrix, \([b]\), are diagonal \(M \times M\) matrices. Integration by parts may be applied twice to put the second integral in the form of the second ortho-normality condition, and thus the stiffness matrix, \([k]\), is also a diagonal \(M \times M\) matrix.

The values of these matrices are given by (4.51). The diagonal elements are referred to as the modal parameters.
\[ m_{ij} = \begin{cases} \int_0^L \rho(x) A(x) \left[ \psi_i(x) \right]^2 dx, & i = j \\ 0, & i \neq j \end{cases} \]

\[ b_{ij} = \begin{cases} \int_0^L b(x) \left[ \psi_i(x) \right]^2 dx, & i = j \\ 0, & i \neq j \end{cases} \]

\[ k_{ij} = \begin{cases} \int_0^L E(x) I(x) \left[ \psi_i''(x) \right]^2 dx, & i = j \\ 0, & i \neq j \end{cases} \]

(4.51)

For instance the first modal mass of a cantilever beam is

\[ m_i = \int_0^L \rho(x) A(x) \left[ \psi_i(x) \right]^2 dx = 0.24 \rho A L \approx \frac{M_{\text{Total}}}{4} \]

(4.52)

On the right-hand side of (4.49), the second integral does obey the ortho-normality condition; however, the first and third integrals do not. Therefore the generalized force matrices, \([F_0] (M \times 1), [F_i] (M \times M)\) is diagonal, whereas \([F_i] (M \times M)\), are in general not diagonal. Therefore, the nonlinearities may couple the modes of vibration. This leads to differential equations which must be solved simultaneously which significantly increases the complexity of the calculation. The values of the generalized force matrices are given in:
\[ F_{0,i} = \frac{C_1 V^2(t)}{2} \int_0^L \psi_i(x) \, dx \]

\[
F_{i,j} = \begin{cases} 
\frac{C_2 V^2(t)}{2} \int_0^L \left[ \psi_i(x) \right]^2 \, dx, & i = j \\
0, & i \neq j 
\end{cases} 
\]

\[ F_{n,j} = \frac{C_n V^2(t)}{2n!} \int_0^L \psi_i(x) \left[ \psi_j(x) \right]^n \, dx \]

We see here that since the higher modes are relatively symmetric above and below the cantilever, their modal generalized forces (specifically \[ F_0 \]) will be much less significant than the first mode. This is partly the reason why the higher modes of vibration are rarely observed. Of course, the principle reason is that the excitation frequencies are generally not high enough to excite these modes.

9. **UNCOUpled NON-DIMENSIONALIZED MODAL EOMs**

If we assume that the off-diagonal terms in the force matrices are negligible, then we may uncouple the modes of vibration. In this case each mode behaves as a damped harmonic oscillator subject to a nonlinear excitation. The uncoupled modal EOMs are given by

\[
\left[ m \ddot{\zeta}(t) + b \dot{\zeta}(t) + k \zeta(t) = F_0 + F_1 \dot{\zeta}(t) + F_2 \zeta^2(t) + \ldots \right]_m
\]

\[
= \left[ \frac{C_1 V^2(t)}{2} \left( \dot{\alpha} + \dot{\beta} \left( \frac{C_2}{C_1} \right) \zeta(t) + \frac{1}{2} \frac{C_3}{C_1} \zeta^2(t) + \ldots \right) \right]_m
\]

(4.54)
Where the parameters $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ are the mode shape integrals.

$$\hat{\alpha} \equiv \int_{0}^{L} \psi_m(x) dx$$
$$\hat{\beta} \equiv \int_{0}^{L} \psi^2_m(x) dx$$
$$\hat{\gamma} \equiv \int_{0}^{L} \psi^3_m(x) dx$$

(4.55)

Now we may non-dimensionalize the uncoupled modal equations using the same scaled variables as for the lumped model

$$\ddot{q} + \frac{1}{Q} \dot{q} + q = \tilde{e} f(\tau) \left( \hat{\alpha} + \hat{\beta} q + \hat{\gamma} q^2 + \ldots \right)$$

(4.56)

where the perturbation parameter, $\tilde{e}$, is again given by

$$\tilde{e} \equiv \frac{C_1 V^2}{2 m \omega^2 d_0}$$

(4.57)

And the dimensionless system constants are given by

$$\tilde{\alpha} \equiv \hat{\alpha}$$
$$\tilde{\beta} \equiv \hat{\beta} \frac{C_2 d_0}{C_1}$$
$$\tilde{\gamma} \equiv \hat{\gamma} \frac{C_3 d_0^2}{2C_1}$$

(4.58)

These are identical to the lumped model except that they are scaled by the mode integrals (4.55).
10. VALIDITY OF THE PARALLEL PLATE MODEL

The parallel plate model is chosen for its simplicity; however, it does not take into account that the microcantilever and counter-electrode may not be flat and parallel, that the electric field may not be uniform in the gap region, and that the electric field may exist outside the plates (fringing field effects). All of these possibilities call into question the validity of the parallel plate model; however, it can be shown that for most typical configurations these differences are minor, and the parallel plate model is valid.

The effect of geometry on capacitance may be examined by considering two well known configurations shown in Figure 4.4 the parallel-plate and the charged parallel cylindrical capacitors.

![Parallel-plate and parallel-cylinder capacitance models](image)

**Figure 4.4:** Parallel-plate and parallel-cylinder capacitance models
The capacitances of these arrangements are well known and are given by

\[
C_{pp} = \frac{\varepsilon b L}{d} \quad \quad C_{pw} = \frac{\varepsilon \pi L}{\cosh^{-1}(D/2a)}
\]  

(4.59)

Where \( \varepsilon \) is the permittivity of the dielectric, \( b \) is the width of the parallel plate capacitor, \( d \) is the parallel plate gap distance, \( D \) is the distance between centers of the cylinders, \( a \) is the cylinder radius, and \( L \) is the lengths of both the capacitors in the direction of the page.

We may immediately compare these two formulae by noting that for the cylinders to be the same distance apart as the parallel plates, \( d = D - 2a \), and to be the same width, \( b = 2a \). We then plot both as shown in Figure 4.5. For large gap distance to width ratios, the two capacitances are essentially identical; however, for small separations, the two configurations vary greatly.

However, since the motion of micro-cantilever depends only on the derivative of the capacitance (through the electrostatic force), the two geometries produce similar results down to approximately, \( d/b = 0.25 \).

For a typical cantilever, the width is approximately \( b = 35 \mu m \) and the nominal gap distance \( d = 10 \mu m \) with an average amplitude of vibration of \( \Delta d \approx 1 \mu m \). The slopes at \( d/b = 0.3 \) are similar; thus, differences in the micro-cantilever response due to variations in geometry are generally negligible for small displacements.
Figure 4.5: Parallel-plate and parallel-cylinder capacitances as functions of gap separation
CHAPTER V

APPROXIMATE SOLUTION METHODS TO NONLINEAR PROBLEMS

1. INTRODUCTION

The purpose of this chapter is to survey the various methods available for determining an approximate solution to a general nonlinear problem. It is usually advantageous to seek an exact, analytical solution; unfortunately, for most nonlinear problems no such solutions are known to exist. In these cases, an approximate solution method must be utilized. In this chapter we will discuss several approximate approaches including numerical integration, straightforward perturbation, the Lindstedt-Poincaré method, the method of multiple scales, averaging, and the method of harmonic balance.

An analytical solution, also referred to as an exact or closed form solution, is one expressed solely in terms of elementary functions in “algebraic form without the necessity of introducing numerical values for parameters or initial conditions” [38]. The general solutions to the harmonic oscillator and Euler-Bernoulli beam derived in the previous chapter are examples of analytical solutions. Once an analytical solution is found, any particular set of parameters can easily be inserted, allowing the entire possible range of solutions to be explored.

Perhaps the most straightforward of the approximation methods is numerical integration. In this method, the response of the system is calculated at discrete time intervals,
effectively reducing the problem to finding the solution of a set of algebraic equations. Numerical integration is extremely powerful because a certain algorithm can generally handle a wide variety of different problems.

However, numerical solutions also have their limitations since specific values for the system parameters and initial conditions must be chosen. Therefore, unlike the analytical solutions, more general features may not be readily apparent. Typically many numerical simulations must be performed in order to ascertain the overall system behavior, and even then there is no guarantee that significant phenomena have not been overlooked.

The other approximate methods considered herein represent a type of middle ground between an exact analytic solution and a strictly numerical integration. Perturbation theory, for instance, is based on the principle that if the nonlinearities are small then the nonlinear solution should be in some sense “close” to the linear solution. To this end, the solutions are assumed to take the form of a power series in some small parameter associated with the nonlinearities. Once this ansatz has been made, an iterative procedure is used to determine closed form corrections to the linearized solutions.

The accuracy of the perturbation methods depends on the strength of the nonlinearity and the order of the terms retained in the series solutions. Of course, an increase in mathematical complexity is associated with higher order terms. In this sense, the perturbation methods are similar to strictly numerical integrations; greater accuracy requires more computational cycles. However, unlike numerical integration, the
perturbation solutions are “built up” from closed form, analytical expressions. Therefore, it is often easier to determine overall response characteristics from perturbation solutions than with numerical solutions alone, particularly the effect of increasing the nonlinearity. It should be noted, that though perturbation solutions are of closed form they are not exact like analytical solutions, since in practice the series is always truncated.

As a means of comparison, each of the approximate methods will be used to seek a solution to an especially interesting and well known nonlinear problem: the Duffing oscillator. As will become apparent, the behavior of the Duffing oscillator is also intimately connected with the response of an electrostatically actuated micro-cantilever.

2. THE GENERAL NONLINEAR PROBLEM

Most nonlinear problems are governed by the general differential equation given in (5.1).

\[ m \frac{d^2x}{dt^2} + \varphi \left( \frac{dx}{dt}, t \right) + f(x, t) = F(t) \]  

(5.1)

In analogy to linear systems, the terms from left to right are referred to as the inertial force, the damping or dissipative force, the spring or restoring force, and the external force or excitation respectively. Systems in which \( \varphi(\dot{x}) = 0 \) are referred to as undamped and those with \( F(t) = 0 \) as unforced or free [39]. If the time dependence is only in the inhomogeneous term on the right then the system is said to be externally excited, whereas, if the time dependence is in the coefficients on the left, then it is parametrically excited [40].
It is often the case that the nonlinearities themselves are small with respect to the linear terms. In these cases, the equation of motion may be reformulated with a small dimensionless parameter, $\varepsilon \ll 1$, which characterizes the strength of the nonlinearities (5.2). Equations of this form lend themselves to perturbation solutions. Systems in which the nonlinearities are larger usually must be solved using numerical or other means.

$$\ddot{x} + x = \varepsilon f(x, \dot{x}, t)$$

(5.2)

If the forces are conservative and also independent of time (autonomous) then the system is Hamiltonian. Certain analytical methods are only suitable to Hamiltonian systems.

3. THE DUFFING OSCILLATOR

The Duffing oscillator is an example of a single degree of freedom forced nonlinear oscillator in which the nonlinearity enters as a small cubic term in the restoring force. The “Duffing Equation” (5.3), was originally derived in 1918 by G. Duffing as a model for the forced vibration of industrial machinery [41].

$$\ddot{x} + \frac{1}{Q} \dot{x} + x + \varepsilon \alpha x^3 = \varepsilon f \cos(\Omega \tau)$$

(5.3)

The Duffing equation provided in (5.3) has been non-dimensionalized, where $x$ is the dependent variable, $Q$ is the quality factor, $\varepsilon$ is a small parameter representing the strength of the nonlinearity, $\alpha$ is a constant indicating whether the nonlinearity has a hardening or softening effect, and $f$ is the magnitude of the driving force which is assumed small and thus set to be of order $O(\varepsilon)$. $\Omega$ is the ratio of the driving frequency to
the natural frequency, and the dimensionless time, \( \tau \), has been scaled so that the natural frequency does not appear explicitly.

It should be noted that unlike most nonlinear problems, the undamped \((Q = \infty)\) and unforced \((f = 0)\) Duffing oscillator possesses an analytical closed form solution in terms of Jacobi’s elliptic functions that is associated with solitons [42]; however, only approximate solutions will be examined in the following sections of this chapter.

4. HARDENING AND SOFTENING SPRINGS

Consider the nonlinear restoring force typical to a Duffing oscillator (5.4).

\[
F(x) = x + \varepsilon \alpha x^3
\]

(5.4)

Where \( \varepsilon \) is again a small dimensionless parameter indicating the nonlinearity, and \( \alpha \) is a dimensionless constant whose magnitude may be positive, zero, or negative. The restoring force as a function of distance is plotted in Figure 5.1 for each of these three cases.

The stiffness, which is the derivative of force with respect to displacement, \( dF/dx \), is a constant in the linear case; however, for \( \alpha \neq 0 \), this is not the case. When \( \alpha \) is positive, the stiffness increases as \( |x| \) increases. This phenomenon is known as spring hardening. Conversely, when \( \alpha \) is negative, the stiffness decreases with increasing \( |x| \). This is referred to as spring softening.
Figure 5.1: The nonlinear restoring force for a Duffing oscillator versus displacement for three cases: $\alpha > 0$ (Hardening), $\alpha = 0$ (Linear), $\alpha < 0$ (Softening). The parameters used were $\varepsilon = 0.01$, $\alpha = 1, 0, -1$.

Establishing whether a system exhibits spring hardening or softening is an important step in determining the qualitative behavior of the response. For instance, systems that harden are usually stable, whereas softening systems have the potential for instability. Also, as we will show, the sign of $\alpha$ will have a significant and obvious effect on the frequency response curve of the forced Duffing oscillator.

Another important feature evident in Figure 5.1 is that the forces are all odd functions about $x = 0$. This indicates that the nonlinearities are symmetric in nature, i.e. the potential is even about $x = 0$. For instance, if the nonlinearities are due to mid-plane stretching of a fixed-fixed beam, then it makes no difference whether the beam is pushed or pulled, the same nonlinearities will arise in either direction. Any symmetric
nonlinearities will generate only odd powers of $x$ in the restoring force, e.g. $x^1$ and $x^3$ in (5.4). Obviously this must be case for the whole function to be odd.

Though only symmetric nonlinearities will be considered in this chapter, it is important to remember that this is not true in general. Specifically, for the electrostatically actuated microcantilever the even order terms are at least as significant as the symmetric terms, as was shown in the previous chapter.

5. NUMERICAL INTEGRATION

Numerical integration is an especially powerful and versatile approach to determining the solutions to arbitrary nonlinear differential equations. As such, there are many specific algorithms available and a great deal of literature on the subject. Only methods of the finite-difference type applicable to initial value problems will be considered here. For boundary value problems other numerical methods must be used, for instance a *shooting method* was used by Nayfeh to determine the modes of an electrostatically actuated micro-beam [43].

Numerical finite difference methods are so called because they utilize finite difference equations to approximate the derivatives. They can broadly be classified into explicit and implicit formulations. Explicit methods calculate the state of the system at a later time based only on the current state of the system. Implicit methods determine the state by solving an equation which depends on both the current and later states. Explicit solutions are generally more straightforward to program and less computationally intensive;
however, certain so called “stiff” problems require an impractically small time step in order to converge to an accurate solution using explicit methods. In these cases, an implicit method is more stable, and consequently, the most commonly used numerical integration methods are implicit.

5.1 Forward Euler Method

The simplest of the explicit finite-difference methods is the forward Euler method. Consider the first order differential equation and initial condition given by

\[ \dot{x}(t) = f(x,t) \quad x(t_0) = x_0 \]  

(5.5)

In the forward Euler method, the derivative at time \( t \) is approximated by the finite difference equation

\[ \dot{x}(t) \approx \frac{x(t + \Delta t) - x(t)}{\Delta t} \]  

(5.6)

Using (5.5), an incremental solution for \( x(t + \Delta t) \) is obtained

\[ x(t + \Delta t) \approx x(t) + \Delta t \cdot f(x(t), t) \]  

(5.7)

Therefore, by repeatedly applying (5.7), it is possible to arrive at an iterative solution to the original differential equation (5.5). Of course the convergence and accuracy of this solution depends on the nature of the problem and upon the choice of the time step, \( \Delta t \).
5.2 Backward Euler Method

An example of an implicit solution method is the backward Euler method. In this method, the following approximation is used.

$$\dot{x}(t) \approx \frac{x(t) - x(t - \Delta t)}{\Delta t}$$  \hspace{1cm} (5.8)

Again solving for \(x(t + \Delta t)\)

$$x(t + \Delta t) \approx x(t) + \Delta t \cdot f(x(t + \Delta t), t)$$  \hspace{1cm} (5.9)

The difference here is that the function \(f\) must now be evaluated at \(x(t + \Delta t)\), which is still an unknown at this point in the iteration. Equation (5.9), which may be quite complicated, must therefore be solved implicitly in order to determine \(x(t + \Delta t)\). However, since (5.9) is an algebraic equation, its solutions may be found using any of the standard root finding algorithms, \textit{e.g.} the Newton-Raphson method. Since this step must be performed for each iteration, the implicit methods are in general more computationally intensive.

Both the forward and backward Euler methods are examples of \textit{first-order} numerical methods. They are in general not as accurate as \textit{higher-order} methods. An example of a second-order method is the central difference method. The most commonly used higher-order method by far is the fourth-order Runge-Kutta method, or simply “RK4”.

92
5.3 Fourth-Order Runge-Kutta Method

For simplicity the explicit form of the Runge-Kutta method is given in (5.10); however, an implicit formulation is more frequently used.

\[ x(t + \Delta t) \approx x(t) + \frac{\Delta t}{6} \left( k_1 + 2k_2 + 2k_3 + k_4 \right) \]

\[
\begin{align*}
    k_1 &= f \left( x(t), t \right) \\
    k_2 &= f \left( x(t) + \frac{\Delta t}{2} \cdot k_1, t + \frac{\Delta t}{2} \right) \\
    k_3 &= f \left( x(t) + \frac{\Delta t}{2} \cdot k_2, t + \frac{\Delta t}{2} \right) \\
    k_4 &= f \left( x(t) + \Delta t \cdot k_3, t + \Delta t \right)
\end{align*}
\] (5.10)

Since the function value is evaluated at four intervals and averaged (increased weighting for the two central slopes), the incremental equation is more accurate than the single time step used in both the Euler methods. All of the numerical integrations presented in this thesis utilized an implicit form of the Runge-Kutta method integrated in the MAPLE™ dsolve command [44].

6. NUMERICAL SOLUTIONS TO THE DUFFING OSCILLATOR

In this section, the behavior of a Duffing oscillator is investigated numerically. Despite the relative simplicity of the equation of motion, the Duffing oscillator exhibits a remarkable variety of interesting behavior unique to nonlinear systems.
It will be shown that when the nonlinearities are small, the Duffing oscillator behaves in a qualitatively similar way to a simple harmonic oscillator. As the nonlinearity is increased, the frequency-response curve bends depending on whether the nonlinearity is hardening or softening. Eventually, the system bifurcates and multiple solutions (stable and unstable) emerge. As nonlinearities are increased even further, the system transitions into chaos.

6.1 Free Damped Duffing Oscillator

Consider the free vibrations of a damped Duffing oscillator whose governing equation of motion is given by (5.11).

\[
\ddot{x} + \frac{1}{Q} \dot{x} - x + \varepsilon \alpha x^3 = 0
\]  
(5.11)

The restoring force consists of a repulsive linear term and an attractive \((\alpha > 0)\) cubic nonlinearity. The symmetric potential leading to this nonlinear restoring force is given in (5.12)

\[
V(x) = -\frac{1}{2} x^2 + \frac{\varepsilon \alpha}{4} x^4
\]  
(5.12)

This potential is plotted in Figure 5.2 where the parameters \(\varepsilon = 0.1, \alpha = 1\) have been chosen.

It is clear from Figure 5.3 that there are two points of stable equilibria around \(x= -3, 3\) and an unstable equilibrium at \(x=0\). For any initial conditions except \(x_0 = 0, v_0 = 0\), the
system will eventually settle in one of the two potential wells. To see this, we perform a numerical integration for two initial positions, \( x_0 = 4.6 \) and \( x_0 = -10 \).

The time response of the displacement, \( x(t) \), and the corresponding phase space portraits (\( v(t) \) vs. \( x(t) \)) of the Duffing oscillator are presented in Figure 5.4. Depending on the initial conditions two distinct equilibrium points (centers) are reached corresponding to the minima of the potential. Since linear systems possess only one minimum, they only exhibit a single position of stable equilibrium. This illustrates that the overall response of nonlinear systems, in contrast with linear systems, may depend sensitively on initial conditions. This is colloquially known as “The Butterfly Effect”.

![Figure 5.2: The symmetric potential of a Duffing oscillator (5.12) with \( \varepsilon = 0.1, \alpha = 1 \).](image)
Figure 5.3: The time response and phase space portraits of the Duffing oscillator (5.11) for two initial conditions, with $\varepsilon = 0.1$, $\alpha = 1$, $Q=10$.

Two different stable equilibria are reached.

6.2 Forced Duffing Oscillator with Small Nonlinearities

We now return attention to the forced damped Duffing oscillator in which the linear restoring force is again attractive (negative valued). In the unforced case, the only
possible steady state solutions were the two stable centers; however, since the external force can add and remove energy from the system, there are more possibilities for the response.

For small nonlinearities ($\varepsilon \ll 1$), the response of the Duffing oscillator approaches that of the linear oscillator. This is the fundamental assumption in the various perturbation methods. We can see that this is indeed the case by numerically determining the response of the Duffing oscillator when the nonlinearity is small, $\varepsilon = 0.01$.

The numerically calculated system response of the Duffing oscillator with weak nonlinearity and forcing is shown in Figure 5.4. It exhibits behavior exactly analogous to a linear oscillator. The time response consists of both a transient part due to the initial conditions and a steady state solution due to the forcing term. The phase space portrait illustrates that the response of the Duffing oscillator approaches a stable periodic limit cycle. Also, the frequency response curve shows a single symmetric resonance peak.

The periodicity of the solution can be evaluated using a Poincaré map, in which the state of the system in phase space is sampled once per period, that is at times $T_n = 2n\pi/\Omega$. If the solution is truly periodic, then all points will overlap on the Poincaré map.
Figure 5.4: The numerically determined potential, time response, phase portrait and frequency response curve of a Duffing oscillator with weak nonlinearity and forcing, both $O(\varepsilon)$ with $\varepsilon = 0.01$. The behavior closely resembles that of a linear oscillator. The other parameters are $\alpha = 1$, $x_0 = v_0 = 0$. 
The Poincaré map for the weakly nonlinear forced Duffing oscillator is presented in Figure 5.5. In the Poincaré map, the transient motion appears as points which approach a steady state value near \( x(T_n) = 1 \) and \( x(T_n) = 0 \). After the transients have decayed, all subsequent points overlap. This indicates that the steady state motion of the Duffing oscillator with weak parameters is strictly periodic. Of course, depending on the phase of the “snapshots” the actual location of the steady state solution in the Poincaré map can be shifted. However, if the steady state motion is periodic, the points will still coalesce regardless of phase.

For small nonlinearities, the steady state response consists of a single harmonic signal at the driving frequency, \( \Omega \). This is evident in the time response of Figure 5.4. It can also be shown quantitatively by numerically evaluating the Fourier coefficients of the time response data. For the time response shown, the first harmonic amplitude is computed to be, \( A_1 = 0.96 \). The second harmonic amplitude is much less, \( A_2 = 0.0044 \).

6.3 Frequency-Response Curve Bending in Duffing Oscillator

As the nonlinearities are increased somewhat, an interesting phenomenon emerges in the frequency-response curve of the forced Duffing oscillator. The resonance curve begins to bend, to the left for softening springs \( (\alpha < 0) \) and to the right for hardening springs \( (\alpha > 0) \). This behavior has been demonstrated numerically as shown in Figure 5.6.
Figure 5.5: Poincaré map of the weakly nonlinear Duffing oscillator.

After the transient motion has decayed, all points overlap at the stable steady state of the system.
Figure 5.6: Numerically determined frequency response curves for the forced damped Duffing oscillator. The resonance peaks bend depending on whether the nonlinearities are softening or hardening. Simulation parameters are $\varepsilon = 0.1$, $Q = 10$, $\sigma_f = 0.1$.

The frequency response curves of Figure 5.6 illustrate a few important characteristics common to many nonlinear systems. In linear oscillators, the frequency of vibration is independent of the amplitude of excitation. This is not the case in nonlinear systems.

Also, notice that there are “jumps” in the nonlinear frequency-response curves. These breaks indicate that there may be multiple stable and unstable solutions. It can be shown that the stable solution which is followed depends on the past history of the system. Nonlinear oscillators, therefore, exhibit hysteresis. These jumps, also called bifurcations, represent an extremely important and unique feature of nonlinear systems.
7. PERIOD DOUBLING, BIFURCATIONS, AND THE ONSET OF CHAOS

Bifurcations are sudden qualitative or topological changes in the system behavior that arise upon increasing the nonlinearities in the system. It has been shown that they are a route to chaos in the Duffing oscillator [45].

Chaos can broadly be defined as the extreme sensitivity of certain dynamical systems to initial conditions. Chaotic systems appear to be random, despite the deterministic nature of their governing equations. Chaotic oscillators are referred to as strange attractors. They are attractors because points that are close to the attractor remain close even if slightly perturbed. However, unlike limit cycles which are asymptotically periodic, strange attractors possess a fractal structure, in which the motion never repeats itself exactly. That is to say, the response is not periodic (aperiodic). Since they are not periodic, the Poincaré map of a strange attractor does not approach a single point but rather form a cluster of points around the attractor.

The onset of chaos in a Duffing oscillator is easily shown numerically. If the nonlinearity is relatively large, $\varepsilon = 0.4$, the phase space portrait and Poincaré map are given in Figure 5.7. As nonlinearity increases and chaos emerges, more harmonics appear in the response which eventually leads to a completely aperiodic solution. In fact, at certain critical values of the nonlinearity, the response suddenly does not repeat itself every period of the driving force, but rather every two periods, then four periods, and so on. This period-doubling bifurcation has been shown to be an important route to chaos in both the Duffing Oscillator and the electrostatically actuated micro-cantilever [46].
**Figure 5.7:** Numerically determined phase portrait and Poincaré map illustrating the onset of chaos in a forced damped Duffing oscillator with 
\[ \varepsilon = 0.4, \; Q = 10, \; \epsilon_f = 0.4. \]

8. PERTURBATION METHODS

Underlying the perturbation methods is the fundamental assumption that for small nonlinearities the solution should be close to that of a linear system. The solutions may then be expressed as a power series in terms of a small parameter characterizing the nonlinearities. Since the power series is in practice finite, the perturbation solutions are only approximate; however the error can be controlled by choosing the order of the perturbation. Perturbation methods were initially developed for problem in celestial mechanics; however, they now find extensive use in various branches of applied mathematics especially nonlinear dynamical systems.
8.1 Straight-Forward Expansion

The first perturbation method considered is the so called straight-forward expansion. It will be shown that this method does not result in a uniformly valid solution. Consider the undamped unforced Duffing equation (5.13).

\[ \ddot{x} + x + \varepsilon x^3 = 0 \]  \hspace{1cm} (5.13)

We assume a solution may be represented as a power series in the variable \( \varepsilon \), (5.14), where \( x_0(t) \) will be the solution to the linear free vibration problem.

\[ x(t) = x_0(t) + \varepsilon x_1(t) + \varepsilon^2 x_2(t) + ... \]  \hspace{1cm} (5.14)

We assume the following initial conditions (5.15).

\[ x(0) = A \quad \dot{x}(0) = 0 \]  \hspace{1cm} (5.15)

Substituting this expansion into (5.13) and only keeping terms up to \( O(\varepsilon^2) \) yields (5.16).

\[ (\ddot{x}_0 + x_0) + \varepsilon(\ddot{x}_1 + x_1 + x_0^3) + \varepsilon^2(\ddot{x}_2 + x_2 + 3x_0^2x_1) = 0 \]  \hspace{1cm} (5.16)

Now since the power series is \textit{linearly independent} the coefficients of each power must separately equate to zero. The zeroth order equation is simply that of a linear free undamped oscillator (5.17).

\[ \ddot{x}_0 + x = 0 \]  \hspace{1cm} (5.17)

The solutions to this are well known and may either be expressed in terms of sin and cos, or exponential functions, \textit{c.f.} Chapter III. In order to satisfy the initial conditions we take the solution as (5.18).
\[ x_0(t) = A \cos(t) \]  \hspace{1cm} (5.18)

Now we may substitute this into the first order equation

\[ \ddot{x} + x = -x_0^3 \]  \hspace{1cm} (5.19)

This can be rewritten using a trigonometric identity.

\[ \ddot{x} + x = -\frac{3}{4} A^3 \cos(t) - \frac{1}{4} A^3 \cos(3t) \]  \hspace{1cm} (5.20)

The first term on the right represents a driving force at the natural frequency of this system. Since there is no damping, this force leads to a solution whose amplitude grows without bound. The exact form of this solution may be determined using Green’s functions as in the previous chapter, resulting in the first order correction to the response (5.21).

\[ x_1(t) = -\frac{3}{8} A^7 \sin(t) - \frac{1}{32} A^3 \left( \cos(t) - \cos(3t) \right) \]  \hspace{1cm} (5.21)

The first term here is known as a \textit{secular term} since it increases linearly with \( t \) without bound. If the secular term were retained, it would destroy the periodicity of the whole solution. Therefore an important step in all perturbation procedures is to set the coefficients in the secular terms to zero. However, in this case there is no such freedom since \( A \) is a real constant.

As it turns out, the emergence of the secular term in (5.21) is due to the fact that the nonlinear oscillator will not in general oscillate at the same frequency as the applied
force. The frequency in nonlinear systems is actually amplitude dependent. This is one of the most important distinctions between linear and nonlinear oscillators.

Secular terms can be seen to have the effect of shifting the frequency of vibration in the following expansion (5.22) [38].

\[
\sin[(\omega + \varepsilon)t] \approx \sin(\omega t) + \varepsilon t \cos(\omega t) - \frac{\varepsilon^2 t^2}{2} \sin(\omega t) + ... \tag{5.22}
\]

### 8.2 Lindstedt-Poincaré Method

The problem of secular terms can be avoided by assuming that the frequency of vibration can also be expanded in terms of terms of the small parameter, \( \varepsilon \), essentially making it amplitude dependent, (5.23).

\[
\omega = \omega_0 + \varepsilon \omega_1 + \varepsilon^2 \omega_2 + ...
\tag{5.23}
\]

We also introduce a scaled time variable, \( \tau \equiv \omega t \), since the system no longer vibrates with the applied frequency. The nonlinear equation of motion is then given by (5.24).

\[
\omega^2 \frac{d^2 x}{d\tau^2} + x + \varepsilon x^3 = 0 \tag{5.24}
\]

Substituting the expansions for both \( x(t) \) and \( \omega \), keeping only terms of first order in \( \varepsilon \), and equating coefficients as before gives the following set of equations (5.25). Here the dot indicates differentiation with respect to \( \tau \).
\[
\omega_0^2 \ddot{x}_0 + x_0 = 0
\]
\[
\omega_0^2 \ddot{x}_1 + x_1 = -2\omega_0 \omega_1 \dot{x}_0 - x_0^3
\]  
(5.25)

The zeroth order solution satisfying the initial conditions is given by
\[
x_0(t) = A \cos(\tau)
\]
(5.26)

The first order equation ten becomes
\[
\ddot{x}_1 + x_1 = \left(2\omega_1 - \frac{3}{4} A^2\right) A \cos(\tau) - \frac{1}{4} A^3 \cos(3\tau)
\]  
(5.27)

The first term on the right would again generate a secular term in the solution; however, in this case, we have the freedom to set the coefficient of this term to zero. Therefore, the first order correct to the natural frequency is given by (5.28).
\[
\omega_1 = \frac{3}{8} A^2
\]
(5.28)

And the solution for the displacement to first order in \(\varepsilon\) is (5.29).
\[
x(t) = \left(A - \frac{1}{32} \varepsilon A^3\right) \cos(\omega t) + \left(\frac{1}{32} \varepsilon A^3\right) \cos(3\omega t) + ...
\]  
(5.29)

The form of the solution in (5.29) is instructive. It shows that the response of nonlinear oscillators can be expanded in terms of a harmonic of Fourier series. This must be case if the solutions are truly periodic. This fact is used to develop an alternative approximate solution procedure known as the harmonic balance method that is used extensively in this study.
The nonlinear equations considered above were autonomous, that is they have no explicit time dependence. Problems involving forced excitations, included the electrostatically actuated microcantilever, do not fall into this class. They are said to be nonautonomous since the time enters explicitly through the forcing function. A general nonautonomous nonlinear equation takes the form of (5.30), where again $\tau = \omega t$ and $\phi$ is the phase of the applied force.

$$\ddot{x}(\tau) + x = \varepsilon f(x, \dot{x}, \tau + \phi)$$  \hspace{1cm} (5.30)

For nonautonomous systems, the steady state periodic solution will have a least period which is equal to that of the driving force, so an expansion of the natural frequency as in (5.23) is not necessary. However, the amplitude and phase difference of the response are unknown. Therefore, we proceed by expanding both in power series in terms of an auxiliary variable, $\mu$, which will later be set to one. (5.31)

$$x = x_0 + \mu x_1 + \mu^2 x_2 + ... \hspace{1cm} (5.31)$$

$$\delta = \delta_0 + \mu \delta_1 + \mu^2 \delta_2 + ...$$

Again we consider the forced undamped Duffing oscillator governed by (5.32).

$$\ddot{x} + \Omega^2 x = \mu \left[ \left( \Omega^2 - 1 \right) x - \varepsilon x^3 + F \cos(\Omega t) \right]$$ \hspace{1cm} (5.32)

Now scaling time by $\tau \equiv \Omega t - \delta$, we obtain (5.33), where the dots again denote differentiation with respect to $\tau$.

$$\ddot{x} + x = \mu \left[ \left( 1 - \frac{1}{\Omega^2} \right) x - \frac{\varepsilon}{\Omega^2} x^3 + \frac{F}{\Omega^2} \cos(\tau + \delta) \right]$$ \hspace{1cm} (5.33)
Substituting the expansions (5.31) and equating the coefficients of the powers in \( \mu \) yields the following set of differential equations to second order in \( \mu \).

\[
\begin{align*}
\ddot{x}_0 + x_0 &= 0 \\
\ddot{x}_1 + x_1 &= \left(1 - \frac{1}{\Omega^2} \right)x_0 - \frac{1}{\Omega^2} \left( \varepsilon x_0^3 - F \cos(\tau + \delta_0) \right) \\
\ddot{x}_2 + x_2 &= \left(1 - \frac{1}{\Omega^2} \right)x_1 - \frac{1}{\Omega^2} \left( 3\varepsilon x_0^2 x_1 - \delta_1 F \sin(\tau + \delta_0) \right)
\end{align*}
\] (5.34)

The solution to the linear equation is as usual given by (5.35), except \( A_0 \) is a constant to be determined in a subsequent step.

\[ x_0(\tau) = A_1 \cos(\tau) \] (5.35)

The first order equation then becomes (5.36) using some trigonometric identities.

\[
\begin{align*}
\ddot{x}_1 + x_1 &= \left[ \left(1 - \frac{1}{\Omega^2} \right)A_1 - \frac{1}{\Omega^2} \left( \varepsilon \frac{3}{4} A_1^3 \cos(\tau) - F \cos(\delta_0) \right) \right] \cos(\tau) \\
&\quad - \frac{F}{\Omega^2} \sin(\delta_0) \sin(\tau) - \varepsilon \frac{1}{4\Omega^2} A_1^3 \cos(3\tau)
\end{align*}
\] (5.36)

Again in order to eliminate secular terms that would destroy the periodicity of the solution, we must set the coefficient of \( \cos(\tau) \) and \( \sin(\tau) \) on the right to zero, leading to (5.37).

\[
\begin{align*}
\sin(\delta_0) &= 0 \quad \cos(\delta_0) = 1 \\
(\Omega^2 - 1)A_1 - \varepsilon \frac{3}{4} A_1^3 + F &= 0
\end{align*}
\] (5.37)
Equation (5.37) is the important *frequency-response* relation for a forced Duffing oscillator. It connects the amplitude of the response with the frequency of and amplitude (through $\varepsilon$) of the excitation.

The solution to the first order equation once the secular terms have been removed is given in

$$x_i(\tau) = A_2 \cos(\tau) + \frac{\varepsilon}{32\Omega^2} A_1^3 \cos(3\tau)$$

(5.38)

Where $A_2$ is a constant used to eliminate secular terms in the next order solution.

If we now solve for the magnitude of $A_1$ in relation (5.37) and plot it versus driving frequency, $\Omega$, we obtain the frequency-response curve for an undamped forced Duffing oscillator. In Figure 5.8 the resonance curve is plotted for a hardening systems with $\varepsilon = 0.1$.

Just as in the numerical simulations, the resonance peak of a hardening Duffing oscillator bends to the right. However, unlike the numeric solutions, the perturbation method clearly shows the three solutions in the bent region, two of which are stable (solid line) and one of which is unstable (dashed line). This also explains the hysteresis effect since the system will tend to stay on the stable solution it is on as long as possible before jumping. Therefore, the response should be different depending on whether the frequency is slow increased from below resonance or slowly decreased from above.
Figure 5.8: Perturbation based analytic solution to the frequency-response curve of a hardening undamped forced Duffing oscillator with $\varepsilon = 0.1$, $\epsilon f = 0.1$.

8.3 Method of Multiple Scales

The method of multiple scales is an alternative perturbation method strongly advocated by Nayfeh [40]. It is based on the assumption that solution should be a function of
several independent time scales. It has certain advantages over the Lindstedt-Poincaré method, such as more conveniently being able to handle nonconservative (damped) problems.

9. AVERAGING METHODS

Consider the general first order differential equation given by (5.39). The following arguments will hold equally well for higher order systems since they may readily be transformed into a system of first order equations.

\[ \dot{x} = \varepsilon f(x,t) \]  \hspace{1cm} (5.39)

The function \( f \) is typically periodic with period \( \tau \). The evolution of such systems occurs on two timescales, a fast oscillatory one due to the periodicity of \( f(x,t) \) and a slow one associated with the small parameter \( \varepsilon \).

The fundamental assumption of the averaging methods is that the coefficients of the fast oscillatory response do not change significantly over a given period. Therefore, it is possible to average over periods and transform the differential equations into algebraic equations that are more readily soluble.

Averaging methods are limited in their ability to handle transient phenomena that may vary more quickly in time. In these cases, other numerical or perturbation methods may be more suitable. However, for determining the steady state behavior of periodic systems, there is perhaps no better choice than averaging.
9.1 Method of Harmonic Balance

One form of averaging that is extensively used is known as the method of harmonic balance, also sometimes referred to as harmonic analysis. In this method the solutions are assumed to be of the form of a harmonic series (5.40).

\[ x(t) = \sum_{n=0}^{N} A_n(\Omega) \cos(n\Omega t) + B_n(\Omega) \sin(n\Omega t) = \sum_{n=0}^{N} \tilde{A}_n e^{in\Omega t} \]  

(5.40)

The two forms of the series are equivalent due to Euler’s identity. In the latter form, the amplitudes \( \tilde{A}_n \) are complex valued functions. Therefore phase information that is contained in the relative magnitude of \( A_n \) and \( B_n \) is contained solely in \( \tilde{A}_n \).

The perturbation solutions in the previous sections all generated harmonic series solutions. Therefore, it is reasonable to assume that (5.40) is a valid form for most systems with relatively weak nonlinearities.

For the undamped forced Duffing oscillator considered in the previous section, we will now seek solution using Harmonic balance. Since the potential is symmetric, only odd order terms will appear in the harmonic series for the response. Also, no sine terms appear because of time reversal symmetry. Therefore, the solution takes the form (5.41).

\[ x(t) = \sum_{n=1,3,5,...} A_n(\Omega) \cos(n\Omega t) \]  

(5.41)
Retaining terms only to third order and inserting this into the Duffing equation (5.32) yields (5.42)

\[
\left(1 - \Omega^2\right) A_i \cos(\Omega t) + (1 - 9\Omega^2) A_i \cos(3\Omega t) + \\
\varepsilon \left(\frac{3}{4} A_i^3 \cos(\Omega t) + \frac{1}{4} A_i^3 \cos(3\Omega t)\right) = f \cos(\Omega t)
\]

Since like power series expansions, Fourier series are linearly independent, the coefficients may separately equated to zero. In this case the coefficients \(\cos(\Omega t)\) gives

\[
\left(1 - \Omega^2\right) A_i + \varepsilon \frac{3}{4} A_i^3 = f
\]

Which we recognize as identical to the frequency-response equation derived using a perturbation method in (5.37). Therefore, to this level of approximation, harmonic balance and perturbation theory yield identical results and both are equally valid.
CHAPTER VI

ANALYTICAL AND NUMERICAL RESULTS

1. INTRODUCTION

This chapter presents the results, obtained using various analytical procedures, for the response of an electrostatically actuated micro-cantilever. First, low-order harmonic balance solutions are derived. These relatively simple closed-form solutions offer basic explanations for many of the interesting properties observed in the experimental HDR results, cf. Chapter II.

Higher-order harmonic balance solutions are then sought; however they become tedious if not impossible to perform by hand. Furthermore, the analytical expressions that are obtained are generally too lengthy to be of much pedagogical value. For these reasons, a computer program was created to carry out the harmonic balance analysis. The program was written in MAPLE™ and is included in the APPENDIX.

The higher-order harmonic responses for both the displacement and current were calculated using this program. They were shown to be in excellent agreement with the experimental frequency responses and polar plots obtained using HDR. Furthermore, the efficiency of the computerized solutions allowed the effect various parameters to be studied quantitatively. The variation of amplitude as a function of gap distance is presented.
Finally a preliminary analysis and some numerical results are presented for the case of large nonlinearities. These nonlinearities may be large as a result of reduced nominal gap distance, reduced damping, or increased applied voltage. These results are then compared to some interesting recent data generated using HDR exhibiting Duffing behavior.

2. CLOSED-FORM HARMONIC BALANCE SOLUTIONS

2.1 Linearized EOM

Consider the dimensionless linearized equation of motion for the first mode of an electrostatically actuated micro-cantilever, derived in Chapter IV.

\[ \ddot{q} + \frac{1}{Q} \dot{q} + q = \tilde{\epsilon} \left( 1 + \Lambda_1 \exp(i\tilde{\Omega}\tau) + \Lambda_2 \exp(i2\tilde{\Omega}\tau) \right) \left( \tilde{\alpha} + \tilde{\beta}q \right) \]  \(6.1\)

where \( q \equiv \zeta / d \) is the dimensionless displacement with \( \zeta \) the real tip deflection of the micro-cantilever and \( d_0 \) the nominal gap distance separating the cantilever and counter-electrode. Since \( \zeta < d_0 \) to prevent crashing, \( q \) must remain less than unity. The dot indicates differentiation with respect to the scaled time, \( \tau \equiv \omega_0 t \), and the small parameter \( \tilde{\epsilon} \) governing the nonlinearity of the system is given by

\[ \tilde{\epsilon} \equiv \frac{C_1 V^2}{2m\omega_0^2 d_0} \]  \(6.2\)
where the $C_n$ are the Taylor series coefficients of the capacitance, $m$ and $\omega_0$ are the first modal mass and natural frequency, and $V^2 \equiv V_{dc}^2 + V_{ac}^2/2$. The dimensionless forcing coefficients are

$$\tilde{\alpha} \equiv 1$$
$$\tilde{\beta} \equiv \frac{C_2 d_0}{C_1}$$

Finally, the dimensionless system constants $\tilde{\alpha}$ and $\tilde{\beta}$ for the lumped model with a general capacitance are given by

$$\Lambda_1 = \frac{2V_{dc}V_{ac}}{V_{dc}^2 + V_{ac}^2/2}$$
$$\Lambda_2 = \frac{V_{ac}^2}{2V_{dc}^2 + V_{ac}^2}$$

The linearized EOM in the unscaled system parameters is thus given by

$$\ddot{\zeta}(t) + \frac{\alpha_0}{Q} \dot{\zeta}(t) + \omega_0^2 \zeta(t) =$$

$$\frac{1}{2m}(C_1 + C_2 \zeta(t)) \left[V_{dc}^2 + \frac{1}{2}V_{ac}^2 + 2V_{dc}V_{ac} \cos(\Omega t) + \frac{1}{2}V_{ac}^2 \cos(2\Omega t) \right]$$

Using the method of harmonic balance, it is possible to obtain analytical solutions to (6.1) and (6.5) corresponding to the mechanical response of the first mode of an electrostatically actuated micro-cantilever. From the mechanical response, it is then straightforward to calculate the electrical response. In this section we will consider only the first two harmonics, but the procedure for generating any higher-order solutions would follow naturally, although with increasing mathematical complexity.
2.2 Validity of Linearization

The EOMs in (6.1) and (6.5) have been *linearized*. In other words, all terms involving powers of \( q \) higher than one have been dropped. This is equivalent to neglecting all capacitive coefficients higher than \( C_2 \). This first-order approximation of the electrostatic force is sufficient for small \( q \), in which case the error is \( O(q^2) \).

However, caution must be used when making such approximations. In nonlinear systems, higher-order terms can affect lower-order harmonics. This can readily be seen in the following trigonometric identities.

\[
\cos^2(\tau) = \frac{1}{2}(1 + \cos(2\tau)) \quad \cos^3(\tau) = \frac{3}{4}\cos(\tau) + \frac{1}{4}\cos(3\tau)
\] (6.6)

Therefore, \( q^2 \) terms can shift the average displacement as well as contributing to the second harmonic. Likewise, \( q^3 \) can affect both the first and third harmonics. Similar relations exist for all high-powers of \( \cos(\tau) \). Therefore it is difficult to determine *a priori* how nonlinearities may affect the harmonic response of the system, and generally many terms must retained in the solutions in order to obtain accurate results.

This is the reason that Nayfeh strongly objects to the harmonic balance technique [40]. Indeed the perturbation methods are typically more valid since they naturally introduce higher harmonics as necessary. However, computerized implementations of the harmonic balance method are more suited to obtaining accurate results for higher harmonics because a large number of terms can easily be retained in the solution.
2.3 Harmonic Series

The fundamental assumption in the harmonic balance method is that the solution may be expanded as a harmonic (Fourier) series (6.7). This implies that the motion is periodic and steady state. Consequently, harmonic balance and other averaging methods are not suitable for determining transient responses.

\[ q(\tau) = A_0 + \sum_{n=1}^{N} A_n \cos\left(n\tilde{\Omega} \tau + \phi_n\right) \]  

(6.7)

where \( n \) is the harmonic number, \( N \) is the highest harmonic considered, \( A_n \) is the real valued harmonic amplitude, \( \phi_n \) is the harmonic phase difference, \( \tilde{\Omega} = \Omega / \omega_0 \) is the dimensionless driving frequency, and \( \tau \) is the scaled time. It is usually much more convenient to consider the complex exponential form of the harmonic series solution.

\[ \tilde{q}_c(\tau) = A_0 + \sum_{n=1}^{N} \tilde{A}_n \exp\left(jn\tilde{\Omega} \tau\right) \]  

(6.8)

where \( \tilde{q}_c \) and \( \tilde{A}_n \) are the complex scaled displacement and complex harmonic amplitude respectively, and \( j \) is the imaginary unit. Since the actual displacement must be real valued, we must consider only the real part of \( \tilde{q}_c \). However, it can readily be shown, that the real amplitude and phase of the solution can be determined directly from the modulus and argument of the complex amplitude, noting that \( \tilde{A}_n = A_n \exp\left(j\phi_n\right) \).

\[ \text{Re}[\tilde{q}_{c,n}] = \text{Re}\left[A_n \exp\left(j\phi_n\right)\exp\left(jn\tilde{\Omega} \tau\right)\right] = A_n \cos\left(n\tilde{\Omega} \tau + \phi_n\right) \]  

(6.9)
In all the subsequent harmonic balance analysis, the complex exponential solution will be assumed. The exponential solution drastically simplifies the harmonic balance analysis, because the trigonometric identities, \( e.g. \) (6.6), do not explicitly need to be considered. Rather, they follow naturally from the real and imaginary parts of the solution as follows

\[
\exp(j\tau) = \exp(j2\tau) = \cos(2\tau) + j\sin(2\tau) = \\
\exp(j\tau) \cdot \exp(j\tau) = [\cos(\tau) + j\sin(\tau)] \cdot [\cos(\tau) + j\sin(\tau)] = \\
\cos^2(\tau) - \sin^2(\tau) + j2\cos(\tau)\sin(\tau)
\]

Equating the real and imaginary parts separately and using the identity \( \cos^2(\tau) + \sin^2(\tau) = 1 \) yields

\[
\cos(2\tau) = \cos^2(\tau) - \sin^2(\tau) = 2\cos^2(\tau) - 1
\]

\[
\sin(2\tau) = 2\cos(\tau)\sin(\tau)
\]

The first equation yields the identity previously stated in (6.6), and the second is another of the so-called double-angle formulas.

Assuming a second-order harmonic series solution \( (N=2) \), and substituting it into the EOM (6.1) with complex forcing gives

\[
-\tilde{\Omega}^2 \tilde{A}_1 \exp(j\tilde{\Omega}\tau) - 4\tilde{\Omega}^2 \tilde{A}_2 \exp(j2\tilde{\Omega}\tau) + j\frac{\tilde{\Omega}\tilde{A}_1}{Q} \exp(j\tilde{\Omega}\tau) + j\frac{2\tilde{\Omega}\tilde{A}_1}{Q} \exp(j2\tilde{\Omega}\tau) + \\
A_0 + \tilde{A}_1 \exp(j\tilde{\Omega}\tau) + \tilde{A}_2 \exp(j2\tilde{\Omega}\tau) = \tilde{\varepsilon} \left( 1 + \Lambda_1 \exp(j\tilde{\Omega}\tau) + \Lambda_2 \exp(j2\tilde{\Omega}\tau) \right) \cdot \\
\left( \tilde{\alpha} + \tilde{\beta} \left( A_0 + \tilde{A}_1 \exp(j\tilde{\Omega}\tau) + \tilde{A}_2 \exp(j2\tilde{\Omega}\tau) \right) \right)
\]
Expanding the driving force on the right-hand-side, ignoring any harmonics higher than
the second, and factoring gives

\[
\left( -\dot{\Omega}^2 + j \frac{\dot{\Omega}}{Q} + 1 \right) \hat{A}_i \exp\left( j\dot{\Omega} \tau \right) + \left( -4\dot{\Omega}^2 + j \frac{2\dot{\Omega}}{Q} + 1 \right) \hat{A}_2 \exp\left( j2\dot{\Omega} \tau \right) \\
(\tilde{\epsilon} \alpha + \tilde{\epsilon} \beta A_0) + (\tilde{\epsilon} \alpha \Lambda_1 + \tilde{\epsilon} \beta \Lambda_0 \Lambda_1 + \tilde{\epsilon} \beta \Lambda_1) \exp\left( j\dot{\Omega} \tau \right) + \\
(\tilde{\epsilon} \alpha \Lambda_2 + \tilde{\epsilon} \beta A_0 \Lambda_2 + \tilde{\epsilon} \beta \Lambda_0 \Lambda_1 + \tilde{\epsilon} \beta \Lambda_2) \exp\left( j2\dot{\Omega} \tau \right)
\]

(6.13)

Since the harmonic terms and the constant are linearly independent (this is the essentially
the same as saying that the equations must hold for all time), we may equate the
coefficients of each harmonics and the constant term to zero separately. This leads to an
expression for the constant term and for each of the \( N \) complex harmonic amplitudes in
terms of the driving frequency and system parameters.

2.4 Average Position

Equating only the constant terms in (6.13) gives

\[
A_0 = (\tilde{\epsilon} \alpha + \tilde{\epsilon} \beta A_0)
\]

(6.14)

We may then expand the fraction since \( \tilde{\epsilon} \) is small.

\[
A_0 = \frac{\tilde{\epsilon} \alpha}{\left(1 - \tilde{\epsilon} \beta \right)} \approx \tilde{\epsilon} \alpha \left(1 + \tilde{\epsilon} \beta + O(\tilde{\epsilon}^2)\right)
\]

(6.15)

Or in terms of the system parameters, shifting back to the unscaled displacement,

\[
\zeta =qd_0.
\]
This represents a shift in the average position of the cantilever toward the counter-electrode.

2.4 First Mechanical Harmonic

Equating the coefficients of the first harmonic terms \( \exp(j\tilde{\Omega} \tau) \) gives

\[
\begin{pmatrix}
-\tilde{\Omega}^2 + j\frac{\tilde{\Omega}}{Q} + 1
\end{pmatrix}\tilde{A}_1 = \left(\tilde{\varepsilon}\tilde{\alpha}_1 + \tilde{\varepsilon}\tilde{\beta}_1 A_0 + A_0 \right)
\]

Solving for \( \tilde{A}_1 \).

\[
\tilde{A}_1 = \frac{\tilde{\varepsilon}\tilde{\Lambda}_1 (\tilde{\alpha} + \tilde{\beta} A_0)}{-\tilde{\Omega}^2 + j\frac{\tilde{\Omega}}{Q} + 1 - \tilde{\varepsilon}\tilde{\beta}}
\]

We now explicitly solve for the real and imaginary parts

\[
\tilde{A}_1 = \frac{\tilde{\varepsilon}\tilde{\Lambda}_1 (\tilde{\alpha} + \tilde{\beta} A_0)}{((-\tilde{\Omega}^2 + 1 - \tilde{\varepsilon}\tilde{\beta}) + j\tilde{\Omega}/Q)} = \frac{\tilde{\varepsilon}\tilde{\Lambda}_1 (\tilde{\alpha} + \tilde{\beta} A_0) ((-\tilde{\Omega}^2 + 1 - \tilde{\varepsilon}\tilde{\beta}) - j\tilde{\Omega}/Q)}{(-\tilde{\Omega}^2 + 1 - \tilde{\varepsilon}\tilde{\beta})^2 - (\tilde{\Omega}/Q)^2} - \frac{\tilde{\varepsilon}\tilde{\Lambda}_1 (\tilde{\alpha} + \tilde{\beta} A_0) \tilde{\Omega}/Q}{(-\tilde{\Omega}^2 + 1 - \tilde{\varepsilon}\tilde{\beta})^2 - (\tilde{\Omega}/Q)^2}
\]

The real amplitude of the first harmonic is therefore
\[ A_i = \sqrt{\text{Re}\left[\tilde{A}_i\right]^2 + \text{Im}\left[\tilde{A}_i\right]^2} = \frac{\tilde{\varepsilon}\Lambda_i (\tilde{\alpha} + \tilde{\beta} A_0)}{\left[(-\tilde{\Omega}^2 + 1 - \tilde{\varepsilon} \tilde{\beta})^2 - (\tilde{\Omega}/Q)^2\right]^{1/2}} \] (6.20)

The phase difference is given by

\[ \phi_i = \tan^{-1}\left(\frac{\text{Im}\left[\tilde{A}_i\right]}{\text{Re}\left[\tilde{A}_i\right]}\right) = \tan^{-1}\left(\frac{-\tilde{\Omega}/Q}{(-\tilde{\Omega}^2 + 1 - \tilde{\varepsilon} \tilde{\beta})}\right) \] (6.21)

These are simply the amplitude and phase responses of a damped driven linear harmonic oscillator, as they must be since we linearized the EOM (6.1). Figure 6.1 shows the real mechanical first harmonic amplitude (6.20) plotted as function of driving frequency for two values of nonlinearity, \( \tilde{\varepsilon} = 0.01 \) and \( \tilde{\varepsilon} = 0.1 \).

A single symmetric resonance peak exists in the first harmonic; however, it is shifted below the unforced natural frequency, 1. This is due to the \( \tilde{\varepsilon} \tilde{\beta} \) term in (6.20). The shift is more pronounced when the nonlinearity is larger. Therefore, the electrostatic actuation force leads to a spring softening effect, which can be seen experimentally in Figure 2.2.

In terms of the unscaled system parameters the amplitude and phase of the first mode are given by (6.22) to first-order in \( \tilde{\varepsilon} \), i.e. ignoring the \( \tilde{\varepsilon} \tilde{\beta} \) term.

\[ A_i = \frac{C V_{dc} V_{ac}}{m \sqrt{(\omega_0^2 - \Omega^2)^2 + \left(\frac{\omega_0 \Omega}{Q}\right)^2}} \] (6.22)

\[ \phi_i = \tan^{-1}\left(\frac{\omega_0 \Omega}{Q (\Omega^2 - \omega_0^2)}\right) \]
Figure 6.1: Mechanical first harmonics from closed form harmonic balance calculations for two levels of nonlinearity $\tilde{\epsilon} = 0.01$ and $\tilde{\epsilon} = 0.1$.

2.5 Second Mechanical Harmonic

For the second harmonic, we follow the same procedure as before, equating coefficients of $\exp(j2\tilde{\Omega} \tau)$ in (6.13).

$$\left(-4\tilde{\Omega}^2 + \frac{2\tilde{\Omega}}{Q} + 1 - \tilde{\epsilon}\tilde{\beta}\right)\tilde{A}_2 = \left(\tilde{\epsilon}\tilde{\alpha}_2 + \tilde{\epsilon}\tilde{\beta}_2 A_0 A_2 + \tilde{\epsilon}\tilde{\beta}_2 \tilde{A}_2\tilde{A}_1\right)$$  \hspace{1cm} (6.23)

In this case, the expression for $\tilde{A}_2$ includes $\tilde{A}_1$ which is itself complex and depends on the driving frequency. Therefore the expressions for the real frequency and phase of the second harmonic are necessarily more complicated, and will not be given explicitly. However, the complex amplitude is given by (6.24).
\[
\tilde{A}_2 = \frac{\tilde{\epsilon}\tilde{a}\Lambda_2 + \tilde{\epsilon}\tilde{\beta} A_0\Lambda_2 + \tilde{\epsilon}\tilde{\beta}\tilde{A}_1\Lambda_1}{-4\tilde{\Omega}^2 + j\frac{2\tilde{\Omega}}{Q} + 1 - \tilde{\epsilon}\tilde{\beta}}
\]  
(6.24)

**Figure 6.2:** Mechanical second harmonics from closed form harmonic balance calculations for two levels of nonlinearity $\tilde{\epsilon} = 0.01$ and $\tilde{\epsilon} = 0.1$

The mechanical second harmonic amplitude (6.24) for two values of nonlinearity, $\tilde{\epsilon} = 0.01$ and $\tilde{\epsilon} = 0.1$ is shown in Figure 6.2. In both cases the second harmonic exhibits a super-harmonic resonance peak near $\tilde{\Omega} \approx 1/2$ when the denominator is a minimum. Like the first harmonic, this resonance is shifted slightly lower by the $\tilde{\epsilon}\tilde{\beta}$ term. When the nonlinearities are larger, in addition to the super-harmonic resonance, there is also a peak near the primary resonance $\tilde{\Omega} \approx 1$; however, this peak is generally small $O(\tilde{\epsilon}^2)$ since $\tilde{A}_1$ is $O(\tilde{\epsilon})$. 

125
3. ELECTRICAL RESPONSE CALCULATIONS

The results thus far have been solely for the mechanical response. For the electrical response, we must determine the current flowing through the cantilever. Again, we will use the linearized model (truncated after $C_2$), and we will consider only the first two harmonics of the current signal. The current through a general capacitor with a dc and ac voltage is

$$i(t) = \frac{dq}{dt} = \frac{d}{dt} \left[ \left( C_0 + C_1 \zeta(t) + \frac{1}{2} C_2 \zeta^2(t) \right) \left( V_{dc} + V_{ac} \exp(j\Omega t) \right) \right]$$  (6.25)

Substituting a harmonic solution and neglecting any harmonics above the second and any terms $O(\hat{\epsilon}^2)$ and above, i.e. products of the mechanical amplitudes we obtain

$$i(t) = j\Omega \left[ C_0 V_{ac} + C_1 (A_0 V_{ac} + A_1 V_{dc}) \right] \exp(j\Omega t) +$$

$$j2\Omega \left[ C_1 (A_0 V_{ac} + A_1 V_{dc}) \right] \exp(j2\Omega t)$$  (6.26)

The first harmonic of the current signal has two terms $C_0 V_{ac} + C_1 A_0 V_{ac}$ which do not depend on the driving frequency. $A_0$ and $A_1$ are of order $O(\hat{\epsilon})$ and thus the latter two terms are negligible compared to the first, $O(1)$. These terms are the source of the parasitic capacitance in the first harmonic. The $C_0$ term contains both the static capacitance of the micro-cantilever and the stray capacitance of all nearby circuit elements. It is usually many orders of magnitude larger than the dynamic capacitance terms, $C_2, C_3, \ldots$. The primary resonance in the first harmonic of the current signal (which arises due to the $A_1$ term) is obscured by the parasitic terms, thus reducing both the Q-factor and SBR.
Because of the time derivative, a factor of $\Omega$ multiplies the first harmonic causing its amplitude to increase linearly with driving frequency, as seen in Figure 6.3. The first harmonic also possesses two terms proportional to $A_1$. These cause the small primary resonance peak to appear which is $O(\tilde{\varepsilon})$.

Figure 6.3: Electrical first harmonics from closed form harmonic balance calculations showing linearly increasing parasitic capacitance.
The second harmonic, shown in Figure 6.4, only contains terms that are proportional to either $A_1$ or $A_2$. Thus, unlike the first harmonic, the higher harmonics do not increase linearly with applied frequency and do not suffer from the parasitic static capacitance, $C_0$. Furthermore, primary and super-harmonic resonance peaks occur in the second harmonic due to the presence of both of the harmonic amplitudes.

Figure 6.4: Electrical second harmonics from closed form harmonic balance calculations absence of parasitic capacitance in higher harmonics as well as primary and superharmonic resonances.
As should be evident from the above expression, higher order harmonics differ greatly in the dominant contributions to their amplitude. For this reason, different harmonics, or different peaks in the same harmonic, will be affected more strongly by different phenomenon. This property may be utilized to design more robust HDR based sensors that measure various properties of a system by simultaneously detecting shifts in multiple resonance peaks of the same cantilever.

4. HIGHER-ORDER COMPUTATIONAL HARMONIC BALANCE SOLUTIONS

The MAPLE™ program included in the APPENDIX, follows a harmonic balance procedure exactly analogous to that given in the previous section for the low-order solutions. However, the number of terms able to be included in the Fourier series solution is far greater. For instance, five harmonics are commonly computed simultaneously without any noticeable effect on the computational time.

The program can output both the amplitude and polar response plots. Though the capacitance model is arbitrary in the program, for the purposes of the following calculations, a parallel-plate model is assumed. The mechanical response for a nonlinearity of $\bar{\varepsilon} = 0.0015$ is shown in Figure 6.5.

The mechanical response for this small nonlinearity exhibits a single peak in each harmonic at a frequency, $\Omega = \omega_0/n$, where $n$ is the order of the harmonic. Also, the resonance in the first harmonic is well defined with a relatively high Q-factor. Therefore,
there is minimum difficulty in measuring the mechanical resonance of a micro-cantilever, for instance with laser reflectometry. This is the reason why harmonics were not utilized previously in standard detection methods.

The electrical (current) response for the same level of nonlinearity is given in Figure 6.6. The electrical response typically exhibits more features than the mechanical response due to mixing of the signals. The first harmonic of the current increases linearly with frequency, which is indicative of parasitic capacitance. The resonance peak is visible in the first harmonic; however, the $Q$-factor is relatively low. In the second and higher harmonics, the $Q$-factor is much improved, which is the reason why HDR is successful as a sensing technology.
Figure 6.5: Computed mechanical frequency response of a micro-cantilever. Compare with experimental AFM results, Figure 2.3.
Figure 6.6: Computed electrical (current) frequency response spectrum of a micro-cantilever. The first harmonic, showing the parasitic effects, has a different scale. Compare with experimental HDR results, Figure 2.4.
We also notice a trend for small nonlinearities: there are generally as many super-harmonic resonance peaks as the order of the harmonic. For instance, there is a single peak visible in the first harmonic at $\omega_0$. There are two peaks visible in the second harmonic, at $\omega_0$ and $\omega_0/2$, and so on. We will see that this is not true for larger nonlinearities.

The computational harmonic balance and experimentally determined mechanical and electrical responses are presented in Figure 6.7. The theory and experiment exhibit an excellent qualitative agreement, validating the model and solution methods chosen. There are a few small discrepancies, for instance there is no primary resonance peak in the mechanical second harmonic theory but there is in the experimental results. This probably indicates that the actual value of the nonlinearity is slightly higher. Also, the electrical third harmonic at $\omega_0/2$ in the theory is greater than the second, which is also not so in the experiment. This can most likely be attributed to a deviation of the actual capacitance away from the parallel-plate model.
Figure 6.7: Comparison of experimental data and harmonic balance results for both the mechanical and electrical responses.
4.1 Polar Plots

The harmonic balance program solves for both the amplitude and phase of each of the harmonics; therefore it is possible to generate a polar plot of the results. The first harmonic is shown plotted on a polar graph in Figure 6.8.

The first harmonic of the current signal exhibits parasitic capacitance which manifests itself as a straight line on a polar plot since amplitude is changing with no associated change in phase. The resonance is seen to be a small circle offset from the origin in the first harmonic.

The polar plot of the second and third harmonics is presented in Figure 6.9. There is no linear offset as in the first harmonic, since there is no parasitic capacitance. Also, we notice that each resonance peak (primary and super-harmonic) exhibit a closed loop as they go through resonance. However, unlike a linear oscillator, the super-harmonic peaks are no longer necessarily circular, but are part of a general class of curves known as limaçons.
Figure 6.8: Polar plot showing first harmonic of current signal.
Figure 6.9: Polar plot showing second and third harmonic of current signal.
5. NONLINEARITIES, DUFFING, AND CHAOS IN A MICRO-CANTILEVER

5.1 Harmonic Balance Simulations

The overall qualitative response of the micro-cantilever system depends sensitively on the level of the nonlinearity. By decreasing the nominal gap distance or increasing the applied voltage, it is possible to tune the nonlinearity of the system. We can get a sense of these qualitative changes by computing the mechanical and electrical responses for a relatively large value of the nonlinearity, \( \tilde{\varepsilon} = 0.1 \), which corresponds to a gap distance of roughly 3 \( \mu \)m, a dc voltage of 6 V, and an ac voltage of 7 V, Figure 6.10.

When nonlinearities are relatively large, more super-harmonic resonances emerge in the mechanical response. An even more drastic change occurs in the electrical response. Near the primary resonance, the response is no longer dominated by the first harmonic, but rather by the third. These changes are related to period doubling bifurcations. As nonlinearities are increased, bifurcations continue to occur, introducing more harmonics. Eventually the response loses all periodicity and becomes chaotic.

Figure 6.11 presents HDR data related to the increase of nonlinearities. In this case the nonlinearity was increased by decreasing the damping. The system is seen to exhibit a transition from a linear resonance (circle) to a deformed loop and finally to an S curve which shows jumping or Duffing behavior. This is indicative of a bifurcation that leads to more than one stable solution, as discussed in Chapter V.
Figure 6.10: Computed response of system with large nonlinearities, $\varepsilon = 0.1$
Figure 6.11: Experimental polar plots for the electrical response illustrating Duffing jumps at high nonlinearities.
Clear MAPLE memory.
> restart;

Set Environmental Variables
> Digits := 25;
Order of capacitance power series, H.
> Eta := 3;
Number of terms in the Fourier expansion of displacement, N.
> N := 4;

Define the parameters of the simulation.
Global Parameters
> epsilon := 0.88542e-11;

Applied Voltages
> Vac := 7;
> Vdc := 6;

Cantilever Dimensions
> L := 0.250e-3;
> w := 0.35e-4;
> h := 0.2e-5;
> d := 0.15e-4;

Material Properties of Silicon Cantilever
> rho := 2330;
> E := 0.100e12;

Effective Parameters
> m := .2427*rho*w*h*L;
> k := E*w*h^3/(4*L^3);

> omega[0] := sqrt(k/m);
> frequency := evalf(omega[0]/(2*Pi));
> z := 0.12e-1;

Applied voltage (Omega is driving frequency).
> Voltage := Vdc+Vac*exp(I*Omega*t);

Assume parallel plate capacitance
> Capacitance := epsilon*w*L/(d-y);

Calculate Taylor coefficients of capacitance
> C[0] := eval(Capacitance, y = 0.);
> for i from 1 by 1 to Eta do;
>   C[i] := (eval(diff(Capacitance, `$(y, i))`, y = 0.))/factorial(i);
> end do:

Manually adjust coefficients (if necessary).

Express capacitance as Taylor power series.

> Capacitance := sum(C[eta]*y^eta, eta = 0 .. Eta);

Force on cantilever is the positive derivative of the energy of the capacitor.

> Force := (diff((1/2)*Capacitance*Voltage^2, y));

Number of terms in Fourier series of displacement, N.

> y := proc (t) options operator, arrow; sum(A[n]*exp(I*n*Omega*t), n = 1 .. N) end proc;

Substitute into equation of motion of a SHO

"(omega[0]=sqrt(k/m), 2 z=b/(sqrt(km)) )."

> diff(y(t), `$(t, 2))+(2*z*omega[0]*diff(y(t), t))+(omega[0]^2*y(t)-(eval(Force, y = y(t)))/m = 0;
> expand(%);
Perform the "Harmonic Balance", i.e. collect the coefficients for like values of n, and set equal to zero. This yields the N equations, X[i], that may be solved simultaneously to determine the coefficients of the displacement, A[n].

\[ EOM := \text{simplify}(%); \]

> for i from 1 by 1 to N do;
> X[i] := EOM;
> for j from 1 by 1 to (N*Eta+2) do;
> if (i = j) then ;
> X[i] := \text{subs}(\text{exp}(I*\text{convert}(j, \text{float}, 2)*Omega*t) = 1, X[i]);
> else;
> X[i] := \text{subs}(\text{exp}(I*\text{convert}(j, \text{float}, 2)*Omega*t) = 0, X[i]);
> end if:
> end do:
> end do:

Solve the N equations, X[i], simultaneously for the A[n].

\[ A_{\text{solved}} := \text{solve}([\text{seq}(X[i], i = 1 .. N)], [\text{seq}(A[i], i = 1 .. N)]); \]

The amplitude of vibration is the absolute value of these equations for A[n].

\[ \text{Amplitude}[i] := \text{abs}(\text{rhs}(A_{\text{solved}}[1, i])); \]
> Phase[i] := argument(rhs(Asolved[1, i]));
> end do:

Plot the amplitudes of each of the harmonics, $A[n]$, as function of driving frequency.

Plot the total amplitude of the displacement, i.e. sum of the $A[n]$'s, as function of driving frequency.

Calculate the charge and current on the cantilever.
> Capacitance*Voltage;
> subs(y = y(t), %);
> Charge := simplify(expand(%));
> diff(Charge, t);
> Current := simplify(expand(%));

Substitute the solutions for the displacement coefficients, $A[n]$, into the charge expression.
> for i from 1 by 1 to N do;
> Charge := subs(A[i] = rhs(Asolved[1, i]), Charge);
> end do:

Separate the harmonics, $X[i]$, of the charge signal.
> for i from 1 by 1 to N do;
> X[i] := Charge;
> for j from 1 by 1 to (N*Eta+2) do;
> if (i = j) then ;
> X[i] := subs(exp(I*convert(j, float, 2)*Omega*t) = 1, X[i]);
> `else``;
> X[i] := subs(exp(I*convert(j, float, 2)*Omega*t) = 0, X[i]);
> end if:
> end do:
> AmplitudeX[i] := abs(X[i]);
> end do:

Substitute the solutions for the displacement coefficients, A[n], into the current expression.

> for i from 1 by 1 to N do;
> Current := subs(A[i] = rhs(Asolved[1, i]), Current);
> end do:

Separate the harmonics, Y[i], of the current signal.

> for i from 1 by 1 to N do;
> Y[i] := Current;
> for j from 1 by 1 to (N*Eta+2) do;
> if (i = j) then ;
Define the plotting parameters.

Range of driving frequencies to plot.

> wstart := 0;
> wstop := 0.26e6;
> with(plots);

Plot the amplitudes of each of the harmonics of the DISPLACEMENT, A[i], as function of driving frequency.

> plot([seq(Amplitude[i], i = 1 .. 3)], Omega = wstart .. wstop, axes = boxed, legend = ["1st Harmonic", "2nd Harmonic", "3rd Harmonic"], color = [black], linestyle = [solid, dash, dot], thickness = 2);

Plot the amplitudes of each of the harmonics of the CURRENT, Y[i], as function of driving frequency.
> plot([(1/5)*AmplitudeY[1], seq(AmplitudeY[i], i = 2 .. 3)], Omega = wstart ..
wstop, axes = boxed, legend = ["1st Harmonic", "2nd Harmonic", "3rd Harmonic"], color = [black], linestyle = [solid, dash, dot], thickness = 2);

Plot the amplitude and phase of the current harmonics on a polar plot as a function of driving frequency.

Save amplitudes to a data file for external plotting

> with(linalg);
REFERENCES


