AN INFORMATION THEORETICAL
MODELING AND ANALYSIS OF SYSTEMS

by

BUNYAMIN OZAYDIN

MURAT M. TANIKE, COMMITTEE CHAIR
CHRISTOPHER A. GIRKIN
YEHIA MASSOUD
CHITTOOR V. RAMAMOORTHY
GREGG L. VAUGHN
EARL B. WELLS

A DISSERTATION
Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMINGHAM, ALABAMA
2012
The growth of the service industry also comes with problems and complications because of its inherent error proneness. As the complexity of systems increases, whether the system of interest is a service product or a natural phenomenon, it only proves that there is always a need for improved modeling approaches.

In this study, we approach systems from an information theoretical perspective. More specifically, we model systems by noisy communication channels and analyze them by a structure of relationships whose elements are commonly used analysis tools in science and engineering. We call this modeling approach and the structure of relationships it is built on as “noisy communication channel modeling and analysis framework.” A noisy communication channel model is a mechanism that accepts a set of inputs and yields a set of outputs.

The formal mechanisms to exploit the relationships in the proposed framework include roots of unity, special permutations and polynomials, Fourier series, and uniform polygons. For our purposes, each of these formal mechanisms represent a modeling and/or analysis tool-set they belong to: complex, combinatorial, Chebyshev, Fourier analyses, and graph theory. Relating these tool-sets to noisy communication channels – hence enabling us to look at the vast set of engineering problems they deal with from an information theoretical perspective – makes the framework powerful. The scope of the study is to establish the relationships among the mentioned formal mechanisms of these tool-sets and the noisy communication channel concept. Although some of the resulting mappings of concepts are included, further research can potentially study mapping of different concepts in informa-
tion theory to concepts in each of the tool-sets, resulting in reinterpretation of the concepts involved.

A quantum energy modeling is included as a case study for validation and better understanding of the framework.
DEDICATION

To my beloved wife, Sifanur. Without her support, this work would not have been completed. To my son, Yusuf, who has been very patient and understanding when told, “Dad needs to study.” To my parents, Mevlude and Alaattin, whose love and support have never wavered.
ACKNOWLEDGEMENTS

There are too many people to whom I owe gratitude throughout my graduate studies at the University of Alabama at Birmingham. First and foremost, I feel the greatest sense of appreciation and admiration to my advisor and committee chair, Dr. Murat Tanik, for his inspiring passion and enthusiasm for science and knowledge, his sincere friendship, unwavering support, and encouragement that helped me throughout the personal and professional hard times. I thank Drs. Vaughn and Massoud for serving as former and current department chairs, respectively, and as members of my committee. I also thank Dr. Ramamoorthy for the comments and feedback he has provided with wisdom, Dr. Wells for his services as a member of my committee, and Dr. Girkin for serving on my committee, supporting me during my employment at the Glaucoma Clinic, and willing to share his research data.

I am especially indebted to two of my colleagues that I had the privilege to work with during my graduate studies, Drs. Seker and Aktunc. I have truly enjoyed the opportunity to collaborate with Dr. Seker on many aspects of this work. He has been a very good friend, a helpful colleague, and a wise mentor. I consider myself lucky to have shared many things with Dr. Aktunc; we have been classmates, colleagues, roommates, and close friends.

I am grateful to Grimes for serving on my committee until his retirement, Dr. Jannett for his help as graduate program coordinator, Drs. Tanju and Yildirim for their encouragement and valuable discussions, and Mrs. Sandra Muhammad her administrative help.

I owe my thanks to the Department of Anesthesiology, my workplace, especially to the department chair, Dr. Jones, and my supervisor, Dr. Parks, for providing me the
time and environment to be able to complete my studies, and my colleagues that I have
the privilege to work with, Mr. Hagood, Mr. Crum, and Mr. Duke, for their support and
coverage when I had to attend my studies. I would like to also acknowledge Dr. Sakawi
and administrative assistant Ms. Kimbrough and other faculty and staff of the Department
of Anesthesiology.

I also owe my thanks to the Department of Ophthalmology and its Clinical Research
Unit, my former workplace, especially to the former department chair, Dr. Kline; my
supervisor, Ms. Keene; and Drs. Owsley, Jackson, Swanner, and Read for their continuous
support and encouragement.

Many thanks to the faculty of the Department of Physics of Marmara University,
especially the current chair, Dr. Yahsi, and my advisor during my undergraduate studies,
Dr. Altun, for instilling in me the love for science and research. I also want to acknowledge
Turk Petrol Foundation and its founder, Mr. Bolak, who initially provided the opportunity
for me to study abroad. I send special thanks to the teachers and staff of my high school,
Yamanlar Koleji.

I am thankful to the current and former students of Dr. Tanik – Drs. Dogru, Bayrak,
Jololian, Sadasivam, and Tanik – and Mr. Demirezen for their comments and contributions
on this thesis.

I owe special thanks to three special people, Ms. Shirley Stanley, Mrs. Cassandra
Page, and Mrs. Oya Tanik who have always been selflessly caring, supportive and encour-
aging, and looked after me and my family. I am grateful to my friends, Ferhat Zengul,
Mevlut Bulut, Bilal Yaldiz, Kemal Cem Gur, Hilmi Benzer, Umit Gunebir, Drs. Ramazan
Aygun and Satilmis Budak, Umit Goker, Suleyman Eris, Seydi Sahin, Levent Onat, Kursat
Arslan, and Sahin Bulut, who have supported and encouraged me throughout my graduate
studies.

I am also very thankful to my sister, Sumeyra, and my brother in law, Yusuf Ziya,
my younger sister, Betul, my parents-in-law, Serife and Omer Guner, my brothers-in-law,
Ibrahim Ethem and Muhammed Musab, and my extended family for their enormous patience and unconditional love.

Last, but not least, I thank fellow students, friends, teachers, and family members from whom I learned a lot but could not include here.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter/Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xii</td>
</tr>
<tr>
<td><strong>CHAPTER</strong></td>
<td></td>
</tr>
<tr>
<td>1 OVERVIEW OF THE STUDY</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Motivation, Objectives, and Approach</td>
<td>2</td>
</tr>
<tr>
<td>1.3 Outline</td>
<td>4</td>
</tr>
<tr>
<td>2 COMMON NOTATION</td>
<td>7</td>
</tr>
<tr>
<td>2.1 Complex Analysis</td>
<td>7</td>
</tr>
<tr>
<td>2.1.1 Periodic Functions</td>
<td>8</td>
</tr>
<tr>
<td>2.1.2 Representations of Complex Numbers</td>
<td>10</td>
</tr>
<tr>
<td>2.1.3 Complex Exponents and Roots</td>
<td>15</td>
</tr>
<tr>
<td>2.1.4 Functions of a Complex Variable</td>
<td>20</td>
</tr>
<tr>
<td>2.2 Combinatorial Analysis</td>
<td>22</td>
</tr>
<tr>
<td>2.2.1 Some Special Permutations of Interest</td>
<td>22</td>
</tr>
<tr>
<td>2.2.2 Permutation Matrices</td>
<td>24</td>
</tr>
<tr>
<td>2.3 Chebyshev Analysis</td>
<td>27</td>
</tr>
<tr>
<td>2.3.1 Definition of The Chebyshev Polynomials</td>
<td>27</td>
</tr>
<tr>
<td>2.3.2 Some Properties of The Chebyshev Polynomials</td>
<td>29</td>
</tr>
<tr>
<td>2.3.4 Chebyshev Polynomial of Second Kind</td>
<td>30</td>
</tr>
<tr>
<td>2.4 Fourier Analysis</td>
<td>32</td>
</tr>
<tr>
<td>2.4.1 Fourier Series</td>
<td>32</td>
</tr>
<tr>
<td>2.4.2 Fourier Integrals</td>
<td>40</td>
</tr>
<tr>
<td>2.4.3 Fourier Transforms</td>
<td>44</td>
</tr>
<tr>
<td>2.4.4 Discrete Fourier Transform</td>
<td>46</td>
</tr>
<tr>
<td>2.5 Uniform Polygon</td>
<td>51</td>
</tr>
<tr>
<td>2.5 Summary</td>
<td>55</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS (Continued)

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 THE NOISY COMMUNICATION CHANNEL FRAMEWORK</td>
<td>57</td>
</tr>
<tr>
<td>3.1 Noisy Communication Channel and Their Representations</td>
<td>58</td>
</tr>
<tr>
<td>3.1.1 Shannon’s Noisy Communication Channel Representation</td>
<td>59</td>
</tr>
<tr>
<td>3.1.2 Error Content Graph of Noisy Communication Channels</td>
<td>60</td>
</tr>
<tr>
<td>3.1.3 Permutation Representation of Noisy Communication Channels</td>
<td>61</td>
</tr>
<tr>
<td>3.1.4 Adjacency Matrix of Noisy Communication Channels</td>
<td>65</td>
</tr>
<tr>
<td>3.2 Relationships Between Noisy Communication Channels and Some Common Modeling and Analysis Tools</td>
<td>66</td>
</tr>
<tr>
<td>3.2.1 Noisy Communication Channels and Complex Analysis</td>
<td>66</td>
</tr>
<tr>
<td>3.2.2 Noisy Communication Channels and Combinatorial Analysis</td>
<td>67</td>
</tr>
<tr>
<td>3.2.3 Noisy Communication Channels and Chebyshev Analysis</td>
<td>77</td>
</tr>
<tr>
<td>3.2.4 Noisy Communication Channels and Fourier Analysis</td>
<td>86</td>
</tr>
<tr>
<td>3.3 Summary</td>
<td>96</td>
</tr>
<tr>
<td>4 CASE STUDY: ATOMIC MODEL</td>
<td>99</td>
</tr>
<tr>
<td>4.1 Introduction to The Quantum Theory</td>
<td>100</td>
</tr>
<tr>
<td>4.2 Bohr’s Quantum Model of Atom</td>
<td>102</td>
</tr>
<tr>
<td>4.2.1 Introduction</td>
<td>102</td>
</tr>
<tr>
<td>4.2.2 Earlier Models of Atom</td>
<td>104</td>
</tr>
<tr>
<td>4.2.3 Bohr’s Model of Hydrogen Atom</td>
<td>105</td>
</tr>
<tr>
<td>4.2.4 Quantum Mechanics and Modifications to Bohr’s Model</td>
<td>106</td>
</tr>
<tr>
<td>4.2.5 The Hydrogen Atom</td>
<td>108</td>
</tr>
<tr>
<td>4.2.6 Atomic Orbitals</td>
<td>118</td>
</tr>
<tr>
<td>4.2.7 Electron Configuration</td>
<td>125</td>
</tr>
<tr>
<td>4.3 Noisy Communication Channels Representing Atomic Orbitals</td>
<td>133</td>
</tr>
<tr>
<td>4.4 Summary</td>
<td>135</td>
</tr>
<tr>
<td>5 CONCLUSION</td>
<td>137</td>
</tr>
<tr>
<td>5.1 Summary and Conclusion</td>
<td>137</td>
</tr>
<tr>
<td>5.2 Future Work</td>
<td>139</td>
</tr>
<tr>
<td>LIST OF REFERENCES</td>
<td>143</td>
</tr>
<tr>
<td>APPENDIX</td>
<td></td>
</tr>
<tr>
<td>A MATHEMATICA IMPLEMENTATION OF THE NOISY COMMUNICATION FRAMEWORK</td>
<td>149</td>
</tr>
<tr>
<td>B EARLY DIAGNOSIS OF GLAUCOMA</td>
<td>160</td>
</tr>
<tr>
<td>B.1 Collected Data</td>
<td>161</td>
</tr>
<tr>
<td>B.2 Goals, Challenges, and Methods to Overcome Challenges</td>
<td>162</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS (Continued)

<table>
<thead>
<tr>
<th></th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>B.3 Roadmap</td>
<td>163</td>
</tr>
<tr>
<td>C IRB APPROVAL FORMS</td>
<td>164</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Chebyshev Polynomials to Degree 10</td>
<td>29</td>
</tr>
<tr>
<td>2</td>
<td>Chebyshev Polynomials of Second Kind to Degree 10</td>
<td>31</td>
</tr>
<tr>
<td>3</td>
<td>$v_p(x)$ Polynomials Up to Degree 5</td>
<td>80</td>
</tr>
<tr>
<td>4</td>
<td>7 by 3 Error Content Graph Permutations and Corresponding 7-Queen Solutions</td>
<td>87</td>
</tr>
<tr>
<td>5</td>
<td>Quantum Numbers</td>
<td>110</td>
</tr>
<tr>
<td>6</td>
<td>All Possible Quantum States for the Hydrogen Atom of $n = 2$</td>
<td>110</td>
</tr>
<tr>
<td>7</td>
<td>Orbital Angular Momentum Values for an Atom of $n = 3$</td>
<td>112</td>
</tr>
<tr>
<td>8</td>
<td>The Relation Between Quantum Numbers and Atomic Orbitals [3]</td>
<td>119</td>
</tr>
<tr>
<td>9</td>
<td>Electron Configuration of the Atoms in Their Ground State</td>
<td>127</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The modeling process (adapted from [1])</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>An information theoretical modeling framework.</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>(a) Cosine and sine functions plotted versus time, (b) Cosine and sine functions plotted versus phase.</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>Graphical representation of a complex number on the complex plane.</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>Sine wave plotted versus cosine wave.</td>
<td>13</td>
</tr>
<tr>
<td>6</td>
<td>5-th roots of unity.</td>
<td>19</td>
</tr>
<tr>
<td>7</td>
<td>A solution for ( n = 7 ).</td>
<td>23</td>
</tr>
<tr>
<td>8</td>
<td>Drawing of an 8 by 3 uniform polygon.</td>
<td>53</td>
</tr>
<tr>
<td>9</td>
<td>Drawing of an 8 by 3 uniform polygon with the alternate method.</td>
<td>53</td>
</tr>
<tr>
<td>10</td>
<td>Venn diagram of polygon sets.</td>
<td>54</td>
</tr>
<tr>
<td>11</td>
<td>Drawing of an 8 by 2 uniform polygon with the alternate method.</td>
<td>55</td>
</tr>
<tr>
<td>12</td>
<td>Polygon examples: (a) an 8-segment polygon, (b) an 8-segment stellated polygon, (c) the 8-segment ordinary polygon (octagon), (d) 8 by 3 star polygon, (e) 8 by 2 compound polygon.</td>
<td>56</td>
</tr>
<tr>
<td>13</td>
<td>A communication system.</td>
<td>58</td>
</tr>
<tr>
<td>14</td>
<td>Block diagram of a communication channel.</td>
<td>58</td>
</tr>
<tr>
<td>15</td>
<td>A communication channel.</td>
<td>59</td>
</tr>
<tr>
<td>16</td>
<td>Shannon's communication channel representation with probabilities.</td>
<td>60</td>
</tr>
<tr>
<td>17</td>
<td>Polygon (error content graph) describing the communication channel in Fig. 15.</td>
<td>61</td>
</tr>
<tr>
<td>18</td>
<td>Roots of unity on the complex plane corresponding the noisy communication channel represented in Fig. 16.</td>
<td>67</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>19</td>
<td>(a) The graph $G^2$ of graph $G$ shown in Fig. 17, (b) a stable set of $G^2$.</td>
<td>69</td>
</tr>
<tr>
<td>20</td>
<td>Venn diagram of permutation sets.</td>
<td>72</td>
</tr>
<tr>
<td>21</td>
<td>(a) 10 by 2 compound polygon, (b) corresponding chessboard configuration for case $k = 4$.</td>
<td>76</td>
</tr>
<tr>
<td>22</td>
<td>(a) Shannon representation of a noisy communication channel, (b) its corresponding 5 by 2 star polygon error content graph, and (c) its corresponding 5-queens solution.</td>
<td>77</td>
</tr>
<tr>
<td>23</td>
<td>Sides of $n$ by 1, $n$ by 2, and $n$ by 3 uniform polygons and their outer circle.</td>
<td>78</td>
</tr>
<tr>
<td>24</td>
<td>(a) 7 by 3 star polygon error content graph a noisy communication channel and (b) its Shannon representation.</td>
<td>83</td>
</tr>
<tr>
<td>25</td>
<td>(a) 8 by 3 star polygon error content graph a noisy communication channel and (b) its Shannon representation.</td>
<td>84</td>
</tr>
<tr>
<td>26</td>
<td>Evolution of 5 by 2 star figures for increasing resolution.</td>
<td>92</td>
</tr>
<tr>
<td>27</td>
<td>Evolution of 11 by 4 star figures for increasing resolution.</td>
<td>92</td>
</tr>
<tr>
<td>28</td>
<td>Polygon plot examples: (a) $u_{6,1}(p, t)$, (b) $u_{7,3}(p, t)$, (c) $u_{8,2}(p, t)$, (d) $u_{9,3}(p, t)$.</td>
<td>93</td>
</tr>
<tr>
<td>29</td>
<td>(a) Shannon representation of noisy communication channel $[2 7 5 8 1 4 6 3]$ and (b) its error content graph.</td>
<td>95</td>
</tr>
<tr>
<td>30</td>
<td>Irregular polygon $[2 7 5 8 1 4 6 3]$ constructed by Fourier series.</td>
<td>97</td>
</tr>
<tr>
<td>31</td>
<td>The relationships.</td>
<td>98</td>
</tr>
<tr>
<td>32</td>
<td>The electromagnetic spectrum.</td>
<td>101</td>
</tr>
<tr>
<td>33</td>
<td>Line spectra series of the hydrogen atom.</td>
<td>103</td>
</tr>
<tr>
<td>34</td>
<td>Angular momentum vector. (a) The allowed projections of the orbital angular momentum for the case of $\ell = 2$, (b) the orbital angular momentum vector lies on the surface of a cone and precesses about the $z$ axis when a magnetic field is applied in this direction.</td>
<td>113</td>
</tr>
<tr>
<td>35</td>
<td>Spin of an electron: (a) spin up direction, (b) spin down direction.</td>
<td>114</td>
</tr>
<tr>
<td>36</td>
<td>The probability function of hydrogen atom in the ground state.</td>
<td>117</td>
</tr>
<tr>
<td>37</td>
<td>The radial probability density functions for several states of the hydrogen atom.</td>
<td>117</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>38</td>
<td>Boundary surface diagrams for s, p, and d orbitals.</td>
<td>120</td>
</tr>
<tr>
<td>39</td>
<td>The probability distributions of the first few spherical harmonics, $</td>
<td>Y_{\ell m}\rangle^2$</td>
</tr>
<tr>
<td>40</td>
<td>(a) Orbital energy levels of the hydrogen atom, (b) orbital energy levels for the many-electron atoms, (c) orbital energies in ascending order.</td>
<td>124</td>
</tr>
<tr>
<td>41</td>
<td>(a) $u_3(t)$, (b) $u_2(t)$, (c) $u_1(t)$, (d) $u_0(t)$, (e) $u_1(t)$, (f) $u_2(t)$, (g) $u_3(t)$, (h) $u'_3(t)$, (i) $u''_4(t)$.</td>
<td>135</td>
</tr>
<tr>
<td>42</td>
<td>The modeling process.</td>
<td>138</td>
</tr>
<tr>
<td>43</td>
<td>An information theoretical modeling framework.</td>
<td>139</td>
</tr>
<tr>
<td>44</td>
<td>Study of natural processes as communication systems.</td>
<td>140</td>
</tr>
<tr>
<td>45</td>
<td>Mathematica output screen-shot for the 5 by 2 error content graph</td>
<td>157</td>
</tr>
<tr>
<td>46</td>
<td>Mathematica output screen-shot for the 6 by 2 error content graph</td>
<td>158</td>
</tr>
<tr>
<td>47</td>
<td>Mathematica output screen-shot for the 9 by 4 error content graph</td>
<td>159</td>
</tr>
</tbody>
</table>
CHAPTER 1
OVERVIEW OF THE STUDY

In this chapter, we will start with an introduction followed by a section where we talk about the motivation to study this subject, our objectives, and the approach we take to achieve them. We conclude this chapter with an outline of the entire dissertation.

1.1. Introduction

Modeling and analysis of systems in general is a difficult problem for engineers and scientists. With the increase of storage and data acquisition capabilities, the amount of data in many different disciplines demanding to be analyzed has never been as huge. Technological improvements and increasing amount of research being performed only promises more of this demand. However, the abundance and variety of data also increase modeling opportunities [4, 5].

On the other hand, the nature of current interdisciplinary research creates a need for analysis of multi-dimensional data (multiple types of data collected about a single phenomena.) Current analysis approaches (such as statistical analysis, pattern recognition and classification techniques, Fourier analysis, and complex analysis) have difficulty analyzing multi-dimensional data. Current modeling approaches, such as simulators, on the other hand, do not include a closed-form analytical modeling; in other words, they are descriptive rather than deductive. The biological and other natural complex systems are being approached either as particles or as waves. Recently, however, different approaches are being studied more and more. For example, in a very recent paper published in Science magazine, a study presented step-by-step instructions to assemble nano-particles into macroscopic materials using biological information processing; that is, DNA processing
[6, 7]. In this study, we also propose another information theoretical approach to address some of those issues.

A model is a simpler realization or idealization of more complex real-world entities. A model may be a physical miniature of a real-world system or a computer simulation of it. Such models aim to increase one’s understanding of the real systems or even better, to reveal some new information about them. Also, it is likely that investigation of the model rather than the more complex real system becomes more feasible.

As R. W. Hamming points out in his paper titled “The Unreasonable Effectiveness of Mathematics,” mathematical models of real-world phenomena are very much effective in understanding the world around us, even though there is no solid reason to explain why they are [8]. For this and other reasons, mathematical modeling or its various forms, such as informational theoretical modeling that is as rigorous as mathematical modeling, is very efficient, especially for computers that are designed to run with rigorous deterministic algorithms. A mathematical modeling process begins with the representation of a real-world problem by a mathematical model, followed by a solution of the model, and then interpretation of the solution in the real world. The process continues with validation, as the real-world interpretation of the mathematical solution is checked against the real-world behavior of the system. The results of the validation lead to adjustments to the model. Finally, the process is said to be complete after it is repeated to a degree of satisfaction [1]. Although the information theoretical modeling presented in this study is a form of mathematical modeling and follows a similar process, it has not yet enjoyed the recognition of more familiar mathematical models, such as differential equations.

1.2. Motivation, Objectives, and Approach

Information theoretical concepts have been used in modeling systems in a variety of fields, such as medicine and finance [9, 10]. Entropy and probability concepts are the parts of information theory that are used in most models found in the literature. In our approach,
however, the communication channel, particularly the noisy communication channel, plays the main role.

Similar to the mathematical modeling process, the information theoretical modeling process we follow is illustrated in Fig. 1. This process begins by developing the noisy communication channel model for a given real-world system we want to analyze. Following model development, the model is analyzed by rigorous mathematics of the information theory. The results produced by this analysis are then interpreted to the real world, and finally, the modeling and analysis framework is put through validation by checking the results against the real-world system. Based on the success of the validation, adjustments are made to the model.

Fig. 1. The modeling process (adapted from [1])

One purpose of modeling is to be able to benefit from the power of a rigorous field as a tool to understand and analyze a real-world system. Therefore, we base our modeling on the mature and mathematically rigorous field of information theory. Considering the multidisciplinary demand of system modeling and analysis, in this study we approach this problem from an interdisciplinary perspective by modeling systems (mechanisms that interpret their inputs and yield their outputs) as noisy communication channels. More explicitly, our information theoretical approach is the application of Shannon’s noisy communication channel concept to model and analyze systems that are currently analyzed by traditional methods [11, 12]. Our motivation leading to this objective is the existence of a deep re-
relationship between the complex analysis in engineering and the communication channel concept in information theory. We aim to reveal this relationship and provide a framework of tools to be used interchangeably approaching the problems of various disciplines.

Through the use of this interdisciplinary approach, we establish a common notation among some of the commonly used analysis tools from a noisy communication channel perspective and, based on this common notation, we develop a modeling and analysis framework that includes relationships between the noisy communication channel concept and each of the analysis tools included in the common notation. The analysis tools included in this framework are complex analysis, combinatorial analysis, Chebyshev analysis, and Fourier analysis (Fig. 2.)

![Fig. 2. An information theoretical modeling framework.](image)

1.3. Outline

In this chapter, an overview of the study is provided, including introduction, motivation and objectives of the study, proposed approach, and outline of this document.

In the second chapter, a common notation is formed consisting of some topics of interest in complex, combinatorial, Chebyshev, and Fourier analyses and graph theory. While forming this common notation, specific features or representations of these mathematical tools are emphasized. We started the study of complex analysis with periodic functions, as
they are also essential in Fourier analysis. Then several representations of complex numbers are presented, followed by complex exponents and roots and complex functions. The combinatorial analysis section includes the topics that are relevant to this study, namely some special set of permutations that correspond to the solutions of the N-queens problem and permutation matrices. The Chebyshev analysis discussion is also limited to mainly Chebyshev polynomials and matrices; definitions and relevant properties are presented. Considering its vast application areas and the large number of tools it consists of, a rather comprehensive Fourier analysis section is included. Fourier series, integrals, and transforms are discussed, along with their different forms and representations, then the discrete Fourier transform, including the fast Fourier transform and Fourier matrices, are presented. The chapter concludes with a section presenting the uniform polygons along with their classifications, definitions, and properties, followed by a summary section. In addition to serving as a survey, the entire second chapter also introduces the subjects in such a manner as to prepare the reader for the third chapter where the noisy communication channel framework is introduced by establishing a common notation among the topics discussed in its sections.

The third chapter introduces an information theoretical modeling and analysis of systems framework and consists of two main sections and a summary section. Its first section presents the concept of the noisy communication channel and its representations, namely Shannon’s, error content graph, and permutation representations. In its second section, thanks to the common notation introduced in the previous chapter, the framework is built by establishing relationships between the noisy communication channel concept (via its uniform polygon representation) and each of the analysis tool-sets: the relationship to complex analysis via roots of unity and their representation on the complex plan, the relationship to combinatorial analysis via some special permutations, the relationship to Chebyshev analysis via eigenvalues of Chebyshev matrices, and the relationship to Fourier analysis via Fourier series representing regular and irregular uniform polygons. Also a
subsection is included that discusses a combined relationship among noisy communication channels, Chebyshev polynomials and matrices, and permutations. The chapter is concluded with a summary section.

A case study is presented in the fourth chapter. The ideas introduced are applied to model atom starting from Bohr’s model and introducing an information theoretical perspective. In its first section, a survey is presented, including all aspects of Bohr’s model and beyond. Then Fourier series representations of atomic energy levels are introduced in the second section.

The fifth chapter concludes the dissertation by offering sections for summary, conclusion, and future study.

Two appendices are included. The first appendix presents a Mathematica implementation of the noisy communication channel framework and explains the functions utilized. It also includes screen-shots of this implementation’s output. The second appendix talks about preliminary and planning stages of a data analysis research.
CHAPTER 2
COMMON NOTATION

In this chapter, various mathematical tools to be used in the construction of the proposed modeling and analysis framework are prepared to form a common notation. First, these tools are introduced briefly, and then their specific features or representations that would enable us to connect them to the framework are studied in more detail. This way, they are made ready to be easily used in the next section, where the relations among them and their relations to the framework are introduced. Although these topics are covered at many places in the literature, here we introduce them with several considerations in mind: a) We investigate these topics from an information theoretical perspective, and b) the notation and nomenclature are developed in a way so that they would easily fit into the relationships structure introduced in Chapter 3.

Four analysis tool-sets – namely, complex analysis, combinatorial analysis, Chebyshev analysis, and Fourier analysis – along with a geometrical tool, namely, uniform polygons are discussed. First, we start with complex numbers and functions, with the focus on the exponential functions and the roots of unity. Second, we summarize permutations, with the focus on some special permutations. Third, we survey Chebyshev polynomials, focusing on the second kind of Chebyshev polynomials and their recursion property. Fourth, we go over Fourier series, emphasizing their exponential form. Lastly, we investigate polygons in general and uniform polygons in particular.

2.1. Complex Analysis

Considering the scope of this study, instead of a deductive, comprehensive survey of complex analysis, we will assume that the reader has a basic understanding of the complex
numbers and directly go into different representations of them. We will begin by introducing periodic functions and some relevant concepts, because they are important with some of the complex number representations and other concepts introduced later in this chapter. We will conclude with a section that summarizes the most commonly used functions of a complex variable. The style and notation here has been composed from [13, 14, 15, 16].

2.1.1. Periodic Functions

Let us think of a signal (wave)

\[ x(t) = A \cos (wt + \phi) \]  

(2.1)

where, \( A \) is the amplitude, \( w \) is the angular (radian) frequency \((\text{rad/sec})\), and \( w = 2\pi f \) with \( f \) being the cyclic frequency, having the unit of \( \text{Cycles/sec} \) or Hertz. \( \phi \) is called the phase shift, representing how much the signal \( x(t) \) shifted from its actual phase \((w.)\). For this specific case, \( x(t) \) is a continuous function of time \( t \).

**Definition 1.** A function \( x(t) \) is said to be periodic by \( T \) if \( x(t) = x(t + kT) \) where \( k = 0, 1, 2, \ldots \).

**Remark 1.** A periodic signal having the cyclic frequency \( f_0 \) cycles per second, has a period \( T = 1/f \). Period is the required time to complete a full cycle.

In Fig. 3a, \( \sin 2\pi t \) and \( \cos 2\pi t \) functions are plotted. Following the notation above, one notices that the cyclic frequencies for both these functions are \( f = 1 \) Hertz (Hz).

**Remark 2.** It is known that \( \sin^2 \theta + \cos^2 \theta = 1 \) and \( \cos 2\theta = \cos^2 \theta - \sin^2 \theta \). Using these two equalities, the following can be derived:
As observed in Fig. 3a, sine and cosine waves are closely related. Derivative of sin θ is equal to cos θ, and derivative of cos θ equals to negative of sin θ; namely, $d\cos \theta/d\theta = -\sin \theta$. That is to say that the cosine function is the slope of the sine function, and sine function is the negative slope of the cosine function.

By observing Fig. 3a, we see that cosine function would overlap sine function if it were to be shifted by 0.25 seconds. The signal drawn with a solid line in Fig. 3a is $\cos 2\pi t$. If we want to represent the cosine signal in the form of $\cos \theta$, we can define $\theta = 2\pi 1t$. Notice that the cyclic frequency for this function is 1Hz. If the time shift $t = 0.25 \text{sec}$ is substituted into $\theta = 2\pi 1t$, we find $\theta = \pi/2$. This substitution is a phase shift to the cosine function. Hence, one can say that the phase shift between sine and cosine functions is $\pi/2$ radians, which is demonstrated more clearly in Fig. 3b.

**Remark 3.** We can rewrite (2.1) to show that the phase-shifted cosine function equals to sine function, or vice versa:

\[
\cos^2 \theta = \frac{1}{2} (1 + \cos 2\theta)
\]
\[
\sin^2 \theta = \frac{1}{2} (1 - \cos 2\theta)
\]
\[ x_1(t) = A \cos(w_0t + \phi) = A \sin(w_0t + \phi + \pi/2) \]
\[ x_2(t) = A \cos(w_0t + \phi - \pi/2) = A \sin(w_0t + \phi) \]  

(2.3)

The amplitude \( A \) is such that the function oscillates between \(-A\) and \( A \). Therefore, (2.1) can be written in terms of either sine or cosine. Here we prefer to express it with cosine.

**Definition 2.** A function that is symmetric with respect to the axis belonging to its range (in this case the vertical axis) is said to be an **even function**. Algebraically, a function is said to be even if \( x(t) = x(-t) \).

**Example 1.** \( \cos \theta = \cos(-\theta) \)

**Definition 3.** A function \( x(t) \) is said to be an **odd function** if it is symmetric with respect to the origin, or algebraically speaking, if \( x(t) = -x(-t) \).

**Example 2.** \( \sin \theta = -\sin(-\theta) \).

**Remark 4.** A function that does not satisfy any of these conditions is called **neither even nor odd**.

2.1.2. **Representations of Complex Numbers**

In this section, we will investigate four different ways to represent complex numbers [14]:

1. Points in two-dimensional space \((x, y)\) components on Cartesian coordinates and \(r, \theta\) on polar coordinates in exponential form.)

2. Ordered pairs of real numbers, with the complex operator \( j \) hidden in the second element of each pair.

3. \(2 \times 2\) matrices with real elements.

4. Equivalence classes of polynomials with real coefficients.
2.1.2.1. Points in two-dimensional space. We can graphically represent a complex number on the complex plane, as shown in Fig. 4. Definitions of sin and cos from trigonometry indicate that \( \sin \theta = \frac{y}{r} \) and \( y = r \sin \theta \). Similarly, \( \cos \theta = \frac{x}{r} \) and \( x = r \cos \theta \). The complex number is represented by \( z = (x, y) \); where \( x = \text{Re} \{z\} \) is the real part, and \( y = \text{Im} \{z\} \) is the imaginary part. The imaginary part of the complex number is represented with a prefix \( j \), where \( j = \sqrt{-1} \). The complex number in Cartesian form can be written as \( z = x + jy \). In Fig. 4, a complex number is represented as a point \( (z) \) in complex plane, having its real and imaginary parts as its coordinates. These points in complex plane are analogous to vectors in two-dimensional space. Just as vectors have length and directions, the Cartesian form of complex numbers can also be expressed similarly in polar form as \( z = re^{j\theta} \), where the length of the vector is \( r \), and it makes an angle of \( \theta \) with the real axis in counterclockwise direction. If the angle were to be in clockwise direction, then it would have a negative value.

Remark 5. One can utilize trigonometry and the Pythagorean theorem to compute Cartesian coordinates \( (x, y) \) in terms of the polar variables \( r \) and \( \theta \) of \( z = re^{j\theta} = x + jy \). By using Pythagorean theorem, we can compute \( x, y, \) and \( r \) as \( x = r \cos \theta \), \( y = r \sin \theta \), and \( r = \sqrt{x^2 + y^2} \). Note that both polar and Cartesian forms are equivalent in representing complex numbers as points in two-dimensional space.
Rather than using $z = re^{j\theta}$, one can use the cos/sin form $e^{j\theta} = \cos \theta + j\sin \theta$ that is provided by Euler for the complex exponential. Thus, the Cartesian pair $\cos \theta$ and $\sin \theta$ represents any point on a circle with radius of 1. By writing Euler’s representation for a more general case, we get $z = re^{j\theta} = r\cos \theta + jr\sin \theta$. In this manner, a complex exponential signal can be written as

$$f(t) = Ae^{j(w_0t + \phi)}$$

(2.4)

In (2.4), $f(t)$ is a complex valued function of time $t$. The magnitude of this function is $|f(t)| = A$, and its argument is $\arg\{f(t)\} = w_0t + \phi$, which corresponds to $\theta$. It is clear that $\theta$ is a function of time. With the help of Euler’s representation $e^{j\theta} = \cos \theta + j\sin \theta$, the complex function can be shown in Cartesian form as

$$f(t) = Ae^{j(w_0t + \phi)} = A\cos (w_0t + \phi) + jA\sin (w_0t + \phi)$$

(2.5)

A plot of the real part of the function versus its imaginary part is shown in Fig. 5. This plot illustrates rotation of the phasor in the complex plane.

2.1.2.2. Ordered pairs of real numbers. Consider two complex numbers, $z_1 = x_1 + jy_1$ and $z_2 = x_2 + jy_2$. We can represent them as $(x_1, y_1)$ and $(x_2, y_2)$. For these two ordered pairs to be equal, the corresponding numbers in both pairs must be equal; namely, if $(x_1, y_1) = (x_2, y_2)$ then $x_1 = x_2$ and $y_1 = y_2$. The addition on complex numbers is defined as $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$, while the multiplication is $(x_1, y_1) \times (x_2, y_2) = (x_1x_2 - y_1y_2, x_1y_2 + y_1x_2)$. The multiplication operation assumes that the complex operator $j$ is “hidden” with the second element of every ordered pair. Subtraction is rather trivial, but division is more complicated than multiplication: $\frac{(x_1, y_1)}{(x_2, y_2)} = (x_1, y_1) \times \frac{1}{(x_2, y_2)}$. 


where \( \frac{1}{(x_2, y_2)} = \left( \frac{x_2}{x_2^2 + y_2^2}, \frac{-y_2}{x_2^2 + y_2^2} \right) \). One can justify this last equality by taking the product of \((x_2, y_2) \times \frac{1}{(x_2, y_2)}\), which equals to \(\left( \frac{x_2y_2 - y_2(-y_2)}{x_2^2 + y_2^2}, \frac{x_2(-y_2) + y_2y_2}{x_2^2 + y_2^2} \right)\). Looking carefully, we see that the previous ordered pair is reduced to \((1, 0)\).

\[\text{Fig. 5. Sine wave plotted versus cosine wave.}\]

2.1.2.3. **Real valued matrices.** Let us consider the matrix \(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\). By using matrix algebra, we find the square of this matrix as \(\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\), where \(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\) is unity matrix of size 2. Hence, we can say that \(j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\), because its square is equal to negative unity. Applying the Cartesian representation of complex numbers \((z = x + jy)\), we can write a complex number \(z\) in matrix form as \(z = x \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + y \begin{pmatrix} 0 & 0+y \\ 0-y & x \end{pmatrix} = \begin{pmatrix} x+0 & 0+y \\ 0-y & x+0 \end{pmatrix}\), which is also equivalent to \(z = \begin{pmatrix} x \\ y \\ -y \\ x \end{pmatrix}\).

2.1.2.4. **Equivalence classes of polynomials with real coefficients.** Let us consider the polynomial \(a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + a_nx^n\), where \(a_k = 0 \ldots n \in \mathbb{R}\). This polynomial is said to have or is said to be of degree \(n\). Here we are particularly interested in \(\mod (x^2 + 1)\) equivalence (residue) classes of polynomials in \(\mathbb{R}[x]\).
Remark 6. For any polynomial \( p(x) \) in \( \mathbb{R}[x] \), one can write \( p(x) = q(x)\left(x^2 + 1\right) + r(x) \), where \( q(x), r(x) \in \mathbb{R}[x] \), and \( 0 \leq \deg r(x) < 2 \). \( p(x) \) is said to be equivalent to \( r(x) \), which is a polynomial of degree 0 or 1 in the form of \( a_0 + a_1x \). Two polynomials, regardless of their degrees, are said to be equivalent when they have the same remainder as a result of the division by \( 1 + x^2 \). Obviously, there will be more than one polynomial falling into same class, and the polynomials that fall into the same class are said to be members of an equivalence class.

Remark 7. Just as in modular arithmetic, two equivalent classes \([a'_0 + a'_1x]\) and \([a''_0 + a''_1x]\) are equal if and only if \((a'_0 + a'_1x) - (a''_0 + a''_1x)\) is divisible by \( x^2 + 1 \). For \((a'_0 - a''_0) + (a'_1 - a''_1)x\), a polynomial of degree 1, to be divisible by \( x^2 + 1 \), it has to be a 0 polynomial. Hence \( a'_0 - a''_0 = 0, \ a'_1 - a''_1 = 0; a'_0 = a''_0, \ a'_1 = a''_1 \). Therefore, we conclude that the set of all residue classes of \( \mathbb{R}[x] \mod (x^2 + 1) \) is \( \{[a_0 + a_1x]; a_0, a_1 \in \mathbb{R}\} \).

Remark 8. Addition operation on this residue class can be expressed as

\[
[a'_0 + a'_1x] + [a''_0 + a''_1x] = [(a'_0 + a''_0) + (a'_1 + a''_1)x].
\]

Multiplication operation can be expressed as

\[
[a'_0 + a'_1x] [a''_0 + a''_1x] = [a'_0a''_0 + (a'_0a''_1 + a'_1a''_0)x + a'_1a''_1x^2].
\]

Since the degree of the resulting polynomial is neither 0 nor 1, we need to divide it by \( 1 + x^2 \) to find its residue class. The quotient of this division is \( a'_1a''_1 \), whereas its residue is \((a'_0a''_1 + a'_1a''_0)x + (a'_0a''_0 + a'_1a''_1)\). Hence, multiplication can be expressed as

\[
[a'_0 + a'_1x] [a''_0 + a''_1x] = (a'_0a''_1 + a'_1a''_0)x + (a'_0a''_0 + a'_1a''_1).
\]

A special subset of the set of residue classes consists of only the equivalent classes where \( a_1 = 0 \) and \( \{a_0; a_0 \in \mathbb{R}\} \). There is a one-to-one correspondence between the ele-
ments of this set and the real numbers. Under the two operations we have just defined, these equivalent classes behave just like real numbers:

\[ [a_0] + [a_0'] = [a_0' + 0x] + [a_0'' + 0x] = [a_0' + a_0''] \]

\[ [a_0'] [a_0''] = [a_0' + 0x] [a_0'' + 0x] \]
\[ = [(a_0 a_0'' - 0) + (0 + 0)x] = [a_0 a_0''] \]

Notice that the set of residue classes contains a subfield, which is an isomorphic copy of the real numbers. Assuming \([a_0 + a_1x]\) notation for the residue classes, this subset with addition operation has all the group properties with identity \([0 + 0x]\) and inverse \([-a_0 + (-a_1)x]\).

Also, with multiplication operation, it has all the group properties with identity \([1]\) and inverse \([a_0/(a_0^2 + a_1^2) - (a_1/(a_0^2 + a_1^2))x]\). We conclude that the set of residue classes is a field; moreover, this field has a subfield that corresponds to real numbers.

Now we will see how this subset represents complex numbers. For a particular equivalent class \([x]\), we see that \([x] [x] = [x^2]\). Since \(x^2 + 1\) divides the difference of \(x^2\) and \(-1\), \([x] [x] = [-1]\). In this field, there is an element \([x]\) whose square is that element which corresponds to \(-1\) in real numbers. Suppose that we name this element \(j\) and we designate element \([a]\) as \(a\). Then we can write \([a + bx]\) as \([a + xb] = [a] + [x] [b] = a + jb\), which is the standard notation of complex numbers.

### 2.1.3. Complex Exponents and Roots

Among the representations we mentioned in the previous section, we will mainly use the exponential form of complex numbers, considering its ease of use in mathematical operations, such as exponents and roots of complex numbers. Let us start by introducing De-Moivre’s Theorem:

**Theorem 1.** If \(n\) is any positive integer and \(z = A (\cos \theta + j \sin \theta)\), then
\[ z^n = [A (\cos \theta + j \sin \theta)]^n = A^n (\cos n\theta + j \sin n\theta) \]

= \[ A^n e^{in\theta} \]

and

\[ w_k = z^{\frac{1}{n}} = A^{\frac{1}{n}} \left\{ \cos \left[ \frac{(\theta + 2\pi k)}{n} \right] + j \sin \left[ \frac{(\theta + 2\pi k)}{n} \right] \right\} \]

\[ = A^{\frac{1}{n}} e^{j(\theta + 2\pi k)/n}, \quad k = 0, 1, \ldots, n - 1 \] (2.6)

where \( A \) is the magnitude, and \( \theta \) is the argument of \( z \) in complex plane.

**Remark 9.** Because of the fact that \( \theta \) is determined within a multiples of \( 2\pi \), and both sine and cosine are periodic with \( 2\pi \), the values of \( k = n, n + 1, \ldots \) are not considered. The values of sine and cosine for the argument \( (\theta + 2\pi k) \) will be equal at \( k = n \) and \( k = 0, k = n + 1 \) and \( k = 1, k = n + 2 \) and \( k = 2 \), etc.

**Remark 10.** We use De-Moivre’s theorem for any complex number \( z = Ae^{j\theta} \) and positive integer \( n \) to express the unique complex number \( z^n \) and the \( n \) complex roots of the equation \( w^n = z \). Each of these \( n \) roots \( (w_k, k = 0, 1, 2, \ldots, n - 1) \) has the magnitude of \( A^{\frac{1}{n}} \). This means that the \( n \) roots lie on a circle of radius \( A^{\frac{1}{n}} \) that is centered at the origin of the complex plane. Each of the \( n \) roots corresponds to a point on this circle, and they are equally spaced on the circle; hence, the arc length between any two consecutive points is equal.

**Theorem 2.** Any complex number \( z = Ae^{j\theta} = A (\cos \theta + j \sin \theta) \) that is different than zero has exactly \( n \) distinct complex roots, as described in De-Moivre’s Theorem, and these roots can be represented by \( n \) points that are equally spaced on a circle, which is centered at the origin of the complex plane with the radius of \( A^{\frac{1}{n}} \).

**Remark 11.** In order to visualize the locations of the roots on the complex plane, it is useful to investigate the arguments of the roots, since their magnitudes are all the same. Let \( \alpha_k \)
represent arguments of the roots and $\Delta \alpha$ represent the difference between each consecutive arguments of the roots. From (2.6), we see that $\alpha_0 = \theta/n$, $\Delta \alpha = 2\pi/n$, and $\alpha_k = \alpha_0 + k\Delta \alpha$. Then we can easily place each root on the circumference of the circle, with radius $A^{1/n}$ centered at the origin of the complex plane at $\alpha_0$ and each $k\Delta \alpha$ further (counter-clockwise direction) from it.

Remark 12. We can also rewrite the generic form of the roots equation in terms of the root where $k = 0$. From (2.6), $w_0 = A^{1/n}e^{i\theta/n} = A^{1/n}e^{i\alpha_0}$. We also know that $w_k = A^{1/n}e^{i\alpha_k}$. Then $w_k$ can be rewritten as

$$w_k = A^{1/n}e^{i\alpha_0}e^{ik\Delta \alpha}$$
$$= w_0 e^{ik\Delta \alpha}$$

Definition 4. The function $z^n$ has a unique value for any positive integer $n$. The function $z^{1/n}$, on the other hand, may take on any one of the $n$ values defined in (2.6). When $z$ is a real number, these possible values of the function $z^{1/n}$ include a real number. This real value is called the principle value of $z^{1/n}$ in the set of complex numbers.

Remark 13. The principle value of $z^{1/n}$ is formulated by the following:

1. When $z$ is a positive real number, or it is a negative real number and $n$ is even, the principle value is $w_0 = A^{1/n}e^{i\theta/n}$;

2. When $z$ is a negative real number and $n$ is odd, it is $w_{(n-1)/2} = A^{1/n}e^{i(\theta + 2\pi(n-1)/2)/n} = A^{1/n}e^{i(\theta + \pi(n-1)/2)/n}$;

3. We shall not attempt to designate principal values of $z^{1/n}$ when $z$ is imaginary.

Remark 14. Let us consider the special case where the complex number $z = Ae^{i\theta}$, whose roots we are considering, is equal to 1. In this case, the magnitude of this complex number is $A = 1$, and its argument is $\theta = 0$. Consequently, $\alpha_0 = 0$, $\alpha_k = k\Delta \alpha$, $w_0 = 1$, and $w_k = e^{ik\Delta \alpha} = e^{ik2\pi/n}$. 
Definition 5. The $n$-th degree roots of complex number 1 are called $n$-th roots of unity. For the cases second and third degrees, they are called square roots of unity and cube roots of unity, respectively.

Remark 15. Regardless of the degree of roots we are dealing with, the principle value of the roots of complex number $z = 1$ is always $w_0 = 1$. $w_1$ is also significant, since the rest of the roots are simply the $k$-th exponents of $w_1$. We can easily show this from $w_k$ formula in the previous remark: $w_1 = e^{j2\pi/n}$ and $w_k = e^{jk2\pi/n} = w_1^k$. Because of this significance, $w_1$ is simply denoted as $w$ in the literature. The set of roots of unity then is denoted as \{$w, w^2, w^3, \ldots, w^n$\}. Please note that $w^n = e^{j2\pi} = 1$, which is also equal to $w_0$.

Example 3. Let’s think on the case when $z = w^3 = 1$, the cube roots of unity. The roots of this equation satisfy $z = w^3 = 1e^{j0} = 1(\cos 0 + \sin 0)$. Therefore, the cube roots of 1 may be expressed as $w = e^{j2\pi/3} = \cos \left[2\pi/3\right] + j\sin \left[2\pi/3\right]$ and its square and cube. More explicitly,

\[\begin{align*}
w &= \cos\frac{2\pi}{3} + j\sin\frac{2\pi}{3} = -1/2 + j\sqrt{3}/2 \\
w^2 &= \cos\frac{4\pi}{3} + j\sin\frac{4\pi}{3} = -1/2 - j\sqrt{3}/2 \\
w^3 &= \cos\frac{6\pi}{3} + j\sin\frac{6\pi}{3} = \cos 2\pi + j\sin 2\pi = 1.
\end{align*}\]

Remark 16. The points that represent $w$, $w^2$, and $w^3$ on the circle centered at origin of the complex plane and having the radius $A^{1/n}$ are vertices of an equilateral triangle inscribed in a unit circle ($A = 1$) and having one vertex at $(1,0)$ on the positive $x$-axis. In general, the $n$-th roots of unity can be represented by the vertices of a regular polygon of $n$ vertices (and sides), inscribed in the unit circle centered at the origin, with one vertex of the inscribed regular polygon lying at $(1,0)$ on the positive $x$-axis. The 5-th roots of unity are shown in Fig. 6.

Remark 17. Looking from a different point of view, the $n$-th roots of unity form a cyclic group of $n$ elements under multiplication with $w_0 = 1$ as the identity element. In cyclic groups, every element of the group can be expressed as a function of a single element
called the generator. For $n$-th roots of unity cyclic group, any root whose powers from 1 to $n$ generates all the elements of its group is called a primitive $n$-th root. As described in Remark 15, $w_1 = e^{j2\pi/n}$ is always a primitive $n$-th root of unity, since its powers generate all other roots. For any positive integer $n$, there exists at least one primitive $n$-th root of unity such that $w^n = 1$. Note that not all $n$-th roots of unity are primitive $n$-th roots of unity.

**Example 4.** For $w^3 = 1$, one can represent all of the three cube roots of unity as $w_1, w_1^2, w_1^3$ or as $w_2, w_2^2, w_2^3$. So, for the cube roots of unity, both $w$ ($w_1$) and $w^2$ ($w_2$) are primitive cube roots of unity. For $w^4 = 1$, only $w$ and $w^3$ are primitive 4th roots of unity and $w^2$ is not, since its powers does not generate $w$ and $w^3$.

**Theorem 3.** If $s$ is a primitive $n$-th root of unity, all $n$-th roots of unity are given by the sequence $s, s^2, s^3, \ldots, s^{n-1}, s^n = 1$.

**Remark 18.** The primitive $n$-th roots of unity are precisely the numbers $s^m$, where $m$ can be any number that is relatively prime to $n$, and $s$ is any primitive $n$-th root of unity [13]. Since we already know that $w = w_1 = e^{j2\pi/n}$ is always a primitive $n$-th root of unity, then the set of primitive $n$-th roots of unity can be defined as $\left\{ e^{jm2\pi/n} \ | \ m \text{ and } n \text{ are relatively prime} \right\}$.
2.1.4. Functions of a Complex Variable

**Definition 6.** Let \( f(x) \) be a function that is defined for variable \( x \in \mathbb{R} \). This function will work when we replace the variable \( x \) with \( z \), which is a complex variable, because the complex numbers form a field [15]. Therefore, if we have a polynomial function of a real variable \( f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \), then for complex argument \( z \), we have \( f(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n \). The function \( f(z) \) is defined as a *function of a complex variable* and maps every point in the complex plane to another point in the same plane.

**Remark 19.** When we want to visualize a real function, we simply plot it. As in the case of a complex function, we have maps of planes to visualize functions of a complex variable.

**Example 5.** \( f(z) = z + z_1 \) translates the complex plane by vector \( z_1 \), whereas \( f(z) = zz_1 \) with \( \|z_1\| = 1 \) rotates the plane by an angle of \( \arg z_1 \).

**Example 6.** \( f(z) = z^2 \) maps each point to another point whose distance from the origin squared and argument doubled (2\( \arg z \)).

**Remark 20.** Consider the rational functions of a complex variable, and let \( f(z) = \frac{g(z)}{h(z)} \) defined for all \( z \), except those that make \( h(z) = 0 \). Let \( a, b, c, \) and \( d \) be complex constants with the condition \( ad - bc \neq 0 \), and let \( f(z) = \frac{az + b}{cz + d} \) be defined as a *fractional linear transformation* of the complex plane to itself at every point but \( z = -d/c \), where \( c \neq 0 \). Examining the limiting conditions, one can also say that this particular function maps \( z = -d/c \) to infinity, and infinity to \( a/c \).

**Remark 21.** Similarly, one can define an exponential function for a complex variable. From the basics of exponential functions, we know that \( e^{x+y} = e^x e^y \); hence, for a complex variable \( z = x + jy \), we will have \( e^z = e^{x+jy} = e^x e^{jy} \). Substituting Euler’s relation \( e^{jy} = \cos y + j \sin y \) in this equation, we get \( e^{x+jy} = e^x (\cos y + j \sin y) \).

**Remark 22.** The function \( w = f(z) = e^z = e^{x+jy} = e^x (\cos y + j \sin y) \) maps the infinite strip \( \{x \mid -\infty \leq x \leq \infty \} \) with width of \( 2\pi \), \( \{y \mid 0 \leq y \leq 2\pi \} \), in \( z \)-plane to the whole \( w \)-complex.
plane except the origin, since $e^z$ can never be zero. One can see that each horizontal line in this strip (that is, $y$ is constant), becomes a ray (constant $\theta$) from the origin with different lengths in complex plane. Similarly, each vertical line (that is, $x$ is constant) in the domain of this function corresponds to a circle in the complex plane. The logarithm, which is the inverse of exponential function, needs to be a multiple valued function. Let us now think of points in $w$-plane, such that $w = e^z (\cos y' + j \sin y')$. We can alternatively write this equation as $w = e^{\log_e e^{y' + j(y' + 2k\pi)}} = e^{y' (\cos y' + j \sin y')}$. Hence, for each value of $k$ and $w \neq 0$,

$$z' = \log_e w' = \log_e e^{y'} + j(y' + 2k\pi), \quad 0 \leq y' \leq 2\pi$$

(2.7)

is an inverse for the exponential function $w = f(z) = e^z$. The result of (2.7) for every $k$ value is called a *branch* of the multiple valued function $g(z) = \log_e z$. Branches can be put together geometrically to define a single valued function (defined not on the complex plane but) on a Riemann Surface [15].

Remark 23. We can use the complex variable $z$ in place of $\theta$, and write $e^{jz} = \cos z + j \sin z$. We can even derive the $\cos z$ and $\sin z$ by using the exponential $z$. The trigonometric rules will, of course, be valid here, too. We can replace the real variable $x$ in a function $f(x)$ with $z$ when the function has a convergent Taylor series [15]. Similarly, the Maclaurin series for $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$ will work nicely when we replace the real number $x$ with complex number $z$.

Remark 24. One can think of many other functions for complex variable $z$. Let us consider the infinite series $\sum \frac{1}{n}$, which is known to diverge. However, if we let $s = x + jy$, a complex number with $x > 1$ (real part greater than 1), then

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \quad \text{with} \quad n^s = e^{s \log_e n} \quad \text{and} \quad \log_e n \in \mathbb{R}$$

(2.8)
converges and defines the function in the half-plane \( x > 1 \). It is called the Riemann zeta function, a famous function used in number theory.

2.2. Combinatorial Analysis

For the purposes of this study, we limit the survey of combinatorial analysis to permutations. Permutations in general is a broad subject, however, considering the scope of this study, we are only interested in permutations as a representation of specific arrangements of a set as used in the field of combinatorics. Let us first define permutation and the notation for permutations we choose to use. The style and notation here has been adapted from [17].

**Definition 7.** Given a set of \( n \) elements \( S = \{1, 2, \ldots, n\} \), a function \( \Omega \) defined from \( S \) to \( S \) is said to be a permutation of degree \( n \), if it is one to one and onto. Basically, a permutation can be denoted by listing the elements of \( S \) in a row with their images below them. The notation

\[
\begin{array}{cccc}
1 & 2 & 3 & \cdots & n \\
x_1 & x_2 & x_3 & \cdots & x_n \\
\end{array}
\]

is used to show that \( x_1 = \Omega(1), x_2 = \Omega(2), \ldots, x_n = \Omega(n) \).

**Definition 8.** The set of all permutations of order \( n \) is called \( \text{PERMUTATION}_{\Omega}(n) \).

**Remark 25.** A simpler form of permutation notation can also be written by omitting the top line:

\[
\begin{array}{cccc}
x_1 & x_2 & x_3 & \cdots & x_n \\
\end{array}
\]

2.2.1. Some Special Permutations of Interest

Because of some of their special relationships to regular polygons of graph theory, we are interested in some special permutations that correspond to the solutions of a
problem known as the N-queens problem. Because of their unexpected applications in various areas of science and engineering, different aspects of these solutions are still being published [18]. In this section, the N-queens problem is first defined, then a permutation representation of the N-queens solutions is presented.

**Definition 9.** The placing of \( n \) queens on a \( n \times n \) chessboard so that no queen attacks any other is called the *N-queens problem*.

![Fig. 7. A solution for \( n = 7 \).](image)

**Remark 26.** There can be only one queen on a diagonal, a row, or a column. Let \( Q \) be a queen on the row \( a \) and column \( b \) of a chessboard (Fig. 7.) Let us represent such a queen as \( Q(a, b) \). Two queens, \( Q_1(a_1, b_1) \) and \( Q_2(a_2, b_2) \) are said to attack each other, if at least one of the following conditions is satisfied:

- \( a_1 = a_2 \) (they are on the same row),
- \( b_1 = b_2 \) (they are on the same column),
- \( a_1 + a_2 = b_1 + b_2 \) (they are on the same diagonal),
- \( a_1 - a_2 = b_1 - b_2 \) (they are on the same diagonal).

If none of the conditions above is satisfied for all possible pairs of placed queens, the placing is said to be a solution.
2.2.1.1. *Permutation representation for the solutions.* Due to our specific interest, we will explain the permutation representation of the N-queens solutions.

**Remark 27.** The same permutation notation we used in (2.9) is used here to represent the positions of queens on a chessboard. Each pair in the permutation \((x_1,1), (x_2,2), (x_3,3), \ldots, (x_n,n)\) represents a position of a queen, like a point in Cartesian coordinates.

The permutation notation can further be simplified as seen below, where \(x_i\) represents the coordinate of column \(x_i\) and row \(i\) (here row numbers are on \(y\) and column numbers are on \(x\) axis)

\[
\begin{bmatrix}
x_1 & x_2 & \cdots & x_n \\
\end{bmatrix}
\]

**Example 7.** The permutation representation of the solution in Fig. 7 is

\[
\begin{bmatrix}
1 & 2 & 3 & 4 & 5 & 6 & 7 \\
1 & 3 & 5 & 7 & 2 & 4 & 6 \\
\end{bmatrix}
\]

or simply

\[
\begin{bmatrix}
1 & 3 & 5 & 7 & 2 & 4 & 6 \\
\end{bmatrix}
\]

**Remark 28.** For a given permutation representation, a queen solution on a chessboard can also be obtained by placing the queens on the Cartesian coordinates of \((x_i,i)\). Here \(x_i\) indicates the column number, whereas \(i\) indicates the row number.

**Definition 10.** The set of all permutations of order \(n\) that correspond to an \(n\)-queens solution is called \(\text{QUEEN}_\Omega(n)\).

2.2.2. *Permutation Matrices*

**Definition 11.** A square matrix \(P\) is called a *permutation matrix* if each row and column of \(P\) has exactly one 1 and all of its other entries are 0 [19]. A permutation matrix can also
be defined as a matrix obtained by permuting rows of an $n \times n$ identity matrix according to some permutation of the numbers 1 to $n$ [20]. Every permutation has a unique permutation matrix, hence there are $n!$ permutation matrices of size $n$.

A permutation matrix is usually used to rearrange columns (or rows) of an $n \times n$ square matrix when it is post-multiplied (or pre-multiplied) by a permutation matrix according to the given permutation vector. We will see how it is used to rearrange columns of Fourier Matrix to perform Fast Fourier Transform later in the Fourier Analysis section.

A permutation matrix is nonsingular, whose determinant is always $\pm 1$, and satisfies $PP^T = I$. Hence permutation matrices are orthogonal. In addition, the product of two permutation matrices is also a permutation matrix.

In this study, we are especially interested in a special subset of permutation matrices, the primary permutation matrices, because their eigenvalues give $n$-th roots of unity, as mentioned in Section 2.1.3. It will suffice to include a list of definitions and theorems following [19]:

**Definition 12.** An $n$-square *circulant matrix* is a matrix of the form

$$
\begin{bmatrix}
c_0 & c_1 & c_2 & \cdots & c_{n-1} \\
c_{n-1} & c_0 & c_1 & \cdots & c_{n-2} \\
c_{n-2} & c_{n-1} & c_0 & \cdots & c_{n-3} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
c_1 & c_2 & c_3 & \cdots & c_0
\end{bmatrix},
$$

where $c_0, c_1, \ldots, c_{n-1}$ are complex numbers.
**Definition 13.** A circulant matrix $P$ of the form

\[
\begin{bmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1 \\
1 & 0 & 0 & \cdots & 0
\end{bmatrix}
\]

is referred as $n \times n$ primary permutation matrix.

**Theorem 4.** An $n$-square matrix $C$ is circulant if and only if $C = PCP^T$, where $P$ is the $n \times n$ primary permutation matrix.

**Theorem 5.** Let $C$ be a circulant matrix and let $f(\lambda) = c_0 + c_1\lambda + \ldots + c_{n-1}\lambda^{n-1}$. Then

- $C = f(P)$, where $P$ is the $n \times n$ primary permutation matrix,
- $C$ is a normal matrix, where $\overline{CC} = C\overline{C}$,
- the eigenvalues of $C$ are $f(w^k)$, where $k = 0, 1, \ldots, n - 1$ and $w$ is the primitive $n$-th root of unity as defined in Section 2.1.3,
- $\det C = f(w^0)f(w^1)\ldots f(w^{n-1})$,
- $\overline{FCF}$ is a diagonal matrix where $F$ is the unitary matrix with its $(i, j)$-entry equal to $\frac{1}{\sqrt{n}}w^{(i-1)(j-1)}$, $i, j = 1, 2, \ldots, n$.

**Definition 14.** A matrix $A$ of order $n$ is said to be reducible if there exists a permutation matrix $P$ such that $P^TAP = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$, where $B$ and $D$ are square matrices of order at least 1. Also, a matrix is said to be irreducible if it is not reducible.

**Remark 29.** The eigenvalues of an $n \times n$ primary permutation matrix $P$ are exactly all the roots of equation $\lambda^n = 1$, that is, $n$-th roots of unity, since $\det(\lambda I - P) = \lambda^n - 1$ by a direct computation. Also, for any positive integer $k < n$, $P^{n-1} = P^T$, $P^n = I_n$, and $P_k = \begin{bmatrix} 0 & I_{n-k} \\ I_k & 0 \end{bmatrix}$. 

26
**Definition 15.** The matrix $P^TAP$ (or $P^{-1}AP$) is said to be permutation similar to $A$ through the permutation matrix $P$.

**Theorem 6.** Every reducible permutation matrix is permutation similar to a direct sum of primary permutation matrices. Moreover, the rank of an $n$-square irreducible permutation matrix minus $I$ is $n - 1$.

2.3. Chebyshev Analysis

In this section, the basic definition and some elementary properties of Chebyshev polynomials are presented [21, 22, 23, 24]. There are several different spelling forms for Chebyshev in the literature, including Chebychev [25], Tchebychef [26], and Tchebycheff. In this document, we will use the spelling “Chebyshev” [27]. Furthermore, the Chebyshev polynomials notation used here generally follows [27].

2.3.1. Definition of The Chebyshev Polynomials

**Definition 16.** If $\mathbb{F}$ is a field, and $n$ is a non-negative integer, then a polynomial of degree $n$ over $\mathbb{F}$ is a sum of the form

$$p_n(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n,$$  \hspace{1cm} (2.11)

where $a_i \in \mathbb{F}$ for $i = 0, \ldots, n$, $a_n \neq 0$, and $x$ is undetermined [28].

In this definition, $\mathbb{F}$ indicates the complex number field, and $x$ indicates a complex variable. In this section, however, we shall suppose that, unless otherwise stated, $\mathbb{F}$ is the real number field and $x$ is a real variable. Polynomials can be completely specified as simple as by the $n + 1$ numbers $a_0, \ldots, a_n$. This feature makes the polynomials more suitable to represent mathematical, geometric, and engineering models.
Let us consider the function

\[ T_n(x) = \cos n\theta, \quad (2.12) \]

where \( n \) is a non-negative integer, \( x = \cos \theta \), and \( 0 \leq \theta \leq \pi \).

Remark 30. Since \( x = \cos \theta \), and \( 0 \leq \theta \leq \pi \), the function \( T_n(x) \) is defined on the interval \(-1 \leq x \leq 1\). We denote this interval for \( x \) as \( I \), so that, for a given \( x \in I \), we can find the unique value of \( \theta = \arccos x \), which satisfies \( 0 \leq \theta \leq \pi \), and \( T_n(x) = \cos n\theta \). Thus \( T_n(x) \) is a single-valued function defined on \( I \), which can be written as

\[ T_n(x) = \cos n(\arccos x), \]

where \(-1 \leq x \leq 1\).

**Theorem 7.** The function in (2.12) can be written in polynomial form, such as [27]

\[ T_n(x) = t_0^{(n)} + t_1^{(n)}x + \cdots + t_n^{(n)}x^n, \quad (2.13) \]

where

\[ t_{n-2k+1}^{(n)} = 0, \quad (k = 0, \ldots, \lfloor n/2 \rfloor) \quad \text{and} \]

\[ t_{n-2k}^{(n)} = (-1)^k \sum_{j=k}^{\lfloor n/2 \rfloor} \binom{n}{j} \binom{j}{k}, \quad (k = 0, \ldots, \lfloor n/2 \rfloor). \]

**Definition 17.** For each non-negative integer \( n \), the polynomial given by (2.13) is called the **Chebyshev polynomial of degree \( n \)**.

**Example 8.** Chebyshev polynomials up to degree 10 are shown in Table 1.
TABLE 1

Chebyshev Polynomials Up to Degree 10

| \( T_0 (x) \) | 1 |
| \( T_1 (x) \) | \( x \) |
| \( T_2 (x) \) | \( 2x^2 - 1 \) |
| \( T_3 (x) \) | \( 4x^3 - 3x \) |
| \( T_4 (x) \) | \( 8x^4 - 8x^2 + 1 \) |
| \( T_5 (x) \) | \( 16x^5 - 20x^3 + 5x \) |
| \( T_6 (x) \) | \( 32x^6 - 48x^4 + 18x^2 - 1 \) |
| \( T_7 (x) \) | \( 64x^7 - 112x^5 + 56x^3 - 7x \) |
| \( T_8 (x) \) | \( 128x^8 - 256x^6 + 160x^4 - 32x^2 + 1 \) |
| \( T_9 (x) \) | \( 256x^9 - 576x^7 + 432x^5 - 120x^3 + 9x \) |
| \( T_{10} (x) \) | \( 512x^{10} - 1280x^8 + 1120x^6 - 400x^4 + 50x^2 - 1 \) |

2.3.2. Some Properties of The Chebyshev Polynomials

1. From (2.13), we see that \( T_n (x) \) is an even function for even \( n \) and an odd function for an odd \( n \)

\[
T_n (-x) = (-1)^n T_n (x).
\]

2. Again from (2.13), the nonzero coefficients of \( T_n (x) \) are integers that alternates in sign, and the leading coefficient \( t_n (n) \) can be written in a simpler form:

\[
t_n (n) = \sum_{j=0}^{[n/2]} \binom{n}{2j} = \frac{1}{2} \left\{ (1 + 1)^n + (1 - 1)^n \right\} = 2^{n-1}.
\]

3. Another property of Chebyshev polynomials is the classical recurrence equation that these polynomials satisfy

\[
T_{n+1} (x) = 2x T_n (x) - T_{n-1} (x).
\]
4. Chebyshev polynomials also have orthogonality property. The following equation is trivial for all non-negative \( m \) and \( k \)

\[
\int_0^\pi \cos k\theta \cos m\theta d\theta = 0, \quad k \neq m,
\]

\[
\int_0^\pi \cos^2 k\theta d\theta = \begin{cases} \frac{\pi}{2}, & k \neq 0, \\ \pi, & k = 0. \end{cases}
\]

(2.14)

if we change the variables according to \( x = \cos \theta \), (2.14) becomes

\[
\int_{-1}^1 T_k(x) T_m(x) \frac{dx}{\sqrt{1-x^2}} = 0, \quad k \neq m,
\]

\[
\int_{-1}^1 T_k^2(x) \frac{dx}{\sqrt{1-x^2}} = \begin{cases} \frac{\pi}{2}, & k \neq 0, \\ \pi, & k = 0; \end{cases}
\]

(2.15)

therefore, we can say that the Chebyshev polynomials \( \{ T_n(x) \}_{n=0}^\infty \) form a sequence of orthogonal polynomials on \(-1 \leq x \leq 1\) with respect to the weight function \((1 - x^2)^{-1/2}\).

### 2.3.3. Chebyshev Polynomial of Second Kind

Chebyshev polynomials of second kind have the same properties as Chebyshev polynomials [29]. Actually, the second kind of Chebyshev polynomials is what we will use to establish their relationships with the rest of the tools in the framework.

**Remark 31.** Let us find the differentiation of \( T_n(x) = \cos n\theta \) with respect to \( x \)

\[
T_n'(x) = \left( \frac{d}{d\theta} \cos n\theta \right) \frac{d\theta}{dx} = (-n \sin n\theta) \frac{d}{dx} \arccos x = (-n \sin n\theta) \left( -\frac{1}{\sqrt{1-x^2}} \right)
\]

\[
= (-n \sin n\theta) \left( -\frac{1}{\sqrt{1-\cos^2 \theta}} \right) = \frac{n \sin n\theta}{\sqrt{\sin^2 \theta}}
\]

\[
T_n'(x) = \frac{n \sin n\theta}{\sin \theta}.
\]

(2.16)
Definition 18. The polynomial of degree $p - 1$

$$u_{p-1}(x) = \frac{1}{p} T_p'(x) = \frac{\sin p\theta}{\sin \theta}, \quad (x = \cos \theta), \quad (2.17)$$

is called the Chebyshev polynomial of second kind.

Remark 32. By replacing $p$ with $p + 1$ in (2.17), we obtain the general form of Chebyshev polynomial of second kind

$$u_p(x) = \frac{\sin (p + 1) \theta}{\sin \theta}. \quad (2.18)$$

Example 9. Chebyshev polynomials of second kind up to degree 10 are shown in Table 2.

<table>
<thead>
<tr>
<th>Chebyshev Polynomials of Second Kind Up to Degree 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u_0(x)$</td>
</tr>
<tr>
<td>$u_1(x)$</td>
</tr>
<tr>
<td>$u_2(x)$</td>
</tr>
<tr>
<td>$u_3(x)$</td>
</tr>
<tr>
<td>$u_4(x)$</td>
</tr>
<tr>
<td>$u_5(x)$</td>
</tr>
<tr>
<td>$u_6(x)$</td>
</tr>
<tr>
<td>$u_7(x)$</td>
</tr>
<tr>
<td>$u_8(x)$</td>
</tr>
<tr>
<td>$u_9(x)$</td>
</tr>
<tr>
<td>$u_{10}(x)$</td>
</tr>
</tbody>
</table>

Just like Chebyshev polynomials, Chebyshev polynomials of second kind too have odd-even, orthogonality, and recurrence properties, where the recurrence equation is given by:

$$u_p(x) = 2xu_{p-1}(x) - u_{p-2}(x). \quad (2.19)$$
2.4. Fourier Analysis

Fourier analysis contains mathematical tools that are used in many fields of science and engineering. It has been approached from different angles in the literature depending on which discipline a particular publication belongs to. We too are going to introduce this subject from our perspective using a notation to connect this set of tools to the noisy communication channel framework. We will first touch on Fourier Series and its generalized and complex forms and cosine and sine series. That is followed by the Fourier Integrals and their complex form. Then, we examine Fourier Transform followed by a summary of Discrete Fourier Analysis consisting of Discrete Fourier Transform, Fast Fourier Transform, and Fourier Matrices. We will also use some of the concepts introduced in Complex Analysis such as periodic functions, complex exponentials, and some concepts of trigonometry.

2.4.1. Fourier Series

From an engineering point of view, Fourier Series is simply expressing a periodic function (as defined in Definition 1) in terms of an infinite number of sines and cosines.

Definition 19. The simplest periodic function is called a harmonic and denoted as

\[ y(x) = A \sin(wx + \phi) \]  

(2.20)

where \( A \), \( w \), and \( \phi \) are constants. This function is called a harmonic of amplitude \( |A| \), angular frequency \( w \), and initial phase \( \phi \) [30].

The period of this harmonic would be \( T = \frac{2\pi}{w} \), since for any \( x \)

\[
A \sin \left[ w \left( x + \frac{2\pi}{w} \right) + \phi \right] = A \sin \left[ (wx + \phi) + 2\pi \right] \\
= A \sin (wx + \phi).
\]
Note that we could have also chosen to use cosine instead of sine in (2.20) and the period would have been the same. Using the well-known trigonometric formula below, we can rewrite the above harmonic as

$$A \sin(wx + \phi) = A \left( \cos wx \sin \phi + \sin wx \cos \phi \right).$$

By setting $a = A \sin \phi$, and $b = A \cos \phi$, we can write a harmonic in the following form:

$$y(x) = a \cos wx + b \sin wx,$$

where $A = \sqrt{a^2 + b^2}$, $\sin \phi = a/A$, and $\cos \phi = b/A$. Also, if we set the period to $T = 2l$, and since we know $T = 2\pi/w$, then we get $w = 2\pi/T = \pi/l$. Hence the harmonic with period $T = 2l$ can be written as

$$a \cos \frac{\pi x}{l} + b \sin \frac{\pi x}{l}.$$

Let us now consider the following harmonics with frequencies $w_k = \pi k/l$ and periods $T_k = 2\pi/w_k = 2l/k$:

$$a_k \cos \frac{\pi k x}{l} + b_k \sin \frac{\pi k x}{l} \quad (k = 1, 2, \ldots).$$

Note that the period $T = 2l$ for the first harmonic is also a period for all other harmonics (for this reason, we denoted this period as $T$, instead of $T_1$.)

**Definition 20.** A sum $s_n(x)$ of the following form is called a *trigonometric polynomial of order $n$ and period $2l*$:

$$s_n(x) = A + \sum_{k=1}^{n} \left( a_k \cos \frac{\pi k x}{l} + b_k \sin \frac{\pi k x}{l} \right).$$
The period of the resulting function is also $T = 2l$, since all of the sine and cosine components share the same period. By assigning different values to $A$, $a_k$, $b_k$ coefficients, one would end up with variety of functions with period $T = 2l$ that are much more complex than simple harmonics. In the event that the order of this polynomial is taken to infinity, the types of functions that can be produced as a result of this sum (assuming it converges) is even more diverse. Consequently, one wonders if any given function with period $T = 2l$ can be represented as a sum of such trigonometric series. Let us assume that $f(x)$ is a function that can be represented as $f(x) = A + \sum_{k=1}^{\infty} (a_k \cos \frac{\pi k x}{l} + b_k \sin \frac{\pi k x}{l})$.

Now let us apply $\frac{\pi x}{l} = t$ or $x = t l / \pi$ variable change to get a simpler form of the infinite trigonometric series:

$$g(t) = f\left(\frac{tl}{\pi}\right) = A + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt).$$  \hspace{1cm} (2.21)

The function $g(t)$ has a period of $2\pi$. Also, we know how to convert $g(t)$ and $f(x)$ to each other. So, it makes sense to solve the expansion for functions of the standard period $2\pi$, since we can always convert the solution back to the generalized form with period $2l$.

Let us assume that the function $f(x)$ with period $2\pi$ has the following expansion:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)$$  \hspace{1cm} (2.22)

The first constant is denoted as $a_0/2$ to simplify the subsequent formulas. Given the function itself, in order to define this expansion, the $a_0$, $a_k$, and $b_k$ coefficients need to be determined for $k = 1, 2, \ldots$. Both sides of the equation are a) integrated from $-\pi$ to $\pi$ to determine $a_0$; b) multiplied by $\cos nx$, then integrated from $-\pi$ to $\pi$ to determine $a_k$; c) multiplied by $\sin nx$, then integrated from $-\pi$ to $\pi$ to determine $b_k$. Here we assume that (2.22) and the series obtained by $\cos nx$ and $\sin nx$ multiplications can be integrated term by term; in other words, for all these series, the integral of the sum equals the sum of the integral.
Considering the scope of this study, we will not talk about the details of integrability here; it is explained in detail in [30].

Solving these three integrals becomes easier because of the integral identities below, which are due to the fact that sine and cosine make up an orthogonal system over the interval $[-\pi, \pi]$ for $n, m \neq 0$:

\[
\begin{align*}
\int_{-\pi}^{\pi} \cos n x \, dx &= 0 \\
\int_{-\pi}^{\pi} \sin n x \, dx &= 0 \\
\int_{-\pi}^{\pi} \cos nx \cos mx \, dx &= \pi \delta_{nm} \\
\int_{-\pi}^{\pi} \sin nx \sin mx \, dx &= \pi \delta_{nm} \\
\int_{-\pi}^{\pi} \sin nx \cos mx \, dx &= 0
\end{align*}
\]  

(2.23)

where $\delta_{nm}$ is Kronecker Delta defined as

\[
\delta_{nm} \equiv \begin{cases} 
0, & n \neq m \\
1, & n = m
\end{cases}.
\]

As a result, the coefficients are obtained as:

\[
\begin{align*}
a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx \quad (n = 0, 1, 2, \ldots) \\
b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx \quad (n = 1, 2, \ldots)
\end{align*}
\]  

(2.24)

We note that since the integral of a function can be defined as the area between the function and x-axis, the above-mentioned integrals for periodic functions would be the same for any interval of length $2\pi$. In other words, we can replace the $[-\pi, \pi]$ interval with $[a, a + 2\pi]$ for any real number $a$ and still obtain same results.

**Definition 21.** The coefficients defined by (2.24) are called Fourier coefficients, and the trigonometric series defined by (2.22) is called Fourier Series of function $f(x)$. 
To generalize this idea to periodic functions with an arbitrary period $2l$, let us remember $\frac{\pi x}{l} = t$ (or $x = \frac{lt}{\pi}$) variable change that we used before to obtain $g(t) = f\left(\frac{lt}{\pi}\right)$ with period $2\pi$, as described in (2.21). Now let us write the Fourier Series and coefficients $g(t)$ of period $2\pi$ using (2.22) and (2.24):

$$g(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nt + b_n \sin nt)$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} g(t) \cos nt \, dt = \frac{1}{\pi} \int_{-\pi}^{\pi} f\left(\frac{lt}{\pi}\right) \cos nt \, dt \quad (n = 0, 1, \ldots)$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} g(t) \sin nt \, dt = \frac{1}{\pi} \int_{-\pi}^{\pi} f\left(\frac{lt}{\pi}\right) \sin nt \, dt \quad (n = 1, 2, \ldots)$$

By resetting $t$ to $\frac{\pi x}{l}$, we obtain Fourier Series and coefficients of a function of period $2l$ as follows:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos \pi nx/l + b_n \sin \pi nx/l)$$

$$a_n = \frac{1}{l} \int_{-l}^{l} f(x) \cos \pi nx/l \, dx \quad (n = 0, 1, \ldots)$$

$$b_n = \frac{1}{l} \int_{-l}^{l} f(x) \sin \pi nx/l \, dx \quad (n = 1, 2, \ldots)$$

(2.25)

2.4.1.1. Cosine and sine series. It is also worthwhile to look at the Fourier Series of even and odd functions, as defined in Definitions 2 and 3, respectively. Given an even periodic function $f(x)$, $f(x) \cos nx$ is an even function as well, since $\cos nx$ is also an even function, and multiplication of two even functions is an even function too. Similarly, $f(x) \sin nx$ is an odd function, since $\sin nx$ is an odd function, and multiplication of an even function and an odd function is an odd function. Considering its area under the curve definition, integration of an odd periodic function between $[-l, l]$ would be zero and integration of an even periodic function in the $[-l, l]$ would be twice its value between $[0, l]$
for any given \( l \). This means that \( b_n \) will be zero for all \( n \) values. Consequently, Fourier Series collapses into \textit{Cosine Series} of an even function of period \( 2l \) and is given by:

\[
    f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos \frac{\pi nx}{l} \\
    a_n = \frac{2}{l} \int_0^{l} f(x) \cos \frac{\pi nx}{l} dx \quad (n = 0, 1, 2, \ldots)
\]

Applying the same logic to the odd functions, \( a_n \) will be zero for all \( n \) values and Fourier Series will collapse into \textit{Sine Series} of an odd function of period \( 2l \) and can be denoted as:

\[
    f(x) = \sum_{n=1}^{\infty} b_n \sin \frac{\pi nx}{l} \\
    b_n = \frac{2}{l} \int_0^{l} f(x) \sin \frac{\pi nx}{l} dx \quad (n = 1, 2, \ldots)
\]

Let us assume that we can write an arbitrary function in terms of its even and odd components: \( f(x) = f_e(x) + f_o(x) \). In order to verify this assumption, the even and odd components should be constructed in a way so that in addition to the above equation they also should satisfy \( f_e(x) = f_e(-x) \) and \( f_o(x) = -f_o(-x) \). It turns out that such even and odd components can be written in terms of \( f(x) \) as expressed in the below formula:

\[
    f(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2}.
\]

Consequently, series expansion of a function can also be obtained by summation of its sine series of even components and cosine series of odd components. This must be done with caution, making sure attention is paid to the points of jump discontinuities [30].

\subsection*{2.4.1.2. Generalized Fourier series.}
Before we move onto the complex form of the Fourier Series, let us briefly mention the generalized Fourier Series. Fourier Series is possible thanks to the complete bi-orthogonal system of \( \cos nx \) and \( \sin nx \) under the integration over the range \([-\pi, \pi]\). Therefore, it is logical to think that the same would also be possible
for other complete orthogonal systems, such as Laplace Series, which is based on spherical harmonics. Generalized Fourier Series and coefficients are defined as

\[ f(x) = \sum_{n=0}^{\infty} a_n \phi_n(x) \]

\[ a_n = \frac{1}{c_n} \int_{R} f(x) \phi_n(x) w(x) \, dx \]

where \( \{ \phi_n(x) \} \) are uni-variate functions of the orthogonal system, \( w(x) \) is the weighting function that normalizes the orthogonal functions \( \int [\phi_n(x)]^2 w(x) \, dx = N_n \) \( R \) is the interval over which the system is orthogonal, and \( c_m \) are given constants that satisfy the orthogonality relationship \( \int_{R} \phi_m(x) \phi_n(x) w(x) \, dx = c_m \delta_{mn} \).

For a complete bi-orthogonal system of functions \( f_1(n,x) \) and \( f_2(n,x) \), generalized Fourier Series and coefficients are given by

\[ f(x) = \sum_{n=0}^{\infty} a_n f_1(n,x) + \sum_{n=0}^{\infty} b_n f_2(n,x) \]

\[ e = \frac{\int_{R} f(x) w(x) \, dx}{\int_{R} f(x) \, dx} \]

\[ a_n = \frac{1}{c_n} \int_{R} f(x) f_1(n,x) w(x) \, dx \]

\[ b_n = \frac{1}{d_n} \int_{R} f(x) f_2(n,x) w(x) \, dx. \]

Here \( c_n \) and \( d_n \) are obtained from the following orthogonality relationships:

\[ \int_{R} f_1(n,x) f_1(m,x) w(x) \, dx = c_n \delta_{nm} \]

\[ \int_{R} f_2(n,x) f_2(m,x) w(x) \, dx = d_n \delta_{nm} \]

\[ \int_{R} f_1(n,x) f_2(m,x) w(x) \, dx = 0 \]

\[ \int_{R} f_1(n,x) w(x) \, dx = 0 \]

\[ \int_{R} f_2(n,x) w(x) \, dx = 0. \]
2.4.1.3. Complex form of Fourier series. Lastly, we will look at the complex form of the Fourier Series. We start with applying Euler’s formula $e^{i\theta} = \cos \theta + i \sin \theta$ (consequently, $\cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2}$ and $\sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i} = i \frac{e^{i\theta} - e^{-i\theta}}{2}$) relating the trigonometric and complex functions to (2.22) and (2.24):

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left( a_n \cos nx + b_n \sin nx \right)$$

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left( a_n \frac{e^{inx} + e^{-inx}}{2} + ib_n \frac{e^{inx} - e^{-inx}}{2} \right)$$

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left( \frac{a_n - ib_n}{2} e^{inx} + \frac{a_n + ib_n}{2} e^{-inx} \right).$$

Let us set $c_0 = a_0/2$, $c_n = (a_n - ib_n)/2$, and $c_{-n} = (a_n + ib_n)/2$. Then we can rewrite $f(x)$ as $f(x) = c_0 + \sum_{n=1}^{\infty} (c_n e^{inx} + c_{-n} e^{-inx})$. Note that for real valued $f(x)$, $c_n$ and $c_{-n}$ are complex conjugate. The complex form of the Fourier Series of a periodic function $f(x)$ is given as

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx} \quad (2.27)$$

In order to determine the complex Fourier coefficients, let us multiply (2.27) with $e^{-inx}$ and integrate it over $[-\pi, \pi]$:

$$\int_{-\pi}^{\pi} f(x) e^{-inx} dx = \sum_{n=-\infty}^{\infty} c_n \int_{-\pi}^{\pi} e^{i(n-m)x} dx.$$ 

Because of sine cosine orthogonality, from (2.23), we know that for $n \neq m$

$$\int_{-\pi}^{\pi} (\cos (n-m)x + i \sin (n-m)x) dx = 0.$$
So, the only integral term left in the right hand side of the above equation will be \( \cos(n - m)x \) for \( n = m \) (which is \( 2\pi \)), since \( \sin(n - m)x \) will also be equal to zero. We can then express the Fourier coefficients in the complex form as

\[
c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} \, dx
\] (2.28)

Similarly, the complex form of Fourier Series and coefficients of a function of period \( 2l \) is given by

\[
f(x) = \sum_{n=-\infty}^{\infty} c_n e^{in\pi x/l}, \quad c_n = \frac{1}{2l} \int_{-l}^{l} f(x) e^{-inx/l} \, dx
\]

respectively.

2.4.2. Fourier Integrals

In the previous section, we deal with periodic functions. Here we will think of a function \( f(x) \) that is defined only on a finite interval \([-l, l]\) whose length is same as the period of the function we talk about in Section 2.4.1. Here \( f(x) \) is defined as a real valued piecewise smooth function.

**Definition 22.** A function is said to be piecewise smooth on the interval \([a, b]\), if either the function and its derivative are both continuous on \([a, b]\) or they have only a finite number of jump discontinuities on \([a, b]\) [30].

At the points of discontinuity, \( f(x) \) should be replaced by \( 1/2(f(x + \varepsilon) + f(x - \varepsilon)) \) as \( \varepsilon \to 0 \). Let us write the Fourier Series for a piecewise smooth function \( f(x) \) from (2.25):

\[
f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos \pi nx/l + b_n \sin \pi nx/l)
\]

\[
a_n = \frac{1}{l} \int_{-l}^{l} f(u) \cos \pi nu/l \, du \quad (n = 0, 1, 2, \ldots)
\]

\[
b_n = \frac{1}{l} \int_{-l}^{l} f(u) \sin \pi nu/l \, du \quad (n = 1, 2, \ldots)
\]
Let us insert $a_n$ and $b_n$ into the series equation and rearrange the equation to get sin and cos parts together:

$$f(x) = \frac{1}{2l} \int_{-l}^{l} f(u) \, du + \sum_{n=1}^{\infty} \frac{1}{n} \int_{-l}^{l} f(u) \left[ \cos \frac{\pi n u}{l} \cos \frac{\pi n x}{l} + \sin \frac{\pi n u}{l} \sin \frac{\pi n x}{l} \right] \, du$$

$$f(x) = \frac{1}{2l} \int_{-l}^{l} f(u) \, du + \sum_{n=1}^{\infty} \frac{1}{n} \int_{-l}^{l} f(u) \cos \left( \frac{\pi n}{l} (u - x) \right) \, du. \quad (2.29)$$

Let us assume that $f(x)$ is absolutely integrable on the whole $x$ axis. Verification of this assumption can be found in [30]. As $l \to \infty$ ($x$ is fixed), the first term in (2.29) goes to 0. Let us set

$$\lambda_1 = \pi/l, \lambda_2 = 2\pi/l, \ldots, \lambda_n = n\pi/l, \ldots,$$

$$\Delta \lambda_n = \lambda_{n+1} - \lambda_n = \pi/l,$$

then let us rewrite (2.29) by replacing $1/l$ with $\Delta \lambda_n / \pi$ and $\pi n/l$ with $\lambda_n$:

$$f(x) = \lim_{l \to \infty} \frac{1}{\pi} \sum_{n=1}^{\infty} \Delta \lambda_n \int_{-l}^{l} f(u) \cos \left( \lambda_n (u - x) \right) \, du. \quad (2.30)$$

**Definition 23.** An integral is an improper integral if either the interval of integration is not finite (type I) or if the function to integrate is not continuous (not bounded) in the interval of the integration (type II.) For an improper integral $\int_{-\infty}^{\infty} f(x) \, dx$ defined in $[0, \infty]$, if the integral exists for all $t \geq a$, then we define $\int_{-\infty}^{\infty} f(x) \, dx = \lim_{t \to \infty} \int_{a}^{t} f(x) \, dx$, provided that the limit exists as a finite number [31].
Let us set \( h(\lambda) = \int_{-l}^{l} f(u) \cos(\lambda (u-x)) \, du \) in the \([0, \infty]\) interval. Using integral definition, \( \int h(\lambda) \, d\lambda = \lim_{n \to \infty} \sum_{i=1}^{n} h(\lambda_i) \Delta \lambda_i \). After we plug in \( h(\lambda_i) \) and apply the limit as \( n \to \infty \), we get

\[
\int h(\lambda) \, d\lambda = \sum_{n=1}^{\infty} \Delta \lambda_n \int_{-l}^{l} f(u) \cos(\lambda_n (u-x)) \, du.
\]

(2.31)

**Definition 24.** As \( l \to \infty \) in (2.30), it goes into the below improper double integral [30]

\[
f(x) = \frac{1}{\pi} \int_{0}^{\infty} \left[ \int_{-\infty}^{\infty} f(u) \cos(\lambda (u-x)) \, du \right] d\lambda
\]

(2.32)

This entire formula is called the Fourier Integral Theorem, and the inner integral is called the Fourier Integral.

2.4.2.1. Cosine and sine integrals. Let us rewrite (2.32) by utilizing a well-known trigonometric equity \( \cos(a-b) = \cos(b-a) = \cos a \cos b + \sin a \sin b \):

\[
f(x) = \frac{1}{\pi} \int_{0}^{\infty} \left[ \int_{-\infty}^{\infty} f(u) \cos \lambda u \cos \lambda x \, du \right. \\
+ \left. \int_{-\infty}^{\infty} f(u) \sin \lambda u \sin \lambda x \, du \right] d\lambda
\]

\[
= \int_{0}^{\infty} \left[ \frac{1}{\pi} \cos \lambda x \int_{-\infty}^{\infty} f(u) \cos \lambda u \, du \\
+ \frac{1}{\pi} \sin \lambda x \int_{-\infty}^{\infty} f(u) \sin \lambda u \, du \right] d\lambda.
\]

We can now write it in a form similar to (2.25):

\[
f(x) = \int_{0}^{\infty} \left[ a(\lambda) \cos \lambda x + b(\lambda) \sin \lambda x \right] d\lambda
\]

(2.33)

\[
a(\lambda) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(u) \cos \lambda u \, du
\]

\[
b(\lambda) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(u) \sin \lambda u \, du.
\]
We notice the resemblance between (2.25) and (2.33): The sum has been replaced by an integral, and the integral parameter \(\lambda\) is continuous. Just like we did in Section 2.4.1.1, we can express the function \(f(x)\) in terms of its odd and even components. Then, cosine and sine integrals can be obtained as follows, respectively:

\[
\begin{align*}
    f(x) &= \frac{2}{\pi} \int_0^\infty \cos \lambda x \left( \int_{-\infty}^\infty f(u) \cos \lambda u du \right) d\lambda, \\
    f(x) &= \frac{2}{\pi} \int_0^\infty \sin \lambda x \left( \int_{-\infty}^\infty f(u) \sin \lambda u du \right) d\lambda.
\end{align*}
\] (2.34)

2.4.2.2. Complex form of Fourier integrals. Let us consider the integral \(O(\lambda) = \int_{-\infty}^{\infty} f(u) \sin (\lambda (u-x)) du\) defined as a continuous function of \(\lambda\). Since this function depends on only sine of its parameter \(\lambda\), it is an odd function. Consequently, its integral on \(-\infty, \infty\) integral would be equal to 0:

\[
\int_{-\infty}^{\infty} O(\lambda) d\lambda = 0.
\] (2.35)

On the other hand, the integral \(E(\lambda) = \int_{-\infty}^{\infty} f(u) \cos (\lambda (u-x)) du\) defined as Fourier Integral in (2.32) is a continuous function of \(\lambda\) and depends on only cosine of its parameter \(\lambda\), hence it is an even function. Consequently, its integral on \([-\infty, \infty]\) can be given as:

\[
\int_{-\infty}^{\infty} E(\lambda) d\lambda = 2 \int_{0}^{\infty} E(\lambda) d\lambda.
\] (2.36)
Now let us rewrite (2.32) as \( f(x) = \frac{1}{\pi} \int_0^\infty E(\lambda) d\lambda \). By changing the integral limits according to (2.36) and adding \( \frac{i}{2\pi} \int \frac{O(\lambda)}{d\lambda} = 0 \), we arrive at:

\[
\begin{align*}
  f(x) & = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} f(u) \cos(\lambda (u-x)) \, du \right] d\lambda \\
  & + \frac{i}{2\pi} \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} f(u) \sin(\lambda (u-x)) \, du \right] d\lambda \\
  f(x) & = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} f(u) (\cos(\lambda (u-x)) \\
  & + i \sin(\lambda (u-x))) \, du \right] d\lambda.
\end{align*}
\]

Definition 25. Let us write the function inside the inner integral in exponential form:

\[
f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} f(u) e^{i\lambda (u-x)} \, du \right] d\lambda.
\]

This entire formula is called complex form of the Fourier Integral Theorem, and the inner integral is called complex form of the Fourier Integral.

2.4.3. Fourier Transforms

Let us try to write the complex form of the Fourier Integral Theorem in a way that the inner and outer integrals have a similar form, by taking \( x \) exponent outside of the inner integral:

\[
f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} f(u) e^{i\lambda u} e^{-i\lambda x} \, du \right] d\lambda.
\]

Let \( F(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u) e^{i\lambda u} \, du \). Then, we get \( f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\lambda) e^{-i\lambda x} d\lambda \).
Definition 26. Given the function \( f(x) \), the new function

\[
F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u) e^{iux} du
\]  
(2.38)

is called the Fourier Transform of \( f(x) \) and denoted as \( \mathcal{F}_u [f(u)](x) \).

Definition 27. If the Fourier transform of a function is given, then the Inverse Fourier Transform of \( F(x) \) is denoted as \( \mathcal{F}_u^{-1} [F(u)](x) \) and given by

\[
f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(u) e^{-iux} du. \]  
(2.39)

Fourier transform is a tool that is widely used in science and engineering to simplify complicated mathematical operations. Among many, two very important properties of Fourier transform are

\[
\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u) e^{iux} du = \frac{x}{i} F(x)
\]
\[
\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left( \frac{d}{dt} \int_{0}^{t} f(t') dt' \right) e^{iux} du = \frac{i}{x} F(x). \]  
(2.40)

These properties show that differentiating the original function \( f(x) \) corresponds to multiplying its Fourier transform \( F(x) \) by \( x/i \) and integrating \( f(x) \) corresponds to dividing \( F(x) \) by \( x/i \). This idea of reducing complicated mathematical operations on the original function to simple algebraic operations on its transform and then taking the inverse transform of the final result is the basis for almost all applications of Fourier transform in science and engineering.

2.4.3.1. Cosine and sine transforms. Given an even function \( f(x) \), the sine part of (2.33) would vanish and be reduced to

\[
F(\lambda) = \sqrt{\frac{2}{\pi}} \int_{0}^{\infty} f(u) \cos \lambda u du.
\]
This transform is called \textit{(Fourier) cosine transform} of the function $f(x)$. From (2.34), we get

$$f(x) = \sqrt{\frac{2}{\pi}} \int_{0}^{\infty} F(\lambda) \cos x \lambda \, d\lambda.$$ 

What we find is same as the Fourier cosine transform. In other words, the functions $f(x)$ and $F(x)$ are cosine transforms of each other.

Similarly, for an even function, we can write \textit{(Fourier) sine transform} as

$$\Phi(\lambda) = \sqrt{\frac{2}{\pi}} \int_{0}^{\infty} f(u) \sin \lambda u \, du.$$ 

From (2.34), we get

$$f(x) = \sqrt{\frac{2}{\pi}} \int_{0}^{\infty} \Phi(\lambda) \sin x \lambda \, d\lambda.$$ 

Just like in the case of cosine transform, $f(x)$ and $\Phi(x)$ are sine transforms of each other.

2.4.4. \textit{Discrete Fourier Transform}

Up to this point, we talked about the Fourier Analysis that deals with continuous functions defined at all points of an interval. However, in real-world applications, we either deal with arrays of data samples collected or discrete samples that represent a function on digital computers. In this section, we will briefly go over the methods that approximate the mentioned continuous functions.

It is worth to note that although this topic appears in the literature as Discrete Fourier Transform, it can be argued that it may be more appropriate to call it more generally as Discrete Fourier Analysis, because the method we visit here can be used to calculate not only the Fourier Transform integral, but also the integral that defines the Fourier Series coefficients [32].
Considering the scope of this work, in this section we will not get into the derivations of the equations. We will only give definitions.

Let us define a discrete function \( f(x_k) \equiv f_k \) as \( N \) samples of the function \( f(x) \) on \([0, x_{N-1}]\), where \( x_k \equiv k\Delta x \) with \( k = 0, 1, \ldots, N - 1 \).

**Definition 28.** Given the function \( f_k \), the new function

\[
F_n = \sum_{k=0}^{N-1} f_k e^{-2\pi i nk/N}
\]  

is called the *Discrete Fourier Transform (DFT)* of \( f_k \) and denoted as \( \mathcal{F}_k \{ f_k \}_{k=0}^{N-1} \}(n).

**Definition 29.** If the discrete Fourier transform of a function is given, then the *Inverse Discrete Fourier Transform (IDFT)* of \( F_k \) is denoted as \( \mathcal{F}_k^{-1} \{ F_k \}_{k=0}^{N-1} \}(n) \) and given by

\[
f_n = \frac{1}{N} \sum_{k=0}^{N-1} F_k e^{2\pi i nk/N}.
\]  

It is worth to note that DFT of a real sequence of numbers will be a sequence of complex numbers, and there exists the below relation, if \( f_k \) are real:

\[
F_{N-n} = \overline{F_n},
\]

where \( \overline{z} \) denotes complex conjugate of a complex number \( z \) and \( F_0 \) is real. As a result, a periodic function will have two peaks in the transformed domain, because the periods of the input data split into positive and negative frequency complex components.

There also exist forward and inverse discrete (Fourier) sine and cosine transforms that can be defined using the methods used in Section 2.4.3.1.

### 2.4.4.1. Fast Fourier transform.

It would take \( 2N^2 \) computations to obtain DFT as it was described above. Over the years, there have been many techniques introduced to
perform Discrete Fourier Transform faster. In the literature these techniques are referred as Fast Fourier Transforms (FFT). Although FFT was first discussed by Cooley and Tukey in 1965 [34], the critical factorization step was described by Gauss in 1805 [35]. Most of the FFT techniques are based on the Danielson-Lanczos Lemma [36].

Considering the scope of this work, in this section we will not get into the derivations of the equations. It will suffice that we give descriptions of Danielson-Lanczos Lemma and how it is utilized in the Cooley and Tukey algorithm and some of the relevant properties of FFT.

The Danielson-Lanczos Lemma suggests that a DFT of length $N$ can be rewritten as the sum of two DFTs, each of length $N/2$, where one of the DFT is formed from the even-numbered points of the original sequence and the other is from the odd-numbered points of it:

$$F_n = \sum_{k=0}^{N-1} f_k e^{-2\pi i nk/N} = \sum_{k=0}^{N/2-1} f_{2k} e^{-2\pi i (k/2) n} + W^n \sum_{k=0}^{N/2-1} f_{2k+1} e^{-2\pi i (k/2) n},$$

where $W \equiv e^{-2\pi i/N}$ and $n = 0, 1, 2, \ldots, N$. Let us denote the left part of the addition above as $F_n^e$ and the right part of it as $F_n^o$. Hence, we have $F_n = F_n^e + W^n F_n^o$. This procedure is also referred as the splitting method.

If $N/2$ is also even, then we can split $F_n^e$ and $F_n^o$ to $F_n^{ee} + W^n F_n^{eo}$ and $F_n^{oe} + W^n F_n^{oo}$ respectively and for $N = 2^m$, where $m$ is an integer, this process can be repeated for $m = \log_2 N$ times until the original transform is broken into $\log N$ transforms of length 1 (since DFT of a sequence of length 1 is itself, there is now no DFTs left to be done.) As a result of this recursive splitting process, for each $k$, we obtain a sequence of length $\log N$ consisting of es and os, which represents a pattern of even and odd splits. Let us denote that pattern as $(eeo \ldots)_k$, then we can write $F_n^{(eeo \ldots)_k} = f_k$. What is left to do now is to find the $k$ indice that corresponds to each $n$. Cooley and Tukey FFT algorithm suggests that the value of $k$ in its binary form can be obtained by reversing the order of $(eeo \ldots)_k$ pattern, i.e. $(\ldots oee)_k$, then
replacing each \( e \) with 0 and each \( o \) with 1 (note that for \( n = \frac{N}{2} \), \( W^n \) coefficient in front of odd part of the transform \( F''_n \) is equal to 1, \( W^{N/2} = e^{-\pi i} = -1 \).

The process takes \( 2N \log N \) computations, instead of \( 2N^2 \) computations required to obtain DFT using its definition, hence this process is called \textit{Fast Fourier Transform} (FFT.)

\[ \text{2.4.4.2. Fourier matrix.} \] Considering the DFT definition in (2.41), we can express this transform in matrix form by letting sampling sequences \( f_k \) and \( F_n \) be represented by two vectors of length \( N \), and the transform operation \( \mathcal{F}_k \left\{ \{ f_k \}_{k=0}^{N-1} \right\} (n) \) be represented by an \( N \times N \) square matrix whose entries are given by \( e^{-2\pi i nk/N} \).

\[
\begin{bmatrix}
F_0 \\
F_1 \\
F_2 \\
\vdots \\
F_{N-1}
\end{bmatrix}
=
\frac{1}{\sqrt{N}}
\begin{bmatrix}
1 & 1 & 1 & \cdots & 1 \\
1 & W & W^2 & \cdots & W^{N-1} \\
1 & W^2 & W^4 & \cdots & W^{2(N-1)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & W^{N-1} & W^{2(N-1)} & \cdots & W^{(N-1)^2}
\end{bmatrix}
\begin{bmatrix}
f_0 \\
f_1 \\
f_2 \\
\vdots \\
f_{N-1}
\end{bmatrix}
\tag{2.43}
\] This square matrix is called \textit{Fourier Matrix}, and we will denote it as \( \mathcal{F}_N \). There are many properties of this matrix, but for our purposes we will only talk about a few of them. It is immediately apparent that the matrix is symmetric around its diagonal. Furthermore, by normalizing it with \( 1/\sqrt{N} \), Fourier matrix becomes a unitary matrix. (A complex square matrix is called \textit{unitary}, if its conjugate transpose is equal to its inverse.) Because of this property, we could have used \( \overline{W}^{nk} \) instead of \( W^{nk} \) to construct the matrix, hence the same transform matrix can also be used for Inverse Fourier Transform.

Most significantly for our work, all of the entries in Fourier matrix are \( N \)-th roots of unity, as defined in Definition 5. As mentioned above, the choice between \( W \) and \( \overline{W} \) is irrelevant for the discussion of Fourier Matrix. Hence the notation in Remark 15, where the primitive \( n \)-th root of unity \( w_1 \) (later denoted as \( w \)) for any integer \( n \) was given as \( e^{j2\pi/n} \), can be applied here directly.

Consequently, all the entries of the Fourier Matrix are \( N \)-th roots of unity, and the powers of \( W^{nk} \) are periodic modulo \( N \). Furthermore, because \( W^n = W^{n-N} \mod N \) (to see this equity, one should consider angle directions on the complex plane and modulo \( N \),) for
each row of the top half of the Fourier Matrix, except the top row and $n = N/2$-th row, there exists its conjugate row at the bottom half.

The above mentioned and some more properties of Fourier Matrix help us to understand FFT techniques visually. As an example, we will include factorization of the Fourier Matrix, so that the number of calculations necessary for matrix multiplication in (2.43) can be reduced to a minimum. We will suffice with inclusion of the method and avoid detailed derivations:

First the Fourier Matrix is post-multiplied by a permutation matrix $P_N$ that would permute the columns of it in such a way that all even columns proceed the odd columns. Such a permutation matrix would look like an identity matrix whose 1 entries on its odd rows (rows for odd $n$) are shifted $N/2$ places. The resultant matrix’ top left and bottom left quadrants turn out to be equal to $\mathcal{F}_{N/2}$, and bottom right quadrant turns out to be $W^{N/2} = -1$ times bottom left quadrant ($W^n = W^{n-N} \mod N.$) For now, let us call the top right quadrant $N/2 \times N/2$ square matrix $A_{N/2}$.

**Example 10.** Let us show this step of Fourier Matrix factorization for $N = 8$. Multiplication of the Fourier Matrix and the even-odd shift permutation matrix would look like:

\[
\begin{array}{l}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & W & W^2 & W^3 & W^4 & W^5 & W^6 & W^7 \\
1 & W^2 & W^3 & W^4 & W^5 & W^6 & W^7 & W^1 \\
1 & W^3 & W^4 & W^5 & W^6 & W^7 & W^2 & W^3 \\
1 & W^4 & W^5 & W^6 & W^7 & W^2 & W^3 & W^4 \\
1 & W^5 & W^6 & W^7 & W^1 & W^2 & W^3 & W^4 \\
1 & W^6 & W^7 & W^1 & W^2 & W^3 & W^4 & W^5 \\
1 & W^7 & W^1 & W^2 & W^3 & W^4 & W^5 & W^6 \\
\end{array}
\begin{array}{l}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & W^2 & W^3 & W^4 & W^5 & W^6 & W^7 & W^1 \\
1 & W^3 & W^4 & W^5 & W^6 & W^7 & W^2 & W^3 \\
1 & W^4 & W^5 & W^6 & W^7 & W^2 & W^3 & W^4 \\
1 & W^5 & W^6 & W^7 & W^1 & W^2 & W^3 & W^4 \\
1 & W^6 & W^7 & W^1 & W^2 & W^3 & W^4 & W^5 \\
1 & W^7 & W^1 & W^2 & W^3 & W^4 & W^5 & W^6 \\
\end{array}
\begin{array}{l}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & W^2 & W^3 & W^4 & W^5 & W^6 & W^7 & W^1 \\
1 & W^4 & W^5 & W^6 & W^7 & W^2 & W^3 & W^4 \\
1 & W^5 & W^6 & W^7 & W^1 & W^2 & W^3 & W^4 \\
1 & W^6 & W^7 & W^1 & W^2 & W^3 & W^4 & W^5 \\
1 & W^7 & W^1 & W^2 & W^3 & W^4 & W^5 & W^6 \\
\end{array}
\begin{array}{l}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & W^2 & W^3 & W^4 & W^5 & W^6 & W^7 & W^1 \\
1 & W^4 & W^5 & W^6 & W^7 & W^2 & W^3 & W^4 \\
1 & W^5 & W^6 & W^7 & W^1 & W^2 & W^3 & W^4 \\
1 & W^6 & W^7 & W^1 & W^2 & W^3 & W^4 & W^5 \\
1 & W^7 & W^1 & W^2 & W^3 & W^4 & W^5 & W^6 \\
\end{array}
\begin{array}{l}
= (\mathcal{F}_{N/2} A_{N/2}) .
\end{array}
\]

Now we will focus on the $A_{N/2}$ square matrix. One can see that there is a relation between $\mathcal{F}_{N/2}$ and $A_{N/2}$; after all, they are even and odd powers of $W$. It turns out that $A_{N/2}$ is just $\mathcal{F}_{N/2}$ pre-multiplied by a diagonal matrix of $N/2$, that is denoted by $D_{N/2}$, with entities of $1, W, W^2, \ldots, W^{N/2-1}$. 

50
Let us formulate what we have done:

\[
\mathcal{F}_N P_N = \begin{bmatrix}
\mathcal{F}_{N/2} & D_{N/2} \mathcal{F}_{N/2} \\
\mathcal{F}_{N/2} & W^{N/2} D_{N/2} \mathcal{F}_{N/2}
\end{bmatrix},
\]

where \( W^{N/2} = -1 \) and \( P_N^{-1} = P_N^T \) (from Section 2.2.2.) Then, we can express the factorization of Fourier Matrix as:

\[
\mathcal{F}_N = \begin{bmatrix}
I_{N/2} & D_{N/2} \\
I_{N/2} & -D_{N/2}
\end{bmatrix} \begin{bmatrix}
\mathcal{F}_{N/2} \\
\mathcal{F}_{N/2}
\end{bmatrix} P_N^{T}.
\]

One sees the similarity of these matrix operations to the operations discussed about FFT.

2.5. Uniform Polygons

As it will be seen more clearly in the next chapter, the uniform polygons play a very crucial role in this dissertation. In addition to providing a visual representation to the some of the mathematical concepts discussed, uniform polygons are the foundation of the relationships structure that the information theoretical framework is based on. Throughout this study we use the term polygon loosely. An extended discussion about “polygons,” in the sense of their relationships with permutations, can be found in [17]. In this section, however, we briefly give the definitions and properties of “uniform polygons.” For the rest of the dissertation, we will use the term “polygon” without quotas to mean our loose definition of polygons following [17, 37].

**Definition 30.** A *polygon* is defined as the sequence of *n* line segments \( A_1 A_2, A_2 A_3, A_3 A_4, \ldots, A_n A_1 \) connecting successive pairs of *n* points \( A_1, A_2, \ldots, A_n \).

**Definition 31.** Given a polygon, the line segments are called *sides*, and the points are called *vertices*. 
Definition 32. Given a polygon, two vertices are called neighbors if they are connected with a side.

Definition 33. The set of all polygons having \( n \) vertices is called \( \text{POLYGON}(n) \).

Definition 34. If the vertices of a polygon occupy a place on a circle and divide a circle into \( n \) equal arcs, it is said to be a stellated polygon. The sides of a stellated polygon can be thought as the chords of a circle. The set of all stellated polygons having \( n \) vertices is called \( \text{STELLATED}(n) \).

Definition 35. \( \text{Rotate Polygon} \) is an operation defined on \( \text{STELLATED}(n) \), and denoted as

\[
P_j = \text{RotatePolygon}(P_i),
\]

where, \( P_i, P_j \in \text{STELLATED}(n) \). The vertices of \( P_j \) are obtained by applying a rotation operation through angle \( 2\pi/n \) to every vertex of \( P_i \).

Definition 36. Let \( P_i \) be a stellated polygon. If \( \text{RotatePolygon}(P_i) \) has the same geometric figure with \( P_i \), then \( P_i \) is said to be a generalized regular polygon. The set of all generalized regular polygons having \( n \) vertices is called \( \text{REGULAR}(n) \).

Definition 37. Let \( n \) and \( q \) be integers. Let \( S \) be a rotation trough angle \( 2\pi/n \) and let \( A_1 \) be any point not on the axis of the rotation. Then, the points

\[
A_i = S^i(A_1), \quad (i = 1, 2, 3, \ldots, n)
\]

are the vertices of a uniform polygon, denoted by \( \{n\}_{q} \), whose sides are the segments \( A_iA_{q+i(\text{mod}\ n)} \), where \( i = 1, 2, 3, \ldots, n \) and \( A_{0(\text{mod}\ n)} = A_n \).

Remark 33. A uniform polygon can be constructed by drawing the segments \( A_1A_{q+1(\text{mod}\ n)} \), \( A_2A_{q+2(\text{mod}\ n)} \), \ldots, \( A_nA_q \), respectively, where \( A_0 = A_n \). The uniform polygon \( \{8\}_{3} \) is drawn in Fig. 8 as an example.
Fig. 8. Drawing 8 by 3 uniform polygon.

**Remark 34.** Alternatively, a uniform polygon can be drawn by drawing the segments $A_1 A_{q+1 (\text{mod } n)}$, $A_{q+1 (\text{mod } n)} A_{2q+1 (\text{mod } n)}$, $\cdots$, $A_{(n-1)q+1 (\text{mod } n)} A_1$, respectively. The uniform polygon $\{8,3\}$ is drawn again in Fig. 9 as an example.

Fig. 9. Drawing 8 by 3 uniform polygon with an alternate method.

**Remark 35.** By following the drawing procedure described in Remark 34 and writing down each vertex visited in order, we can represent a uniform polygon as a permutation in the form of $$[A_1 \ A_{q+1 (\text{mod } n)} \ A_{2q+1 (\text{mod } n)} \ \cdots \ A_{(n-1)q+1 (\text{mod } n)}].$$ We notice this permutation can be rotated and still represent the same polygon only changing the starting vertex. We also notice that the process described above is reversible; that is, for a given permutation a corresponding uniform polygon can be drawn.

**Definition 38.** The set of all uniform polygons having $n$ vertices is called $\text{UNIFORM} (n)$.

**Definition 39.** Uniform polygons with $q = 1$ are called *ordinary polygons* or $n$-gons, and the set that contains the ordinary polygons having $n$ vertices is called $\text{ORDINARY} (n)$.
Definition 40. Uniform polygons \( \{ \frac{n}{q} \} \), \( n \) and \( q \) being relatively prime integers, are called star polygons, and the set of all star polygons having \( n \) vertices is called \( \text{STAR}(n) \).

Definition 41. Uniform polygons \( \{ \frac{n}{q} \} \), \( n \) and \( q \) being not relatively prime integers, are called compound polygons. and the set of all compound polygons having \( n \) vertices is called \( \text{COMPOUND}(n) \).

Remark 36. A Venn diagram of the polygon sets is given in Fig. 10. Polygons and their classification are discussed in more detail in [17].

Remark 37. While drawing a compound polygon with the alternative method, one ends up drawing same regular polygon over and over without reaching vertex \( A_n \). In these cases, the regular polygon drawn first is called the first cycle. After completing the cycle at the vertex \( A_i \), one needs to go vertex \( A_{i+1} \) without drawing a side and continue drawing to vertex \( A_{i+q+1} \). Here \( i \) is determined by the number of the cycle being drawn. An example drawing of compound polygon \( \{ \frac{8}{2} \} \) is given in Fig. 11. The greatest common divisor of \( n \) and \( q \) gives the number of cycles necessary to draw a compound polygon.

Remark 38. By following the drawing procedure described in Remark 34 and writing down each vertex visited in order, we can represent a compound polygon as a permutation in the form of

\[
\begin{bmatrix}
A_1 & A_{q+1} \pmod{n} & A_{2q+1} \pmod{n} & \cdots & A_{(n/\gcd(n,q)-1)q+i} \pmod{n} & A_{i+1} & \cdots & A_n
\end{bmatrix}.
\]
Fig. 11. Drawing 8 by 2 uniform polygon with the alternate method.

Remark 39. \( \{ \frac{n}{q} \} \) and \( \{ \frac{n}{n-q} \} \) uniform polygons corresponds to the same figure.

As a summary of this uniform polygons, section an example of each type of these polygons is given in Fig. 12.

2.6. Summary

In this chapter, we have surveyed various analysis tool-sets to be used in the construction of the proposed modeling and analysis framework. Then we have presented their specific features or representations that would enable us to connect them to each other and to the framework introduced in the remainder of the dissertation. The definitions and classification provided in [38, 37, 39], and [17] are adjusted for notational and conceptual precision in establishing the framework. The following have been examined so far: a) complex numbers and functions, with the focus on the exponential functions and the roots of unity, b) permutations with focus on some special permutations, c) Chebyshev polynomials with focus on the second kind of Chebyshev polynomials, d) topics that are studied under Fourier analysis, and e) polygons in general and uniform polygons in particular.
Fig. 12. Polygon examples: (a) an 8-segment polygon, (b) an 8-segment stellated polygon, (c) the 8-segment ordinary polygon (octagon), (d) 8 by 3 star polygon, (e) 8 by 2 compound polygon.
CHAPTER 3
THE NOISY COMMUNICATION CHANNEL FRAMEWORK

We call the information theoretical modeling and analysis framework that we introduce in this chapter “the noisy communication channel framework.” Noisy communication channels have been used in modeling software components in [40, 41]. An extended version of that study has been recorded in [42, 43, 44, 45, 46]. Others have also used different aspects of communication channels and their capacity, as well as other concepts of information theory, such as entropy and code distance, for modeling and analysis purposes [47, 48, 49, 50]. In this research, we extend this idea and create a framework of noisy communication channels to model and analyze systems in general. We develop this framework through relationships between noisy communication channels and some analysis methods that are commonly used in science and engineering, most of which were surveyed in the previous chapter. We use some of our preliminary work as basis for the framework [51, 52].

The framework based on a representation of a noisy communication channel as a bipartite graph, which can also be represented as an error content graph. It should also be noted that graph representations of concepts are useful, even if it is beyond the scope of this study, through these representations of graph theory can be included later to extend the framework [53]. Although the error content graph can be represented in any arbitrary form, in order to connect this approach to the concepts and relationships mentioned before, the error content graph is chosen to be in the form of regular and irregular uniform polygons. This choice also allows us to make another important connection to be included in the framework, that is, the connection to the roots of unity. This relationship makes the framework even more powerful by opening a door to the world of complex analysis.
According to this framework, a system with a set of inputs and a set of outputs can be modeled by a noisy communication channel by considering its input set as source signals to the noisy channel and its output set as the signals interpreted by the receiver.

3.1. Noisy Communication Channels and Their Representations

A communication channel can be defined as the medium through which the information is transmitted from the information source to the receiver [54]. The noise that is present in the channel causes alteration of the symbol at the output of the transmitter during the transmission. Therefore, a symbol can be received as something other than the one that was sent. This altering behavior defines the error introduced by the channel. Claude E. Shannon has defined the channel and its capacity, more specifically his zero error capacity theorem, in his famous work [55].

A simplified representation of a communication system is shown in Fig. 13, with the dashed box representing the communication channel. The communication channel representation shown in Fig. 13 can be reduced to a simpler form, as in Fig. 14.

![Fig. 13. A communication system.](image)

![Fig. 14. Block diagram of a communication channel.](image)

The relations between source and receiver symbols are shown in Fig. 15. The symbols on the left correspond to the desired symbols to be received (also the symbols that were sent), and the symbols on the right side show which symbol(s) each sent one maps to. The information contents of the lowercase and uppercase symbols are the same. The
lowercase and uppercase notations are to distinguish between the source and receiver side of the communication channel. If a lowercase letter corresponds to the same letter (uppercase) on the other side, then no error is encountered during the information transmission. In some channel models, each symbol that is sent may map to more than two symbols at the receiver side.

Fig. 15. A communication channel.

We have described the noisy communication channel concept, and now we introduce four different ways that noisy communication channels are represented; namely, as probability matrices, error content graphs, permutations, and adjacency matrices.

3.1.1. Shannon’s Noisy Communication Channel Representation

Shannon represents a noisy communication channel in terms of its communication error probabilities, and based on these probabilities, derives the capacity of such communication channel [54]. Since channel capacity is outside of the scope of this study, we will describe only the graphical representation shown in Fig. 16.

In this representation, the vertices represent input and output signals, and the arrows represent the probability of a particular input to be received as a particular output. No arrow between a particular input-output pair simply means that there is no possibility of that input signal being received as that output signal. For a given input signal, the sum of
all probabilities must be equal to zero. We notice that the noise is represented by $p_{io}$, where $i \neq o$.

This approach also enables us to utilize the matrix representation of a noisy communication channel in terms of its communication error probabilities. The probability matrix representation of the noisy communication channel shown in Fig. 16 is as follows:

$$
\begin{bmatrix}
    p_{aa} & p_{ab} & 0 & 0 & 0 \\
    0 & p_{bb} & p_{bc} & 0 & 0 \\
    0 & 0 & p_{cc} & p_{cd} & 0 \\
    0 & 0 & 0 & p_{dd} & p_{de} \\
    p_{ea} & 0 & 0 & 0 & p_{ee}
\end{bmatrix}
$$

### 3.1.2. Error Content Graph of Noisy Communication Channels

Closely examining the channel model in Fig. 15 and focusing on the errors occurring, one can see that the error pairs are as follows: $a \rightarrow B$, $b \rightarrow C$, $c \rightarrow D$, $d \rightarrow E$, and $e \rightarrow A$. Each of these symbols may be considered as a point on a circle, and the error pairs can be considered as line segments joining these points. This process would produce the polygon shown in Fig. 17 [38]. The polygon in Fig. 17 is a representation of the content transmission in the communication channel shown in Fig. 15. Thus, this polygon is
said to be the error content transmission graph or shortly the error content graph of the communication channel shown in Fig. 15.

\[ \text{Fig. 17. Polygon (error content graph) describing the communication channel in Fig. 15.} \]

Without going into a rigorous proof, we can deduce by observing Fig. 17 that any noisy communication channel can be represented by a polygon. For our purposes, we will limit this representation set with uniform polygons. This representation can be expressed as a mapping of noise probabilities that define a particular noisy communication channel to sides of a polygon, as it is defined in Definition 30.

### 3.1.3. Permutation Representation of Noisy Communication Channels

We have seen how noisy communication channels are represented by uniform polygons. In this section, we take that representation a step further and include the permutation representation of noisy communication channels, as well. First, we describe how to generate permutations that correspond to uniform polygons following the notation used in Sections 2.2 and 2.5.

**Definition 42.** $\Omega_j(n, q)$ is defined as a permutation of the form

\[ \Omega_j(n, q) = \begin{bmatrix} a_1 & a_2 & \ldots & a_{n/h} \end{bmatrix}, \]  

where, $1 \leq j \leq n$, $a_k = [j - 1 + q(k - 1)] \mod n + 1$, and $h = \gcd(n, q)$. 

61
Example 11. Let \( n = 8, q = 2, \) and \( h = \gcd(8, 2) = 2. \) For \( j = 1, j = 4, \) and \( j = 7, \) the permutations \( \Omega_1(8, 2), \Omega_4(8, 2), \) and \( \Omega_7(8, 2) \) are

\[
\begin{align*}
\Omega_1(8, 2) &= \begin{bmatrix} 1 & 3 & 5 & 7 \end{bmatrix}, \\
\Omega_4(8, 2) &= \begin{bmatrix} 4 & 6 & 8 & 2 \end{bmatrix}, \\
\Omega_7(8, 2) &= \begin{bmatrix} 7 & 1 & 3 & 5 \end{bmatrix}.
\end{align*}
\]

Definition 43. Consider a set of permutations \( \Omega_1, \Omega_2, \ldots, \Omega_N \) having the form

\[
\begin{align*}
\Omega_1 &= \begin{bmatrix} a_{11} & a_{12} & a_{13} & \ldots \end{bmatrix}, \\
\Omega_2 &= \begin{bmatrix} a_{21} & a_{22} & a_{23} & \ldots \end{bmatrix}, \\
& \quad \vdots \\
\Omega_N &= \begin{bmatrix} a_{N1} & a_{N2} & a_{N3} & \ldots \end{bmatrix}.
\end{align*}
\]

The \textit{CONCATENATE} operator generates a permutation \( \Omega \) having the form

\[
\Omega = \text{CONCATENATE}(\Omega_1, \Omega_2, \ldots, \Omega_N) = \begin{bmatrix} a_{11} & a_{12} & a_{21} & \ldots & a_{N1} & \ldots \end{bmatrix}.
\]

Example 12. Let us concatenate \( \Omega_1(8, 2) - \Omega_4(8, 2) \) and \( \Omega_4(8, 2) - \Omega_7(8, 2) \) permutations obtained in the previous example

\[
\begin{align*}
\Omega &= \text{CONCATENATE}(\Omega_1(8, 2), \Omega_4(8, 2)) \\
&= \begin{bmatrix} 1 & 3 & 5 & 7 & 4 & 6 & 8 & 2 \end{bmatrix} \\
\Omega &= \text{CONCATENATE}(\Omega_4(8, 2), \Omega_7(8, 2)) \\
&= \begin{bmatrix} 4 & 6 & 8 & 2 & 7 & 1 & 3 & 5 \end{bmatrix}.
\end{align*}
\]
Definition 44. \( \chi_k \) is a function defined from \( \text{UNIFORM}(n) \) to \( \text{PERMUTATION}_\Omega(n) \) producing the corresponding permutation for a particular uniform polygon, and denoted as

\[
\Omega = \chi_k(P)
\]

where \( P = \left\{ \frac{n}{q} \right\} \in \text{UNIFORM}(n) \) and \( \Omega \in \text{PERMUTATION}_\Omega(n) \) such that

\[
\Omega = \text{CONCATENATE} \left( \Omega_k(n, q), \Omega_{k+t}(n, q), \Omega_{k+2t}(n, q), \ldots, \Omega_{k+(h-1)t}(n, q) \right),
\]

with \( h = \gcd(n, q), \quad t = (q-1)(q+1), \quad \text{and} \quad k \leq n \) indicating the starting vertex. Note that indices of \( \Omega \) permutations are modulo \( n \).

Example 13. Let \( P \) be the compound \( \left\{ \frac{14}{2} \right\} \), and \( k = 4 \). Then \( h = \gcd(14, 2) = 2 \), and \( t = (2-1)(2+1) = 3 \). The permutation \( \Omega \) is given by

\[
\Omega = \text{CONCATENATE} \left( \Omega_4(14, 2), \Omega_7(14, 2) \right).
\]

Using (3.44) we find these permutations to be

\[
\Omega_4(14, 2) = \begin{bmatrix} 4 & 6 & 8 & 10 & 12 & 14 & 2 \end{bmatrix}, \\
\Omega_7(14, 2) = \begin{bmatrix} 7 & 9 & 11 & 13 & 1 & 3 & 5 \end{bmatrix}.
\]

Thus, \( \Omega \) becomes

\[
\Omega = \begin{bmatrix} 4 & 6 & 8 & 10 & 12 & 14 & 2 \\
7 & 9 & 11 & 13 & 1 & 3 & 5 \end{bmatrix}.
\]

Remark 40. Without detailed mathematical deduction, let us briefly touch upon the permutation representation of irregular uniform polygons, as well, since some noisy communication channels correspond to irregular uniform polygons. What is different about irregular
uniform polygons is that there is not a constant $q$ value, instead, there is a vector of $q$ values defining the modal difference between a pair of vertices connected by a side.

**Definition 45.** $\chi_k'$ is a function defined from $UNIFORM(n)$ to $PERMUTATION_\Omega(n)$ producing the corresponding permutation for a particular uniform polygon, and denoted as

$$\Omega' = \chi_k'(P')$$

where $P' = \{ [n] \} \in UNIFORM(n)$ and $\Omega' \in PERMUTATION_\Omega(n)$ such that

$$\Omega'(n, [q]) = \begin{bmatrix} a_1 & a_2 & a_3 & \ldots & a_{n-1} \end{bmatrix},$$

where $a_k = a_{k-1} + q[k - 1]$ for $k > 1$, and $a_1 = 1$. Note that indices of $a$ are modulo $n$. In order to write the permutation starting from a vertex other than 1, the permutation needs to be rotated until the desired vertex is in the first place.

**Remark 41.** Negative $q$ values are also valid, indicating clockwise direction.

**Example 14.** Let $P'$ be the irregular polygon $\{ [3, 1, 3, -2, 8, -3, -1, -3, 2] \}$. The permutation $\Omega$ is given by


**Remark 42.** In Section 3.1.2, we saw that noisy communication channels can be represented as uniform polygons, and in this section, we see that regular, compound and other irregular uniform polygons can be represented by permutations. Therefore noisy communication channels can be represented by permutations.

**Example 15.** We can deduce that the permutation representation of the noisy communication channel in Fig. 16, whose uniform polygon representation is given in Fig. 17, is

$$\begin{bmatrix} a & b & c & d & e \end{bmatrix}.$$
3.1.4. Adjacency Matrix of Noisy Communication Channels

The adjacency matrix of a noisy communication channel can be obtained both from its permutation matrix or uniform polygon. This shows that the process of going from polygon to adjacency matrix is reversible.

**Definition 46.** A $2n$ by $2n$ matrix whose first $n$ rows and first $n$ columns represent inputs and last $n$ rows and last $n$ columns represent outputs of a particular noisy communication channel, and its elements are 0, except if they correspond to an input-output pair that is connected in the Shannon representation, is called the **adjacency matrix** for that noisy communication channel. All non-zero elements are 1.

**Remark 43.** We can obtain the adjacency matrix of a noisy communication channel by replacing all of the non-zero elements of its $n$ by $n$ probability matrix with 1, then placing it to the right top quadrant of a $2n$ by $2n$ matrix where elements of its other quadrants are 0.

**Example 16.** Here is the adjacency matrix of the noisy communication channel in Fig. 16 from its probability matrix given in Section 3.1.1:

\[
\begin{bmatrix}
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

**Remark 44.** Using a similar method, we can also obtain the adjacency matrix of a noisy communication channel from its permutation representation. The elements of all the quadrants other that the top right quadrant will be 0. The diagonal elements of the $n$ by $n$ right-top quadrant matrix will be 1, since they represent a message being transmitted correctly. Then, each consecutive (including last and first) elements pair of the noisy communication
channel permutation would correspond to an element of 1 on the $n$ by $n$ right-top quadrant matrix, where the rest of its elements (except the diagonal) remain 0. A Mathematica implementation of this Remark is included in Appendix A.

3.2. Relationships Between Noisy Communication Channels and Some Common Modeling and Analysis Tools

In this section, a structure of relationships is built on the common notation presented in Chapter 2 and the definition and representations of the noisy communication channel concept introduced earlier in Chapter 3.

In the previous section, we defined the concept of noisy communication channel and illustrated its error content as a bipartite graph, then as a uniform polygon. In this section, we look at uniform polygons as representation of roots of unity to use them to connect noisy communication channels to complex analysis; as representations of special permutations to use them to connect noisy communication channels to combinatorial analysis; as geometric figures where trigonometric functions can be applied to use them connect noisy communication channels to Chebyshev analysis; and as plots of function series to use them connect noisy communication channels to Fourier analysis.

3.2.1. Noisy Communication Channels and Complex Analysis

Thanks to the common notation already established in Chapter 2 and uniform polygons representation of noisy communication channels discussed above, we are able to show this relationship rather easily. In Theorem 1, roots of unity was defined, and the visualization of roots of unity on the complex plane was described in Remark 16. Here, we use this visualization in order to represent noisy communication channels (already represented with polygons in the previous section) with roots of unity, for instance, the roots of unity shown on the complex plane that correspond to the noisy communication channel represented in Fig. 16, as shown in Fig. 18.
3.2.2. Noisy Communication Channels and Combinatorial Analysis

In this section, we further investigate the permutation representation of noisy communication channels introduced in Section 3.1.3, focusing on a special set of permutations, namely the ones that correspond to the N-queens solutions introduced in Section 2.2.1. A function from uniform polygons representing error content graphs of corresponding noisy communication channels to permutations will be defined.

3.2.2.1. Coding aspect. Before we introduce this function, let us briefly touch upon another aspect of the special permutations that correspond to N-queens solutions, since it is related to coding, which is another major subject in information theory. Following Berge [38], whom we mentioned about in Section 3.1.2 where we introduced error content graphs of noisy communication channels, there can be only two non-adjacent letters picked at a
time from the graph shown in Fig. 17. Picking three letters is out of the question, since two would result in being adjacent vertices of the graph. For example, if $a$ and $d$ are picked, the following two letter code-words can be constructed: $aa, ad, da,$ and $dd$. No matter what error combination occurs in the channel during transmission, these “combined” inputs or two-word letters cannot be confused with one another. The number of words in this vocabulary is $(\alpha(G))^2 = 4$, where $\alpha(G)$ is defined as stability number of graph $G$.

**Definition 47.** The stability number of a graph $G$ is defined as the maximum cardinality of its stable sets. In other words, the stability number of a graph is defined as finding a stable set that has a maximum number of elements.

**Definition 48.** A stable set of a graph is a subset of the graph $G$ that has no adjacent vertices in it.

For the communication channel shown in Fig. 15, Berge [38] stated that the number of non-adjacent, two-letter code words increases when we consider the graph $G \cdot G = G^2$, where $G$ is the graph shown in Fig. 17. Following Berge, the graph $G^2$ is shown in Fig. 19a.

In the case of using $G^2$, the non-adjacent code words correspond to elements of a stable set of the graph $G^2$. The two-letter code words are formed by the marked vertices in Fig. 19b. For this case, the set of two-letter code words consists of $aa, bc, ce, db,$ and $ed$. Obviously, this code is a richer one. Using such a set of non-confusable inputs for a computational element brings forth the possibility of realizing reversible computation. Although we do not get into much detail in this study, it is important to note that the reversible computation is one of the features of next generation computational elements [56, 57, 58, 59].

Taking a closer look at Fig. 19b: marking down the non-adjacent points; in other words, finding the stable sets corresponds to finding the queen solutions for that grid size. Hence, finding the stable sets of a simple graph with $n^2$ vertices (i.e., stable sets of the
Fig. 19. (a) The graph $G^2$ of graph $G$ shown in Fig. 17, (b) a stable set of $G^2$.

diagram

graph $G^2$, where $n$ is number of vertices in graph $G$) is equivalent to finding an N-queens solutions for the grid size of $n$.

3.2.2.2. Noisy communication channels corresponding to the solutions. Keeping in mind that uniform polygons are one of the ways the noisy communication channels can be represented (Section 3.1.2,) we here define a function from uniform polygons representing the error content graphs of corresponding noisy communication channels to permutations. Noticing that no corresponding ordinary polygon satisfies an N-queens solution, we only include star polygons and compound polygons in this section.

Let us start with solution permutations corresponding to star polygons. Since $\text{STAR}(n)$ is a subset of $\text{UNIFORM}(n)$, we can use the $\chi_k$ function defined in Definition 44 with its simpler form, because of $h = \gcd(n, q)$ being 1.

**Definition 49.** $\zeta_k$ is a function defined from $\text{STAR}(n)$ to $\text{PERMUTATION}_\Omega(n)$, and denoted as

$$\Omega = \zeta_k(P)$$

where $P = \left\{ \frac{n}{q} \right\} \in \text{STAR}(n)$ such that $\Omega = \Omega_k(n, q)$, with $k \leq n$.

**Definition 50.** The range set elements of this function $\zeta_k$ are called star permutations, and this set is denoted as $\text{STAR}_\Omega(n)$. 

Remark 45. For a particular star polygon $P = \{n_q\}$, there are $n$ different permutations in $\textit{STAR}_\Omega(n)$. These permutations are in same order, but every one of them starts from a different vertex of $P$. Notice that all of these permutations correspond to the same noisy communication channel.

Example 17. Let $P$ be the star polygon $\{7_3\}$. The star permutations of this polygon are

\[ \zeta_1(P) = \Omega_1(7, 3) = \begin{bmatrix} 1 & 4 & 7 & 3 & 6 & 2 & 5 \end{bmatrix}, \]
\[ \zeta_2(P) = \Omega_2(7, 3) = \begin{bmatrix} 2 & 5 & 1 & 4 & 7 & 3 & 6 \end{bmatrix}, \]
\[ \zeta_3(P) = \Omega_3(7, 3) = \begin{bmatrix} 3 & 6 & 2 & 5 & 1 & 4 & 7 \end{bmatrix}, \]
\[ \vdots \]
\[ \zeta_7(P) = \Omega_7(7, 3) = \begin{bmatrix} 7 & 3 & 6 & 2 & 5 & 1 & 4 \end{bmatrix}. \]

Definition 51. $\zeta = \zeta_1$ is a function defined from $\textit{STAR}(n)$ to $\textit{PERMUTATION}_\Omega(n)$, and denoted as

\[ \Omega = \zeta(P) \]

where $P = \{n_q\} \in \textit{STAR}(n)$ such that $\Omega = \Omega_1(n, q)$.

Definition 52. The range set elements of this function $\zeta$ are called representative permutations, and this set is denoted as $\textit{REPSTAR}_\Omega(n)$.

Example 18. Let $P$ be the star polygon $\{7_3\}$. The representative permutation of this polygon is

\[ \Omega = \zeta(P) = \Omega_1(7, 3) = \begin{bmatrix} 1 & 4 & 7 & 3 & 6 & 2 & 5 \end{bmatrix}. \]

Remark 46. The function $\zeta$ is a one to one onto function from $\textit{STAR}(n)$ to $\textit{REPSTAR}_\Omega(n)$. Therefore, the cardinalities of these two sets are equal.
Remark 47. For a particular star polygon, the permutations that are rows of the circulant matrix [60] generated by representative permutation of the polygon are star permutations. Therefore, using the representative permutation, the star permutations of the particular polygon can be obtained.

Example 19. Let $P$ the star polygon $\{7\}$. The representative permutation of this polygon is $\begin{array}{cccccc}1 & 4 & 7 & 3 & 6 & 2 & 5\end{array}$. The circulant matrix of this permutation is

$$
\begin{bmatrix}
1 & 4 & 7 & 3 & 6 & 2 & 5 \\
5 & 1 & 4 & 7 & 3 & 6 & 2 \\
2 & 5 & 1 & 4 & 7 & 3 & 6 \\
6 & 2 & 5 & 1 & 4 & 7 & 3 \\
3 & 6 & 2 & 5 & 1 & 4 & 7 \\
7 & 3 & 6 & 2 & 5 & 1 & 4 \\
4 & 7 & 3 & 6 & 2 & 5 & 1 \\
\end{bmatrix}
$$

It is obvious that the elements of this matrix are star permutations.

Definition 53. The intersection set of $REPSTAR_\Omega(n)$ and $QUEEN_\Omega(n)$ is called $REP COMMON_\Omega(n)$.

Definition 54. The intersection set of $STAR_\Omega(n)$ and $QUEEN_\Omega(n)$ is called $COMMON_\Omega(n)$.

Remark 48. The relations among these permutation sets are demonstrated as a Venn diagram in Fig. 20.

Next, we will find the elements of $COMMON_\Omega(n)$ set. In order to do that, we need to exclude all the possible permutations that cannot be in this set.

Remark 49. Here are some conditions to determine whether a permutation corresponds to a star polygon and also to a queen solution [17]:

1. If $n = 0, 2, 3, 4 \pmod{6}$, then there is no star polygon that corresponds to an $n$-queens solution.
Fig. 20. Venn diagram of permutation sets.

2. Since \( \{ \frac{n}{q} \} \) and \( \{ n \frac{n}{q} \} \) are the same polygons and correspond to the same \( n \)-queens solutions, let \( q \) be smaller than or equal to \((n - 1)/2\).

3. According to the definition of star polygon, for a polygon \( P = \{ \frac{n}{q} \} \), \( n \) and \( q \geq 2 \) must be relatively prime numbers.

4. Let \( P = \{ \frac{n}{q} \} \) be a star polygon. If \( n \) and \( j - 1 \) are not relatively prime, then \( P \) does not correspond to an \( n \)-queens solution.

5. Let \( P = \{ \frac{n}{q} \} \) be a star polygon. If \( n \) and \( j + 1 \) are not relatively prime, then \( P \) does not correspond to an \( n \)-queens solution.

Remark 50. By using these conditions and finding all proper \( n, q \) couples, a \( \text{REPCOM-MON}_\Omega (n) \) set can be constructed. Since every representative permutation can generate its star permutations, \( \text{COMMON}_\Omega (n) \) can also be constructed.

Example 20. Let us find the permutations that are elements of \( \text{COMMON}_\Omega (7) \). First, we need to find the elements of \( \text{REPCOMMON}_\Omega (7) \) using the above-mentioned five conditions. Since \( 7 = 1 \pmod{6} \), the first condition is satisfied. According to the second condition, \( q \leq 3 \). To satisfy the third condition, the possible \( q \) values are 2 and 3. These
two values also satisfy the fourth and fifth conditions. Then, we can find the representative permutations for these two star polygons using the $\zeta$ function:

$$
\begin{bmatrix}
1 & 3 & 5 & 7 & 2 & 4 & 6 \\
1 & 4 & 7 & 3 & 6 & 2 & 5
\end{bmatrix}
$$

respectively. These two permutations are the elements of $REPCOMMON_\Omega(7)$. The star permutations can be found using either the $\zeta_k$ function or circulant matrices of representative permutations. Here are the elements of $COMMON_\Omega(7)$

$$
\begin{bmatrix}
1 & 3 & 5 & 7 & 2 & 4 & 6 \\
2 & 4 & 6 & 1 & 3 & 5 & 7 \\
3 & 5 & 7 & 2 & 4 & 6 & 1 \\
\vdots
\end{bmatrix}
\begin{bmatrix}
1 & 4 & 7 & 3 & 6 & 2 & 5 \\
2 & 5 & 1 & 4 & 7 & 3 & 6 \\
3 & 6 & 2 & 5 & 1 & 4 & 7 \\
\vdots
\end{bmatrix}
\begin{bmatrix}
7 & 2 & 4 & 6 & 1 & 3 & 5 \\
7 & 3 & 6 & 2 & 5 & 1 & 4
\end{bmatrix}
.$$
Example 21. Let $P$ be the compound $\{\frac{8}{2}\}$. Then $h = \gcd(8, 2) = 2$, and the permutation $\Omega = f_c(P)$ is given by

$$\Omega = \text{CONCATENATE} (\Omega_2(8, 2), \Omega_1(8, 2)).$$

$\Omega_1(8, 2)$ and $\Omega_2(8, 2)$ can be found using (3.44). Thus, $V$ is given by

$$\Omega = \begin{bmatrix} 2 & 4 & 6 & 8 & 1 & 3 & 5 & 7 \end{bmatrix}.$$

Theorem 9. Let $P_1 = \{\frac{n}{q}\}$ be a star polygon, where $n = 1 + 6\alpha$, or $n = 5 + 6\alpha$. Let $(n - 1)$ be divisible by $q$. Then $V = f_c(P_2)$ provides a solution to the $(n - 1)$-queens problem, where $P_2 = \{\frac{n - 1}{q}\}$ is a compound polygon.

Proof: Let $h = \gcd(n - 1, q)$. Since $(n - 1)$ is divisible by $q$, it is obvious that $h = q$. From the previous definition the permutation $\Omega$ is given as

$$\Omega = \text{CONCATENATE} (\Omega_h(n - 1, q), \Omega_{h-1}(n - 1, q), \ldots, \Omega_1(n - 1, q)).$$

Since $h = q$, the permutations can be written as

$$\Omega_h(n - 1, q) = \begin{bmatrix} q & 2q & 3q & \ldots \end{bmatrix},$$

$$\Omega_{h-1}(n - 1, q) = \begin{bmatrix} q - 1 & 2q - 1 & \ldots \end{bmatrix},$$

$$\vdots$$

$$\Omega_1(n - 1, q) = \begin{bmatrix} 1 & q + 1 & 2q + 1 & \ldots \end{bmatrix}.$$

On the other hand, star polygon $P_1$ places the queens on the board as permutation

$$\begin{bmatrix} 1 & q + 1 & 2q + 1 & \ldots & mq \mod n & \ldots \end{bmatrix}.$$
If column 1 and row 1 of the chessboard representation of this solution are removed from
the board, the permutation $\Omega$ is obtained:

$$
\begin{bmatrix}
q & 2q & 3q & \ldots \\
\end{bmatrix}.
$$

Thus $\Omega = f_c(P_2)$ provides a solution to the $(n-1)$-queens problem.

**Example 22.** Let $P$ be the compound polygon $\{12\}_4$. Since the star polygon $\{13\}_4$ provides
a 13-queens solution, $f_c(P)$ provides a 12-queens solution.

**Definition 55.** The integer set $E(w)$ is defined as

$$
E(w) = \begin{cases}
\{\} & , w = 1 \\
\{i \mid i = wa, \text{ where } a = 1, 2, 3, \ldots\} & , w \geq 2
\end{cases}.
$$

**Definition 56.** Let $q$ be a positive integer. Further, assume that $c_1, c_2, c_3, \ldots, c_l$ are the
prime factors of $(q+1)$ and $d_1, d_2, d_3, \ldots, d_m$ are the prime factors of $(q-1)$. Then the set
$S(q)$ is defined as

$$
S(q) = E(q) - (E(c_1) \cup \ldots \cup E(c_l) \cup E(d_1) \cup \ldots \cup E(d_m)).
$$

**Example 23.** Here are the sets $S(2), S(3), S(4),$ and $S(5)$

$$
S(2) = E(2) - (E(1) \cup E(3)) = \{2, 4, 8, 10, \ldots\},
$$

$$
S(3) = E(3) - E(2) = \{3, 9, 15, 21, \ldots\},
$$

$$
S(4) = E(4) - (E(3) \cup E(5)) = \{4, 8, 16, 28, \ldots\},
$$

$$
S(5) = E(5) - (E(2) \cup E(3)) = \{5, 25, 35, 55, \ldots\}.
$$

**Theorem 10.** Let $P = \{n\}_q$ be the compound polygon, where $n \in S(q)$. Then, for all
\begin{align*}
[(h-1)(q+1)] + 1 \leq k \leq n + q - q[(h-1)(q+1)],
\end{align*}
the permutation $\Omega = \chi_k(P)$ provides
an $n$-queens solution.
The proof of this theorem can be found in [17].

**Example 24.** Let $P$ be $\{\frac{10}{2}\}$. Then, for all $3 \leq k \leq 8$, the permutation $\Omega = \chi_k(P)$ provides an $n$-queens solution, i.e., for case $k = 4$

$$
\Omega = \begin{bmatrix} 4 & 6 & 8 & 10 & 2 & 7 & 9 & 1 & 3 & 5 \end{bmatrix}.
$$

$P$ and corresponding chessboard configuration is given in Fig. 21.

Fig. 21. (a) 10 by 2 compound polygon, (b) corresponding chessboard configuration for case $k = 4$.

In summary, we have introduced how uniform polygon permutations corresponding to N-queen solutions can be obtained when the uniform polygon in question is a star polygon and when it is a compound polygon. Knowing that these uniform polygons can be thought of as error content graphs of noisy communication channels, we can use this deduction to find noisy communication channels that correspond to special permutations that represent N-queens solutions. We notice that, for each representative permutation, there exists one uniform polygon and one corresponding noisy communication channel; however, there exist as many N-queens solutions as the number of permutations that can be generated from the representative permutation via circulation. To better understand the process, let us consider the following example.
Example 25. Let us consider the noisy communication channel in Shannon’s representation, as shown in Fig. 22a. Following the procedure described in Section 3.1.2, we obtain the error content graph of this noisy communication channel, as shown in Fig. 22b. This error content graph is obviously the \( \left\{ \frac{5}{2} \right\} \) star polygon. According to Definitions 51 and 52, the corresponding representative permutation is \([1 \ 3 \ 5 \ 2 \ 4]\). Now we can use Remark 49 and verify that this permutation and others in its circulant matrix do correspond to N-queens solutions, one of which is shown in Fig. 22c. We can also do this process in reverse; i.e., we can start from a solution and find its corresponding noisy communication channel.

![Figure 22](image)

Figure 22. (a) Shannon representation of a noisy communication channel, (b) its corresponding 5 by 2 star polygon error content graph, and (c) its corresponding 5-queens solution.

3.2.3. Noisy Communication Channels and Chebyshev Analysis

In this section, we show the relationship between the Chebyshev polynomials of second kind and its associated polynomial \( u_p + u_{p-1} \) and some of the properties of uniform polygons that correspond to the error content graphs of noisy communication channels. The notation and style in this section is adopted from [25]. Details of this topic were presented in [61, 62].
Theorem 11. Let \( l_q \) be a side and \( R \) be the circumradius of a uniform polygon of \( \{n_q\} \). Then the ratio of \( l_q/R \) is

\[
\frac{l_q}{R} = 2\sin \left( \frac{q\pi}{n} \right). \tag{3.45}
\]

Fig. 23. Sides of \( n \) by 1, \( n \) by 2, and \( n \) by 3 uniform polygons and their outer circle.

Proof: Let us prove this theorem by using Fig. 23. Let \( l_q \) be the length of a side of the uniform polygon \( \{n_q\} \). In the figure, therefore, \( l_1 \) is the length of a side of uniform polygon \( \{n\} \), \( l_2 \) is the length of a side of uniform polygon \( \{n\} \), and so on. Here \( \overline{AC}, \overline{AH}, \) and \( \overline{AK} \) are the sides of uniform polygons \( \{n\}, \{n\}, \) and \( \{n\} \), respectively. Since \( OA = OC = R \), the \( AOC \) triangle is isosceles. Let \( B \) be the middle point of \( \overline{AC} \). Then \( \overline{OB} \) becomes the altitude of the triangle and is perpendicular to \( \overline{AC} \). Let \( D \) be the point at which the extension of \( \overline{OB} \) intersects the circle. So the arcs \( AD \) and \( DC \) are equal to each other and \( 2\pi/n \). Then the angle \( \widehat{AOB} \) is also equal to \( \pi/n \). Using the same logic, we can show that \( F \) is the middle point of \( \overline{AH} \), \( \overline{OF} \) is parallel to \( \overline{FC} \), and the arc \( AC \) and the angle \( \widehat{AOF} \) are equal to \( 2(2\pi/n)/2 = 2\pi/n \). By applying the same idea to the uniform polygon \( \{n\} \), we can also show that \( \widehat{AOG} = 3(2\pi/n)/2 = 3\pi/n \). By using the definition of \( \sin \) function in a perpendicular triangle, we can state that: \( l_q/2R = \sin \left( \frac{q\pi}{n} \right) \), which leads us to (3.45).
Theorem 12. The zeros of the Chebyshev polynomial of second kind $u_p(x)$, as defined in (2.18), are

$$
\cos \frac{q\pi}{p+1}, \quad (q = 1, 2, \ldots, p).
$$

(3.46)

Proof: Zeros of $u_p(x)$ are the $x$ values that satisfy $u_p(x) = \sin(p+1)\theta = 0$. Since $x = \cos \theta$, the zeros of $u_p(x)$ are $x = \cos \frac{q\pi}{p+1}$.

Remark 51. Let $U_p(x)$ be an order of $p$, symmetric, tridiagonal matrix as in the following form

$$
U_p(x) = \begin{bmatrix}
2x & 1 & 0 & 0 & \cdots & 0 \\
1 & 2x & 1 & 0 & \cdots & 0 \\
0 & 1 & 2x & 1 & \cdots & 0 \\
0 & 0 & 1 & 2x & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1 & 2x
\end{bmatrix}.
$$

Because of the recurrence property of Chebyshev polynomials in (2.19), the following result can be derived [25]

$$
u_p(x) = \det U_p(x).
$$

(3.47)

Theorem 13. The eigenvalues of $U_p(x)$ are

$$
2(x - 1) + 4\sin^2 \frac{q\pi}{2p+2}, \quad (q = 1, 2, \ldots, p).
$$

Proof: In order to calculate the eigenvalues of a matrix, we rewrite the eigenvalue matrix by subtracting the eigenvalue $\lambda$ from each element of the central diagonal, equal its determinant to 0, and solve this equation for $\lambda$. The eigenvalue matrix that can be obtained
by this method is \( U_p \left( x - \frac{\lambda}{2} \right) \), and the determinant of this matrix is \( U_p \left( x - \frac{\lambda}{2} \right) \). Equaling this determinant to zero and solving it for \( \lambda \) means finding zeros of \( U_p \left( x - \frac{\lambda}{2} \right) \) in terms of \( \lambda \). From (3.46) \( x - \frac{\lambda}{2} = \cos \frac{q\pi}{p+1} \) and \( \lambda = 2x - 2\cos \frac{q\pi}{p+1} \). By applying the double-angle formula of \( \cos \) function and rearranging the equation, we obtain the eigenvalues of \( U_p (x) \) as \( 2(x - 1) + 4\sin^2 \frac{q\pi}{2p+2} \).

**Theorem 14.** Let \( v_p (x) = u_p(x) + u_{p-1}(x) \). Then the zeros of \( v_p (x) \) are

\[
\cos \frac{2q\pi}{2p+1}, \quad (q = 1, 2, \ldots, p).
\]

**Proof:** \( v_p (x) = u_p(x) + u_{p-1}(x) = \frac{\sin(p+1)\theta + \sin p\theta}{\sin \theta} \). By applying trigonometric summation and double-angle formulas we can derive \( v_p (x) \) as

\[
v_p(x) = \frac{2 \sin \left( \frac{2p+1}{2} \right) \theta \cos \frac{\theta}{2}}{2 \sin \frac{\theta}{2} \cos \frac{\theta}{2}} = \frac{\sin (p + 1/2) \theta}{\sin \frac{\theta}{2}}.
\]

By using the same method we used to find the zeros of \( u_p (x) \), the zeros of \( v_p (x) \) are \( x = \cos \frac{q\pi}{p+1/2} = \frac{2q\pi}{2p+1} \).

\( v_p (x) \) polynomials up to degree 5 are given in Table 3.

**TABLE 3**

<table>
<thead>
<tr>
<th>( v_p (x) ) Polynomials Up to Degree 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( v_1 (x) )</td>
</tr>
<tr>
<td>( v_2 (x) )</td>
</tr>
<tr>
<td>( v_3 (x) )</td>
</tr>
<tr>
<td>( v_4 (x) )</td>
</tr>
<tr>
<td>( v_5 (x) )</td>
</tr>
</tbody>
</table>
**Remark 52.** Let \( V_p(x) \) be a symmetric, tridiagonal \( p \times p \) matrix as, in the following form

\[
V_p(x) = \begin{bmatrix}
2x & 1 & 0 & 0 & \cdots & 0 \\
1 & 2x & 1 & 0 & \cdots & 0 \\
0 & 1 & 2x & 1 & \cdots & 0 \\
0 & 0 & 1 & 2x & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1 & 2x + 1
\end{bmatrix}.
\]

Again, because of the recurrence property of Chebyshev polynomials of second kind (2.19), the following result can be derived [25]

\[
v_p(x) = \det V_p(x).
\]

**Theorem 15.** The eigenvalues of \( V_p(x) \) are

\[
2(x - 1) + 4\sin^2 \frac{q\pi}{2p+1}, \quad (q = 1, 2, \ldots, p).
\]

**Proof:** Let us use the same method that we used to find the eigenvalues of \( U_p(x) \).

Then, from (3.48), \( x - \frac{2}{2} = \cos \frac{2q\pi}{2p+1} \) and \( \lambda = 2x - 2\cos \frac{2q\pi}{2p+1} \). By applying the double-angle formula of \( \cos \) function and rearranging the equation, we obtain the eigenvalues of \( V_p(x) \) as \(2(x - 1) + 4\sin^2 \frac{q\pi}{2p+1} \).

**Corollary 1.** Let \( l_q \) be a side and \( R \) be the circumradius of a uniform polygon of \( \{n_q\} \). The eigenvalues of the matrix \( U_p(1) \), if \( n = 2p + 2 \), and the eigenvalues of the matrix \( V_p(1) \), if \( n = 2p + 1 \), are the ratios \((l_q/R)^2\). Here \( q = 1, 2, \ldots, p \).

**Proof:** Using the eigenvalue theorems (Theorem 13 and Theorem 15) and geometrical ratio of uniform polygons theorem (Theorem 11), the proof is trivial.

In summary: given a uniform polygon, we have introduced how its corresponding Chebyshev polynomials, matrices, and eigenvalues (in short, Chebyshev properties) are
produced. Knowing that uniform polygons can be thought as error content graphs of noisy communication channels, we can use this deduction to find Chebyshev properties corresponding to noisy communication channels. We notice that for noisy communication channels represented by irregular uniform polygons, there exists an array of \( q \) values; hence, there exists an array of eigenvalues and side lengths. To better understand the process, let us consider the following examples, first for an odd \( n \) and second for an even \( n \).

**Example 26.** Let \( n \) be 7. Since \( 7 = 2p + 1 \), we shall use \( V_p(1) \) and its corresponding Chebyshev polynomial \( v_p(x) \). Here \( p = 3 \), and possible \( q \) values are 1, 2, and 3. Let the circumradius \( R \) be 1. Then, eigenvalues of \( V_3(1) \) gives the values of \( l_1^2, l_2^2, \) and \( l_3^2 \).

\[
V_3(x) = \begin{bmatrix}
2x & 1 & 0 \\
1 & 2x & 1 \\
0 & 1 & 2x + 1
\end{bmatrix}
\]

and the eigenvalues are

\[
\lambda_1 = l_1^2 = 4 \sin^2 \frac{\pi}{7} = 0.753, \quad \lambda_2 = l_2^2 = 4 \sin^2 \frac{2\pi}{7} = 2.445,
\]

\[
\lambda_3 = l_3^2 = 4 \sin^2 \frac{3\pi}{7} = 3.802.
\]

The length of the sides of the uniform polygons can be calculated either directly from polygon geometry or by taking square roots of these eigenvalues:

\[
l_1 = 0.868, \quad l_2 = 1.564, \quad l_3 = 1.950.
\]

For the case \( q = 3 \), uniform polygon representation and Shannon representation of the noisy communication that corresponds to the Chebyshev properties \( v_3(x) \), \( V_3(1) \), and \( \lambda_3 \) are shown in Fig. 24.
Figure 24. (a) 7 by 3 star polygon error content graph a noisy communication channel and (b) its Shannon representation.

Example 27. Let \( n \) be 8. Since \( 8 = 2p + 2 \), we shall use \( U_p(1) \), and the corresponding Chebyshev polynomial \( u_p(x) \). Here \( p = 3 \), and possible \( q \) values are 1, 2, and 3. Let the circumradius \( R \) be 1. Then, eigenvalues of \( U_3(1) \) give the values of \( l_1^2, l_2^2, \) and \( l_3^2 \).

\[
U_3(x) = \begin{bmatrix}
2x & 1 & 0 \\
1 & 2x & 1 \\
0 & 1 & 2x
\end{bmatrix}
\]

and the eigenvalues are

\[
\lambda_1 = l_1^2 = 4 \sin^2 \frac{\pi}{8} = 0.586, \quad \lambda_2 = l_2^2 = 4 \sin^2 \frac{2\pi}{8} = 2, \\
\lambda_3 = l_3^2 = 4 \sin^2 \frac{3\pi}{8} = 3.414.
\]
The length of the sides of the uniform polygons can be calculated either directly from polygon geometry or by taking square roots of these eigenvalues:

\[ l_1 = 0.765, \quad l_2 = 1.414, \quad l_3 = 1.848. \]

For the case \( q = 3 \), the uniform polygon representation and Shannon representation of the noisy communication that corresponds to the Chebyshev properties \( u_3(x) \), \( U_3(1) \), and \( \lambda_3 \) are shown in Fig. 25.

![Diagram](image)

Figure 25. (a) 8 by 3 star polygon error content graph of a noisy communication channel and (b) its Shannon representation.

### 3.2.3.1. Noisy communication channels, Chebyshev polynomials, and special permutations

Given a noisy communication channel, and utilizing the concepts that have been just introduced, we can now introduce step-by-step instructions showing how to find its corresponding Chebyshev polynomials and permutations that represent N-queens solutions. Since ordinary polygons do not represent any N-queens solution, we exclude \( q = 1 \) case and are only interested in error content graphs that can be represented by either star polygons or compound polygons. Since the theorems concerning Chebyshev polynomials...
allow only \( q \leq (n - 1)/2 \) values, we also exclude \( q = n/2 \) case. The instructions are as follows:

1. The error content graph (uniform polygon) for the noisy communication channel is chosen.

2. Using Corollary 1, a corresponding \( U \) or \( V \) matrix is obtained.

3. Using Remarks 51 and 52, a corresponding Chebyshev polynomial is obtained.

4. Eigenvalue equation is obtained, as explained in the proofs of Theorems 14 and 15.

5. Using Theorems 14 and 15, the eigenvalues and side lengths (Corollary 1) are obtained. (Eigenvalues can also be calculated from the eigenvalue equation.)

6. Using Definitions 44 and 8, corresponding permutations are obtained.

7. Using Remark 26, chessboard representations of permutations are obtained.

8. Using Remark 49, for star polygons, it is determined whether the example corresponds to a solution.

9. Using Theorems 9 and 10, for compound polygons, it is determined whether the example corresponds to a solution.

**Example 28.** Let us see these steps on an example:

1. Let the noisy communication channel shown in Fig. 24 be the choice for this example.

2. Using Corollary 1, we find \( V_3(x) = \begin{bmatrix} 2x & 1 & 0 \\ 1 & 2x & 1 \\ 0 & 1 & 2x + 1 \end{bmatrix} \) and \( V_3(1) = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \end{bmatrix} \).

3. Using Remark 52, we find \( v_3(x) = 8x^3 + 4x^2 - 4x - 1 \) and \( v_3(1) = 7 \).
4. Using Theorem 15, we find eigenvalue equation to be \(-\lambda^3 + 7\lambda^2 - 14\lambda + 7\).

5. Using Theorem 15, the eigenvalues and side lengths (Corollary 1) are found to be 
\[ \lambda_3 = l_3^2 = 4\sin^2 \frac{3\pi}{7} = 3.802 \text{ and } l_3 = 1.950. \]

6. Using Definition 44, corresponding representative permutation is 
\[
\begin{bmatrix}
1 & 4 & 7 & 3 \\
6 & 2 & 5
\end{bmatrix}
\]
By circulation, other permutations shown in Table 4 are obtained.

7. Using Remark 26, chessboard representations of permutations are obtained, as shown in Table 4.

8. According to Remark 49, all of these chessboard representations correspond to a 7-queens solution.

For the noisy communication channels that correspond to irregular uniform polygons, we follow the same instructions; however, a vector of \(q\) values are considered instead of a single \(q\) value. Therefore, there will be a vector of ratio squares corresponding to a vector of eigenvalues and vectors of other Chebyshev parameters (characteristic equations, matrices, and polynomials.)

3.2.4. Noisy Communication Channels and Fourier Analysis

In this section, we again use the error-content graph representation (uniform polygons) of noisy communication channels to show how Fourier series connects to the framework. We will first start with briefly introducing Fourier series representation of uniform polygons. This topic is covered in detail in [63].

From (2.27) (notice that the indice is changed from \(n\) to \(k\) to avoid confusing it with number of vertices of a uniform polygon,) a periodic function \(f(t)\) with \(w = 1\) can be represented by the complex (exponential) form of Fourier series

\[
f(t) = \sum_{k=-\infty}^{\infty} c_k e^{jkt},
\]
### TABLE 4

7 by 3 Error Content Graph Permutations and Corresponding 7-Queen Solutions

<table>
<thead>
<tr>
<th>Permutations</th>
<th>Chessboard</th>
<th>Permutations</th>
<th>Chessboard</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1 4 7 3 6 2 5]</td>
<td><img src="image1" alt="Chessboard" /></td>
<td>[2 5 1 4 7 3 6]</td>
<td><img src="image2" alt="Chessboard" /></td>
</tr>
<tr>
<td>[3 6 2 5 1 4 7]</td>
<td><img src="image3" alt="Chessboard" /></td>
<td>[4 7 3 6 2 5 1]</td>
<td><img src="image4" alt="Chessboard" /></td>
</tr>
<tr>
<td>[5 1 4 7 3 6 2]</td>
<td><img src="image5" alt="Chessboard" /></td>
<td>[6 2 5 1 4 7 3]</td>
<td><img src="image6" alt="Chessboard" /></td>
</tr>
<tr>
<td>[7 3 6 2 5 1 4]</td>
<td><img src="image7" alt="Chessboard" /></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
with $|c_k e^{i k t}| = |c_k|$. Fourier series of a regular uniform polygon is given in [64] as

$$f_n(t) = \sum_{k \equiv 1(n)} \frac{e^{i k t}}{k^2}, \quad (3.49)$$

where $k = 1 + nl$ with $l = -p \ldots p$, and $l, p \in \mathbb{Z}$.

We now introduce an algorithm to construct uniform polygons using Fourier series by utilizing (3.49). The number of elements of the summation increases as we increase the $l$ value. This suggests a higher resolution of the polygon figure plotted on the complex plane as $l$ increases. The following parameters are defined to be used in the algorithm:

- $q \in \mathbb{Z}^+$ corresponds to the difference between two vertices connected by a side (we will call this value as number of jump speed,)

- $n \in \mathbb{Z}^+$ corresponds to the number of vertices of the uniform polygon,

- $l \in \mathbb{Z}^+$ determines the summation interval $[-p, p]$ in which the Fourier series will be calculated,

- $N \in \mathbb{Z}^+$ is the sampling factor for the numerical algorithms and it determines how many samples $t \in [t_0, 2\pi + t_0]$ will have.

Remark 53. With the sampling factor $N$, there will be $N$ discrete values of $t$, starting from $t_0$:

$$\left\{ t_0, t_1 = t_0 + \frac{2\pi}{N}, t_2 = t_0 + 2\frac{2\pi}{N}, \ldots, t_{N-1} = t_0 + (N-1)\frac{2\pi}{N}, t_N = t_0 + N\frac{2\pi}{N} \right\}.$$  

Let us write values of $f \left( \frac{t}{N} \right)$, starting from $f \left( \frac{t_0}{N} \right)$:

$$\left\{ f \left( \frac{1}{N} t_0 \right), f \left( \frac{1}{N} (t_0 + 2\pi) \right), f \left( \frac{1}{N} (t_0 + 2(2\pi)) \right), \ldots, f \left( \frac{1}{N} (t_0 + N(2\pi)) \right) \right\}.$$
Since $f(cx) = f(c(x + mT))$, where $f(x)$ is a periodic function with a period $T$, $c$ is a constant, and $m$ is a positive integer, $f\left(\frac{t}{N}\right)$ is equal to $f\left(\frac{t_0}{N}\right)$ for each of the $N$ discrete values of $t$.

Let us now introduce all of the parameters we defined into a numerical algorithm and write each term of the Fourier series equation as

$$u(l, t) = \frac{e^{i(q + nl)t/N}}{(q + nl)^2}, \quad (3.50)$$

where $l = -p, -p + 1, \ldots, -1, 0, 1, \ldots p - 1, p$. Notice the variable change from $t$ to $t/N$ between the functions $f$ and $u$. Since the resulting Fourier series calculation is the sum of each Fourier term at time $t_0$, we can write the summation for the discrete algorithm as

$$u(t) = \sum_{l=-p}^{p} e^{i(q + nl)t/N} \frac{1}{(q + nl)^2}. \quad (3.51)$$

Then, an analytical algorithm, where $t$ is continuous, would be expected to realize the following formula:

$$u(t) = \sum_{l=-p}^{p} e^{i(q + nl)t} \frac{1}{(q + nl)^2}. \quad (3.52)$$

Symbolic languages like Mathematica allows us to work with this continuous function of $t$, making it easier by removing sampling from the equation. That is what we used in the implementation included in Appendix A.

3.2.4.1. Fourier series of noisy communication channels represented by regular uniform polygon error content graphs. We now investigate the numerical algorithm introduced above in more detail. In this section, we only cover the noisy communication channels whose error content graph representations correspond to regular uniform polygons.
Expanding (3.49) to draw a pentagon \((n = 5 \text{ and } q = 1,\) let us set \(p = 1\),

\[
f_5(t) = \frac{1}{(1+5(-1))^2} e^{i(1+5(-1)) \frac{t}{N}} + \frac{1}{(1+5(0))^2} e^{i(1+5(0)) \frac{t}{N}} \\
+ \frac{1}{(1+5(1))^2} e^{i(1+5(1)) \frac{t}{N}}.
\]

Similarly, for \(p = 2\);

\[
f_5(t) = \frac{e^{i(1+5(-2)) \frac{t}{N}}}{(1+5(-2))^2} + \frac{e^{i(1+5(-1)) \frac{t}{N}}}{(1+5(-1))^2} + \frac{e^{i(1+5(0)) \frac{t}{N}}}{(1+5(0))^2} \\
\times \frac{e^{i(1+5(1)) \frac{t}{N}}}{(1+5(1))^2} + \frac{e^{i(1+5(2)) \frac{t}{N}}}{(1+5(2))^2},
\]

\[
= \frac{e^{i(-9) \frac{t}{N}}}{(-9)^2} + \frac{e^{i(-4) \frac{t}{N}}}{(-4)^2} + \frac{e^{i(1) \frac{t}{N}}}{(1)^2} + \frac{e^{i(6) \frac{t}{N}}}{(6)^2} + \frac{e^{i(11) \frac{t}{N}}}{(11)^2}.
\]

One can see that number of Fourier terms increases as \(p\) increases, which also leads larger

\(l\). In other words, resolution increases as \(p\) increases.

Now, let us introduce the \(q, n, l,\) and \(N\) variables we defined above into the algorithm

and write our new equation:

\[
u_{n,q}(l,t) = \frac{e^{i(q+nl) \frac{t}{N}}}{(q+nl)^2} \tag{3.53}
\]

where \(t \in [t_0, 2\pi + t_0]\), and \(l = -p, -p + 1, \ldots, -1, 0, 1, \ldots, p - 1, p\).
The variable $l$ denotes rows of the resulting matrix. The Fourier series matrix defined by (3.53) has the following form:

$$
u_{n,q}(p,t) = \begin{bmatrix}
\frac{e^{j(q+q(0-p))}}{[q+n(0-p)]^2} \\
\frac{e^{j(q+n(1-p))}}{[q+n(1-p)]^2} \\
\frac{e^{j(q+n(2-p))}}{[q+n(2-p)]^2} \\
\vdots \\
\frac{e^{j(q+n(p-2))}}{[q+n(p-2)]^2} \\
\frac{e^{j(q+n(p-1))}}{[q+n(p-1)]^2} \\
\frac{e^{j(q+n(p))}}{[q+n(p)]^2}
\end{bmatrix}. \tag{3.54}
$$

The matrix in (3.54) has a row for each value of $l \in [-p, p]$, having a total of $2p + 1$ rows. The matrix $u_{n,q}(p,t)$ will have $N$ columns, each corresponding to an instant of time. Since the resulting Fourier series calculation is the sum of each Fourier term for every time instant $t_0$, the equation for $u_{n,q}(p,t)$ can be rewritten as

$$u_{n,q}(p,t) = \sum_{l=-p}^{p} \frac{e^{j(q+nl)}}{(q+nl)^2}. \tag{3.55}$$

Changing limits of the sum, one gets

$$u_{n,q}(p,t) = \sum_{l=0}^{2p} \frac{e^{j(q+n(l-p))}}{(q+n(l-p))^2}. \tag{3.56}$$

Fig. 26 shows star figures with $q = 2, n = 5$, for $p = 1, 2, 5$, and 10, and Fig. 27 shows star figures with $q = 4, n = 11$, for $p = 1, 2, 5$, and 10.
Example 29. Fig. 28 shows some examples to draw polygons of $u_{n,q}(p,t)$, namely $u_{6,1}(p,t)$, $u_{7,3}(p,t)$, $u_{8,2}(p,t)$, and $u_{9,3}(p,t)$. The resolution parameter is constant for all of them $(p = 10)$.

Remark 54. Note that, when $n$ and $q$ are not relatively prime, the number of vertices in the resulting polygon will not be $n$, but the sides of polygon will be drawn on top of the existing sides, as is the case for $u_{8,2}(10,t)$ and $u_{9,3}(10,t)$.

Remark 55. From Section 3.1.2, we know that every noisy communication channel has a graph representation in the form of a uniform polygon. This section has presented Fourier series representations for the regular uniform polygons. Therefore, there exists a Fourier series representation for every noisy communication channel with regular uniform polygon error content graph, and we will have further investigation of the noisy communication channels with irregular uniform polygon error content graphs.

3.2.4.2. Fourier series of noisy communication channels represented by irregular uniform polygon error content graphs. The algebraic advantage of using Fourier series to
Fig. 28. Polygon plot examples: (a) $u_{6,1}(p,t)$, (b) $u_{7,3}(p,t)$, (c) $u_{8,2}(p,t)$, (d) $u_{9,3}(p,t)$.

represent regular polygons does not suffice, since most noisy communication channels are represented by irregular polygons. Therefore, in this section, the algorithm for representing regular polygons with Fourier series will be taken one step further to represent irregular polygons with Fourier series [44].

Let $u_{n,q}(p,t)$ denote a regular polygon that has $n$ vertices and a jump speed of $q$. The time is $t \in [0,2\pi]$ such that

$$0 \leq t_0 < t_1 < \ldots < t_i < t_{i+1} < \ldots < t_{n-1} < t_n = t_0 + 2\pi$$

and the function $u_{n,q}(p,t)$ is affine linear in every interval $[t_i, t_{i+1}]$.

Similarly, let $\varphi_{n,[q]}([p],t)$ denote an irregular polygon that has $n$ vertices, where the jump speed $q$ is not constant but is a vector that has $n$ elements. Moreover, neither $[p]$ has to be constant for each side of the irregular polygon. One can assume that the resolution parameter $p$ is used to represent some parameter in the system being modeled. For example,
let \( q = [c, d, c, e, \ldots, h, q, d] \) where each distinct element of \( q \) corresponds to a different jump speed.

Length of jump vector \( q \) is \( n \), the number of vertices, revealing the fact that the irregular polygon \( \Phi_{n,[q]}([p],t) \) is composed of pieces out of the regular polygons

\[
\begin{align*}
\Phi_{n,[q]}([p],t) &= u_{n,c}([p_1],t), u_{n,d}([p_2],t), u_{n,e}([p_3],t), \ldots, u_{n,h}([p_{n-2}],t), \\
& \quad u_{n,q}([p_{n-1}],t), u_{n,d}([p_n],t).
\end{align*}
\]

Then, the irregular polygon \( \Phi_{n,[q]}([p],t) \) can be written as follows:

\[
\Phi_{n,[q]}([p],t) = \sum_{i=0}^{n-1} t_i u_{n,q_i}([p_i],t) \text{ where } t_i \leq t < t_{i+1} \quad (3.57)
\]

with \( t_0 = 0 \), and \( t_n = 2\pi \). The overall time span, \( t \in [0, 2\pi] \), is divided into \( n \) equal intervals.

The transformation \( t_i \) in (3.57) rotates the regular polygon \( u_{n,q_i}([p_i],t) \) in the counterclockwise direction by an amount of \( \frac{2\pi q_i}{n} \) and scales so that the radius of the circle on which the regular polygon was inscribed is 1. We need this scaling, since the original algorithm for regular polygons provides different radii for different jump speeds and different \( n \s\). As can be seen from (3.54) or (3.55), it is hard to identify what point on the complex plane maps to what time. The solution for this problem is to start the time from 0 (\( t = 0 \)) for all of the regular polygons, rotate them, scale them, and extract the pieces needed for constructing the irregular polygon. The rotation of regular polygons, according to the corresponding jump speed in the irregular polygon, enables using the first side of the corresponding regular polygon in constructing the irregular one. Hence, rotation eases the construction process. The resulting polygon is made up of different Fourier series concatenated one after another for specific (different) time intervals. The resolution vector \([p]\) can be set to have the same elements for the underlying regular polygons. In the following example, the resolutions \( p_i \) are set to be the same for all of the regular polygons to have a
uniform irregular polygon as a result of (3.57). However, this alone is not enough to have a uniform irregular polygon, as discussed in Example 30 below.

When dealing with numerical algorithms, not only is the resolution parameter kept constant for every regular polygon, but the time evolutions are sampled so that each of the sides in the irregular polygon will have the same number of points. Hence, it is not the regular polygons with total number of points that are the same, but regular polygons with sides having the same number of points as any other side from another polygon. For instance, in Example 30, each side has a total of 150 points.

**Remark 56.** Remark 55 suggested that there exists a Fourier series representation for every noisy communication channel with a regular uniform polygon error content graph. This section presented that the noisy communication channels with irregular uniform polygon error content graphs also have Fourier series representation, as given in (3.57).

**Example 30.** Consider the noisy communication channel $\begin{bmatrix} 2 & 7 & 5 & 8 & 1 & 4 & 6 & 3 \end{bmatrix}$ in permutation form (Section 3.1.3.). Both a Shannon representation and uniform polygon representation are shown in Fig. 29.

![Fig. 29. (a) Shannon representation of noisy communication channel $\begin{bmatrix} 2 & 7 & 5 & 8 & 1 & 4 & 6 & 3 \end{bmatrix}$ and (b) its error content graph.](image)

95
The jump speed for the line that goes from 2 to 7 is 5 \((7 - 2 = 5)\). Similarly, the speed for the line that goes from 7 to 5 is \(-2\) \((5 - 7 = -2)\). Since the counterclockwise direction was chosen to plot the polygon, the negative speed corresponds to moving clockwise with the specified amount. In order to prevent any confusion that may arise, \(\text{mod} n\) of the speed vector will be taken. The speed vector includes the jump speed from the last element of the solution vector to the first one which, for this case, would be from 3 to 2, resulting in \(-1\) \((2 - 3 = -1)\). The raw speed vector for this solution is \([5, -2, 3, -7, 3, 2, -3, 1]\). The raw speed vector \(\text{mod} n\) (in this case \(\text{mod} 8\)) provides \([5, 6, 3, 1, 3, 2, 5, 7]\), which is the vector used to rotate the regular polygons. This vector, which has all positive elements, corresponds to the jump speeds for the irregular polygon. Hence, to construct the irregular polygon shown in Fig. 29, pieces from the following regular polygons need to be extracted and put together:

\[u_{8,5}(t), u_{8,6}(t), u_{8,3}(t), u_{8,1}(t), u_{8,3}(t), u_{8,2}(t), u_{8,5}(t), u_{8,7}(t)\].

The resolution parameters \(p_i\) are all set to be 10; therefore, they have been dropped out of the regular polygons seen above. One side of each polygon is the piece needed to construct the irregular polygon.

Fig. 30 shows the desired irregular polygon constructed with Fourier series that corresponds to the noisy communication channel in Fig. 29 by concatenating the chunks corresponding to the sides of regular polygons. The curvy and smooth corners can be rendered sharp, and lines can be made straight if one increases the resolution parameter.

### 3.3. Summary

In this chapter, the noisy communication channel framework is presented. In the first section, the noisy communication channel concept as a system with a set of inputs and a set of outputs and its three representations (Shannon’s bipartite graph representation, rep-
Fig. 30. Irregular polygon \[ \begin{bmatrix} 2 & 7 & 5 & 8 & 1 & 4 & 6 & 3 \end{bmatrix} \] constructed by using Fourier series.

representation of error content graphs as uniform polygons, permutation representation, and adjacency matrix representation) is introduced. These representations are used in the second section of the chapter to relate noisy communication channels to the analysis toolsets mentioned in Chapter 2.

In the second section, the framework of relationships is presented. This framework of relationships is summarized in Fig. 31. As parts of this framework, uniform polygon error content graphs of the noisy communication channels are mapped to a) roots of unity of the complex analysis on the complex plane, b) special permutations from the combinatorial analysis that correspond to solutions of N-queens problem, c) Chebyshev properties (polynomials, matrices, and eigenvalues) of Chebyshev analysis, and d) Fourier series of Fourier analysis.

This work has been implemented in different stages either using Matlab with discrete numerical algorithms or Mathematica with continuous symbolic algorithms. In order to avoid repetition and to be able to represent the entire framework in a single output, all of the implementations have been merged into one Mathematica implementation given in Appendix A.
Fig. 31. The relationships.
CHAPTER 4

CASE STUDY: ATOMIC MODEL

The information theoretical modeling and analysis framework summarized in Fig. 31 is
introduced in Chapter 3. In this chapter, we present a case study where we approach
modeling of the atom using the notation, concepts, and relationships that are part of the
noisy communication channel framework. Like other natural systems, the atom has also
been approached either as a particle or a wave. Here, we will approach the atom from
an information theoretical perspective, representing energy levels of electrons with noisy
communication channels and Fourier series.

This chapter starts with an introduction to quantum theory, then it continues with
an introductory section about models and the behavior of the atom followed by previous
atomic models, Bohr’s atomic model, developments in quantum mechanics, and how these
developments leaded some modifications to Bohr’s atomic model. After this survey, the
topics that are used in noisy communication channel modeling, namely quantum numbers,
atomic orbitals, electron configuration and phenomenon that effect the configuration, are
included. Finally, the chapter concludes with a section describing the noisy communication
channel model that is based on Fourier series of uniform polygons and quantum numbers
and a section of summary.

Detailed information about the topics surveyed in this section can be found in chap-
ters 16, 34, 40, 41, and 42 of [2], chapter 7 of [3], and [65]. The style and notation used in
this chapter closely follows [66, 67].
4.1. Introduction to The Quantum Theory

The quantum theory starts with Planck’s quantized energy idea. While he was studying thermodynamics of solids, he discovered that the energy of radiation emitted by heated solids at various temperatures is not continuous; instead, its values are only integer multiples of a constant value. Radiation is the emission and transmission of energy through space in the form of waves [3]. Solids emit the radiation in the form of light that behaves like electromagnetic wave. In order to understand energy emission, let us give a brief information about waves, particularly electromagnetic waves.

A wave is considered as the motion of a disturbance. The wavelength ($\lambda$) of a wave is defined as the minimum distance between any two points on the wave that behave identically. The frequency ($f$) of a wave is defined as the rate at which the wave repeats itself [2]. The frequency is also said to be the inverse of the period ($T$) of a wave; that is, the time difference between two points that the wave behaves identical. According to these definitions, the velocity ($v$) of a wave is given by

$$v = \lambda f. \quad (4.58)$$

**Example 31.** Let the period of an electromagnetic wave be $125 \times 10^{-10}$ seconds. Then its frequency would be $f = 1/T = 8 \times 10^7$ Hz. Since the speed of this wave is equal to the speed of the light, using (4.58), its wavelength would be $\lambda = c/f = 3 \times 10^8 / 8 \times 10^7 = 3.75 m$.

An electromagnetic wave has electric field and magnetic field components, such that two planes that these fields lie on are mutually perpendicular. The most important property of electromagnetic waves is their speed that is the speed of light, $c$, in vacuum. According to 4.58, since the velocity is constant for electromagnetic waves, wavelength and frequency are two inversely proportional variables. The electromagnetic spectrum in terms of these two variables is shown in Fig. 32.
Let us go back to Planck’s experiments and his quantization of energy theorem. The experiments showed that the amount of a radiation energy emitted depends on the wavelength of the radiation. Considering (4.58), since the wavelength of a radiation is inversely proportional to its frequency, the amount of radiation energy emitted also depends on its frequency. In 1900, Planck suggested an explanation to discrete energy radiation, which classical physics could not explain, with an assumption: Atoms and molecules can emit (or absorb) energy only in discrete quantities, like packages. He formulated allowed emitted (or absorbed) energies by the atom with the following equation:

$$E = nhf,$$

where $h = 6.63 \times 10^{-34} J s$ is now called Planck’s constant. Planck’s theory emphasized the wave characteristic of light and showed that energy does not have to be continuous.

In 1905, Einstein used this quantized energy theory to explain the photoelectric effect. The photoelectric effect was first introduced by Maxwell, but it could not be explained classically. Experiments showed that electrons were emitted from the metal surfaces that are exposed to light of a certain minimum frequency. No matter how intense the light was sent to the surface, no electron was emitted until the frequency of light increased to a certain value that the frequency at is called threshold frequency. Einstein suggested thinking of a light beam as wave-like but rather as a stream of particles, called photons. He could be able to explain the photoelectric effect using this photon idea. Along with the wave
characteristic of the light, now we also have to deal with the particle characteristic of it –
not with one or the other, but both of them together as photon.

Then in 1919, Einstein concluded his photon idea by claiming that a photon of
energy $E$ travels in a single direction (unlike a spherical wave) and carries a momentum
equal to $E/c$. In 1922, Arthur Holly Compton and Peter Debye independently realized that
the scattering of x-ray photons from electrons could be explained by treating photons as
point-like particles with an energy of $hf$ and a momentum of $hf/c$, and by conserving both
the energy and the momentum of the photon-electron pair in a collision. This could only be
explained by Einstein’s claim. This phenomena, known as Compton effect, strengthened
the idea of the photon as a point particle with its known momentum.

4.2. Bohr’s Quantum Model of Atom

4.2.1. Introduction

All substances emit radiation when they are energized. Heating is one of many
energizing methods. The emission out of a material can be obtained by heating a sample
of it. This radiation, either continuous or discrete, is called emission spectra. Atoms in
their gas forms emit light only at specific wavelengths, so that an emission spectra of such
is called line spectra, whereas the emission spectra of the sun or a heated solid is called
continuous spectra. It was shown by experiments that the emission spectra is unique for
every element. Although this knowledge had been around, the line spectra behavior of the
atom could not be understood until Niels Bohr offered an explanation of the line spectra of
hydrogen in 1913.

Between 1860 and 1885, scientists accumulated a large amount of data about the
emission spectrums of elements. In 1885, Johann Jacob Balmer introduced a formula that
correctly predicted the wavelengths of four visible emission lines of hydrogen. His formula can be described as

\[ \frac{1}{\lambda} = R_H \left( \frac{1}{2^2} - \frac{1}{n^2} \right). \]

Here, \( R_H = 1.0974 \times 10^7 m^{-1} \) is the Rydberg constant, and \( n \) may take the values of 3, 4, 5, \ldots. Later on, this emission spectrum was explained by Bohr. This line series is suggested to be the emission lines when electrons move from a higher energy level to the first exited energy level. This series is named the Balmer series after this discovery of his. Other series of line spectra were found afterwards and named after their discoverers as Lyman, Paschen, and Brackett series. This series are shown in Fig. 33. A generalized formula of the emitted and absorbed lights of atoms, when an electron of an atom moves from \( n_i \), the initial energy level, to \( n_f \), final energy level, can be given as

\[ \frac{1}{\lambda} = R_H \left( \frac{1}{n_f^2} - \frac{1}{n_i^2} \right). \]

\( n_f \) is equal to 2 for Balmer series, 1 for Lyman series, 3 for Paschen series, and 4 for Bracket series.

Fig. 33. Line spectra series of hydrogen atom.
4.2.2. Earlier Models of Atom

Before starting Bohr’s model, let us briefly mention the earlier models of atom. The atom, at the time of Newton, was thought to be a tiny, hard, indestructible sphere. But this description could not provide any information about its electrical behavior. Around the beginning of the twentieth century, Joseph John Thompson described the atom as a positive volume including negative charges, just like the seeds of a watermelon. Since the atom is neutral, the total negative charges in the atom were said to be equal to the total positive charge of the volume. In 1911, Ernest Rutherford and his students performed an experiment by sending positively charged alpha particles to a thin metal foil. At the screen behind the foil, they observed that most of the particles passed through the foil with a slight or no scattering, as if the foil was empty. Some of the alpha particles, however, either deflected or scattered with large angles. If Thompson’s model was correct, positive alpha particles should not even touch such a large, positive volume. Rutherford, therefore, concluded that all of the positive charges (protons) are concentrated at the center of the atom in a very small volume, compared to the volume of the whole atom, and the negative charges (electrons) are in a relatively larger volume around this positive center. Then he called this positively charged center the nucleus of the atom. In order to explain why electrons are not pulled into the nucleus, he proposed a model similar to the solar system, such that protons are in the center of the atom analogous to the sun and electrons are circling around the nucleus analogous to the planets of the solar system. This model explained Rutherford’s experiment but still had some problems: It could not explain unique line spectrums of elements. Also, according to Rutherford’s model, electrons circulate around the nucleus, and this circular motion causes an acceleration because of the direction change of the velocity vector. According to the classical electromagnetic theory, accelerating particles must radiate electromagnetic waves and lose some of their energy. This energy loss should cause the radius of the electron orbit to get smaller, and eventually, the electron should fall into the nucleus. Rutherford’s model fails to explain why this does not happen either.
4.2.3. Bohr’s Model of Hydrogen Atom

In 1913, Niels Bohr postulated that classical radiation theory did not hold for atomic-sized systems to overcome the problems that Rutherford’s model could not explain. He applied Planck’s ideas of quantized energy levels to orbiting atom electrons to explain the energy loss of circulating electron. He postulated that electrons in atoms are generally confined to stable, non-radiating energy levels and orbits called stationary states. Furthermore, he applied Einstein’s concept of photon to arrive at an expression for the frequency of the light emitted when the electron jumps from one stationary state to another. The basic ideas of Bohr’s theory, as it applies to a hydrogen atom, are follows [2]:

1. The electron moves in circular orbits about the proton under the influence of the Coulomb force of attraction. This part is no different then classical theory.

2. Only certain orbits are stable. These stable orbits are ones in which the electron does not radiate. Hence the energy is fixed or not stationary, and classical mechanics may be used to describe the electrons motion.

3. Radiation is emitted by the atom when the electron ”jumps” from a more energetic initial stationary state to a lower state. This “jump” cannot be visualized or treated by classically. In particular the frequency $f$ of the photon emitted in the jump is independent of the frequency of the electron’s orbital motion. The frequency of the light emitted is related to the change in the atom’s energy and is given by the Planck-Einstein formula

$$E_i - E_f = hf,$$

where $E_i$ is the energy of the initial state, $E_f$ is the energy of the final state, and $E_i > E_f$.

4. The size of the allowed electron orbits is determined by an additional quantum condition imposed on the electron’s orbital angular momentum. Namely, the allowed orbits
are those for which the electron’s orbital angular momentum about the nucleus is an integral multiple of $\hbar = h/2\pi$,

$$mvr = n\hbar \quad n = 1, 2, 3, \ldots$$

Using these assumptions, the radius of the stationary orbits, the allowed energies, and the emission wavelengths can be calculated with the following equations, respectively:

$$r_n = \frac{n^2 \hbar^2}{mke^2}$$

$$E_n = -\frac{ke^2}{2a_0} \left( \frac{1}{n^2} \right)$$

$$\frac{1}{\lambda} = \frac{f}{c} = \frac{ke^2}{2a_0 hc} \left( \frac{1}{n_i^2} - \frac{1}{n_j^2} \right).$$

Here $e$ is the charge of an electron, and $a_0$ is the radius of the first orbit for $n = 1$, the so-called Bohr radius. The wavelength equation is very close to the Balmer’s experimental equation with Rydberg constant.

4.2.4. Quantum Mechanics and Modifications to Bohr’s Model

Bohr could explain emission spectra; however, he could not explain his postulate saying that electrons are restricted to orbit around the nucleus only at certain distances. In 1924, Louis de Broglie gave a logical explanation to this restriction by thinking that since light waves can behave like a stream of particles, or photons, then particles may behave like waves as well. He suggested a wave model for small particles, including electrons of the atom. According to his model, electrons in an atom behave like waves. These waves are called standing waves because they do not travel along their axis. That is, the points at
which the amplitude of the wave is zero, the nodes, do not move at all. There are two nodes at each end of a standing wave, and there may be nodes between these two.

De Broglie claimed that the wavelength of the electron must exactly fit the circumference of its orbit in the atom, otherwise the wave would eventually cancel itself while its orbital motion. Thus the relation between orbital circumference and wavelength of electron is given by

$$2\pi r = n\lambda,$$

where \( r \) is orbital radius, and \( n = 1, 2, 3, \ldots \). De Broglie concluded that waves can behave like particles and particles can exhibit some wave properties with the relation of

$$\lambda = \frac{h}{m\nu},$$

where \( h \) is the Planck constant, \( m \) is mass of the particle, \( \lambda \) is its wavelength, and \( \nu \) is its velocity.

With the discovery of the wave properties of the electron, the question of how to locate a wave so that we can specify the position of electron in atom arose. In 1927, Werner Heisenberg introduced his uncertainty principle as follows [2]: If a measurement of position is made with precision \( \Delta x \), and simultaneous measurement of momentum is made with precision \( \Delta p \), then the product of two uncertainties can never be smaller than a number of order of \( \hbar \). That is,

$$\Delta x \Delta p \geq \hbar.$$

Thus, it is impossible to measure the exact position and momentum of a particle simultaneously. If we apply this principle to the electron of a hydrogen atom, we see that electron does not orbit the nucleus in a specific path, as it was stated in Bohr’s model, because if it
did, we could determine both the position and the momentum (from its kinetic energy) of the electron simultaneously, which is contrary to Heisenberg uncertainty principle.

After Heisenberg’s discovery, defining an electron’s position in an atom is even more difficult. Apparently it is impossible to locate the electron precisely. In 1928, Erwin Schrödinger proposed a wave equation that described the manner in which matter waves change in space and time. The waves are represented by their physical quantities that vary with time and position. For example, sound waves are represented by pressure variation, and waves on string are represented by displacement. In the case of matter waves, the wave function, $\psi$, represents the wave’s behavior. Schrödinger’s equation incorporates both particle behavior, in terms of mass, and wave behavior, in terms of wave function, that depends on the location of the particle (i.e., electron) in the space of the system (i.e., atom). Wave function itself does not have a physical meaning. However, the square of the wave function is related to the probability of finding the particle in a certain part of the space.

In order to understand the atomic composition, let us use Bohr’s model with the modifications for the simplest atom (with one proton and one electron), hydrogen.

4.2.5. The Hydrogen Atom

One of the situations that Bohr’s original model could not explain was the existence of closely spaced lines in emission spectra. The number of these closely spaced lines was even increased when the atoms were placed in a strong magnetic field. After the developments in quantum mechanics this phenomena was explained by using quantum numbers which we will discuss shortly. The solution of the Schrödinger equation for the hydrogen atom under the electrostatic potential energy $U(r) = -\frac{k e^2}{r}$, where $k$ is Coulomb constant, and $r$ is the radial distance from the proton to the electron, yields two pieces of information. It specifies the possible energy states that the electron can occupy and identifies the corresponding wave functions of the electron associated with each energy state. These energy
states and wave functions are characterized by a set of quantum numbers. According to the quantum mechanics, the energies of the allowed states for the hydrogen atom are:

\[ E_n = -\left(\frac{ke^2}{2a_0}\right)\frac{1}{n^2} \quad n = 1, 2, 3, \ldots \]  

(4.59)

By solving Schrödinger’s equation, we obtain energy levels and corresponding wave functions with which we can calculate the probability of finding the electron in a certain region (|\psi|^2) from. Thus, using these energy levels and the ability of calculating these probabilities, we can construct a comprehensive view of the hydrogen atom. For the atoms that have more than one electrons, however, solving the Schrodinger equation becomes a complex problem, even for the helium atom, which has only two electrons. Instead of solving Schrodinger equation, another approach, using the energies and the wave functions obtained for the hydrogen atom, is taken for many-electron atoms.

To distinguish the quantum mechanical description from Bohr’s model, we replace the term orbit with the term orbital. An orbital can be thought of as the wave function (\(\psi\)) of an electron in an atom. The square of the wave function, |\psi|^2, defines the distribution of electron density in space around nucleus [3].

4.2.5.1. Quantum numbers. In order to solve Schrodinger’s equation for a three-dimensional hydrogen atom problem, three quantum numbers, corresponding to three degrees of freedom for the electron, are needed. These three and also spin quantum numbers are shown in Table 5.

All states with the same principle quantum number are said to form a shell. These shells are represented by the letters K, L, M, N, … for n = 1, 2, 3, 4, …. Likewise, the states having the same values of n and l are said to form a subshell and represented by the letters s, p, d, f, g, h, … for \(\ell = 1, 2, 3, 4, 5, 6, \ldots\).

Example 32. The quantum numbers of the possible states in the hydrogen atom corresponding to the principal quantum number n = 2 are given in Table 6.
TABLE 5
Quantum Numbers

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Name</th>
<th>Values</th>
<th>Number of States</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>Principal quantum number</td>
<td>Any number</td>
<td>Any number</td>
</tr>
<tr>
<td>$\ell$</td>
<td>Orbital quantum number</td>
<td>$0 - (n - 1)$</td>
<td>$n$</td>
</tr>
<tr>
<td>$m_{\ell}$</td>
<td>Orbital magnetic quantum number</td>
<td>$(-\ell) - \ell$</td>
<td>$2\ell + 1$</td>
</tr>
<tr>
<td>$m_s$</td>
<td>Spin magnetic quantum number</td>
<td>$1/2$ or $-1/2$</td>
<td>$2$</td>
</tr>
</tbody>
</table>

TABLE 6
All Possible Quantum States for the Hydrogen Atom of $n = 2$

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\ell$</th>
<th>$m_{\ell}$</th>
<th>$m_s$</th>
<th>Subshell</th>
<th>Shell</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1/2</td>
<td>2s</td>
<td>L</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>-1/2</td>
<td>2s</td>
<td>L</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1/2</td>
<td>2p</td>
<td>L</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>-1/2</td>
<td>2p</td>
<td>L</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1/2</td>
<td>2p</td>
<td>L</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>-1/2</td>
<td>2p</td>
<td>L</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>-1</td>
<td>1/2</td>
<td>2p</td>
<td>L</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>-1</td>
<td>-1/2</td>
<td>2p</td>
<td>L</td>
</tr>
</tbody>
</table>

Let us now briefly mention the physics behind these quantum numbers. As we have seen in energy equation 4.59, the principle quantum number, $n$, is a significant factor in describing the energy levels of an atom.

The orbital quantum number, $\ell$, is seen in the discrete orbital angular momentum formula. The classical angular momentum formula of a particle whose mass is $m$ and velocity is $v$, while it moves in a circle of radius $r$, is given by $L = mvr$. The direction of
angular momentum vector is perpendicular to the plane of the circle. According to classical physics, the angular momentum may have any value, whereas it is restricted to multiples of $\hbar$ according to Bohr’s model. Therefore, $mvr = n\hbar$. Considering the electron as a particle, according to this equation, the electron of a hydrogen atom would have one unit of angular momentum, which is not correct. On the other hand, if angular momentum is said to be zero in Bohr’s model, then it would mean that the electron is oscillating through nucleus which is not acceptable. Therefore, Bohr’s model needs to be modified. This modification is made by considering the wave behavior of atoms. According to quantum mechanics, an atom in a state characterized by the principal quantum number, $n$, can take only the following discrete values of orbital angular momentum:

$$L = \sqrt{\ell(\ell + 1)}\hbar \quad \ell = 0, 1, 2, \ldots, n-1$$

We see that the angular momentum is 0 for $\ell = 0$. This cannot be described in a purely particle-like model. In quantum mechanics interpretation, the electron cloud for the $\ell = 0$ state is spherically symmetric and has no fundamental axis of rotation. Therefore the angular momentum is zero.

We have seen that orbital angular momentum is quantized and determined by the orbital quantum number, $\ell$. Since angular momentum is a vector, to be able to define it, its direction is also needed. An orbiting electron can be considered as a current loop with a corresponding magnetic moment. In the case of existence of a magnetic field, such magnetic moment would interact with it. Let there be a weak external magnetic field, say in $z$ direction, effecting the atom. According to quantum mechanics, the angular momentum component in $z$ direction is quantized by the orbital magnetic quantum number, as seen in the following equation:

$$L_z = m_\ell \hbar \quad m_\ell = -\ell, \ldots, 0, \ldots, \ell$$
The fact that the direction of \( \vec{L} \) is quantized with respect to an external magnetic field is often referred to as space quantization.

**Example 33.** Table 7 shows the orbital angular momentum values and their corresponding components in the direction of external magnetic field for an atom of \( n = 3 \):

<table>
<thead>
<tr>
<th>( \ell )</th>
<th>( L )</th>
<th>( L^2 )</th>
<th>( m_\ell )</th>
<th>( L_z )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>( \sqrt{2} \hbar )</td>
<td>( 2\hbar^2 )</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>( \sqrt{6} \hbar )</td>
<td>( 6\hbar^2 )</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>( \hbar )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>( 2\hbar )</td>
</tr>
</tbody>
</table>

A vector model describing space quantization for \( \ell = 2 \) is given in Fig. 34a. Note that \( z \) direction is chosen to be the direction of the external magnetic field. Also note that the \( x \) and \( y \) components of the angular momentum vector are not certain. In Fig. 34b, a three-dimensional view is shown. In this figure, \( \vec{L} \) must lie on the surface of a cone that makes an angle, \( \theta \), with the direction of the external magnetic field. The angle, \( \theta \), can be expressed as

\[
\cos \theta = \frac{L_z}{L} = \frac{m_\ell}{\sqrt{\ell(\ell+1)}}.
\]

Since \( m_\ell < \sqrt{\ell(\ell+1)} \), \( \theta \) cannot be zero, and \( L \) has to be always greater than \( L_z \). Therefore, \( \vec{L} \) cannot point in a single direction but rather traces out a cone in space. This means
that the angular momentum vector cannot be determined with certainty, which is already expected, because of the uncertainty principle of Heisenberg.

Fig. 34. Angular momentum vector. (a) The allowed projections of the orbital angular momentum for the case of \( \ell = 2 \), (b) the orbital angular momentum vector lies on the surface of a cone and precesses about the z axis when a magnetic field is applied in this direction [2].

The spin magnetic quantum number, \( m_s \), was first introduced in 1925 by Samuel Goudsmidt and George Uhlenbeck, following a suggestion made by Wolfgang Pauli, to resolve the dilemma of two very closely spaced lines in the line spectra of certain gases, referred as fine structure. They proposed this new set of quantum numbers in describing a quantum state to explain the existence of the fine structure. In order to describe the spin quantum number, the spin behavior of the particles has to be described. Although it is not correct, it is convenient to think of an electron spinning on its axis while it orbits the nucleus. An electron can spin on its axis in two directions, as shown in Fig. 35: Spin up direction, as in Fig. 35a, or spin down direction, as in Fig. 35b.

There is a slight energy difference that causes the fine structure between two electrons, each with one of these two spin directions. The quantum numbers associated with
Fig. 35. Spin of an electron: (a) spin-up direction, (b) spin-down direction.

The spin of electrons are \( m_s = \frac{1}{2} \) for the spin-up state and \( m_s = -\frac{1}{2} \) for the spin-down state. The electron cannot be considered spinning since, according to quantum mechanics, it cannot be precisely located. This fact deduces the incorrectness of an electron’s spin behavior; however, all experimental evidence supports the fact that an electron does have some intrinsic property that can be described by the spin magnetic quantum number.

The fine structure is attributed to the existence of the spin magnetic moment of the electron. In 1921, Stern and Gerlach performed an experiment and demonstrated the space quantization. In the experiment, a beam of neutral silver atoms was sent through a nonuniform magnetic field. The same process was repeated for different atoms. The classical expectation would be a continuously spread-out deflected beam. In classical interpretation, the net magnetic force on the atom is along the direction of the magnetic field, let it be \( z \) direction, and is proportional to the magnetic moment, \( \mu_z \). According to quantum mechanics, however, the deflected beam would have several components that define various possible values of \( \mu_z \). With their experiments, Stern and Gerlach were able to shave split beams instead of continuously spread-out beams. Hence, the space quantization was at least quantitatively verified. If we assume that \( \mu_z \) is due to the orbital angular momentum and proportional to \( m_\ell \), then the number of possible \( \mu_z \) values should have been \( 2\ell + 1 \).
Actually, the Stern-Gerlach experiment showed that the beam always splits up only two components. In 1927, Phipps and Taylor repeated the Stern-Gerlach experiment using a beam of hydrogen atoms. It was important because a hydrogen atom has only one electron in its ground level, thus both $\ell$ and $m_\ell$ are zero. If $\mu_z$ was due to the orbital angular momentum, then it was not expected for the beam to be deflected; however, the beam split into two components again, so this deflection must be due to something else.

In 1925, Goudsmit and Uhlenbeck proposed that the electron has an intrinsic angular momentum other than its orbital angular momentum. Classically, this intrinsic angular momentum is attributed the electron’s spinning about its own axis and hence is called electron spin. Therefore, the total angular momentum of an electron contains an orbital component $\vec{L}$ and a spin component $\vec{S}$. In 1929, Dirac solved the relativistic wave equation of an electron in a potential well using a relativistic form of total energy that confirmed the spin idea. He showed that the electron spin could be described by a single quantum number $s = 1/2$. The magnitude of spin angular momentum $\vec{S}$ for the electron is then given by

$$S = \sqrt{s(s+1)}\hbar = \frac{\sqrt{3}}{2}\hbar. \quad (4.61)$$

As seen in the equation, spin angular momentum is also quantized. The $z$ component of the spin angular momentum is specified by spin magnetic quantum number, $m_s$, and given by

$$S_z = m_s\hbar = \pm \frac{1}{2}\hbar.$$ 

The spin magnetic moment of the electron due to its spin angular momentum, $\vec{S}$, and its $z$ component are given by

$$\mu_s = -\frac{e}{m}S, \quad \mu_{sz} = \pm \frac{e\hbar}{2m},$$

respectively.
We have completed the quantum numbers and what they mean in physics. Now, we will explore the wave functions of the hydrogen atom.

4.2.5.2. The wave functions for the hydrogen atom. In this section, the wave functions (depending on the state that the electron is in) for the hydrogen atom and the probability functions for its first few electron states are described. The wave functions are acquired from Schrodinger equation and the probability functions are obtained from these wave functions, as mentioned before.

Since the potential energy of the hydrogen atom is depended only on the radial distance, \( r \), we would expect that the wave function would also mainly depend on this distance; especially when \( \ell \) is zero and there is no additional effect other than the potential. This indeed is the case for the \( s \) states. The wave function for the \( 1s \) state is given by

\[
\psi_{1s}(r) = \frac{1}{\sqrt{\pi a_0^5}} e^{-\frac{r}{a_0}},
\]

where \( a_0 \) is the Bohr radius:

\[
a_0 = \frac{\hbar^2}{\mkappa e^2} = 0.0529 \text{nm}.
\]

Here \( k \) is the Coulomb constant. The wave function depends on the radial distance, and it is spherically symmetric as expected. Using the absolute square of the wave function, the radial probability density function, as the probability of finding the electron in a spherical shell of radius \( r \), for the hydrogen atom in its ground state is given by

\[
P_{1s}(r) = 4\pi r^2 |\psi_{1s}|^2 = \left( \frac{4r^2}{a_0^3} \right) e^{-\frac{2r}{a_0}}.
\]

The details of calculation can be found in [2]. The probability function can be seen in Fig. 36.
Fig. 36. The probability function of hydrogen atom in the ground state.

The next state for the hydrogen atom is 2s state. The wave function for the electron of a hydrogen atom in 2s state is given by

\[
\psi_{2s}(r) = \frac{1}{4\sqrt{2\pi}} \left( \frac{1}{a_0} \right)^{3/2} \left[ 2 - \frac{r}{a_0} \right] e^{-r/2a_0}.
\]

\(\psi_{2s}\) also depends only on \(r\) and is spherically symmetric. Radial distribution plots of this state and some other states are shown in Fig. 37.

Fig. 37. The radial probability density functions for several states of the hydrogen atom.

The wave functions corresponding to the states other than \(s\) states are not spherically symmetric. The three wave functions corresponding to the states for which \(n = 2\) (\(m_l = 1, 0, -1\)) can be expressed as appropriate linear combination of the three \(p\) states. Each state
has a wave function that has distinct directional characteristics, indicated by the notations \( p_x \), \( p_y \), and \( p_z \). The electron clouds for these three states have identical structures but differ in their orientation with respect to the \( x \), \( y \), and \( z \) axes. The non-spherical wave functions for these states are:

\[
\begin{align*}
\psi_{2p_x} &= xF(r) \\
\psi_{2p_y} &= yF(r) \\
\psi_{2p_z} &= zF(r)
\end{align*}
\]

where \( F(r) \) is some exponential function of \( r \). Wave functions with a highly directional character, such as these three, play an important role in chemical bonding, the formation of molecules, and chemical properties [2].

4.2.6. Atomic Orbitals

One of the things that Bohr’s original model could not explain was the existence of closely spaced lines in emission spectra. The number of these closely spaced lines increased even when the atoms placed in a strong magnetic field. After the developments in quantum mechanics, this situation was explained using quantum numbers that we have described in the previous section. In this section, how these quantum numbers relates to the atomic orbitals in hydrogen and many-electron atoms is described. The energy levels of these orbitals are also described.

Table 8, which can be considered as an extension of Table 6 with atomic orbital designations, shows the relation between quantum numbers and atomic orbitals. There is actually nothing new about this table. We have already discussed the names of these orbitals, the numbers of orbitals corresponding to the quantum numbers, and the number of electrons in those orbitals. In this section, however, we will briefly describe s, p, d, and higher energy orbitals from a chemistry point of view.
TABLE 8

The Relation Between Quantum Numbers and Atomic Orbitals [3].

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\ell$</th>
<th>$m_\ell$</th>
<th>Number of orbitals $(2\ell + 1)$</th>
<th>Number of electrons $(2 \times)$</th>
<th>Atomic orbital designations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>$1s$</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>$2s$</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>-1, 0, 1</td>
<td>3</td>
<td>6</td>
<td>$2p_x, 2p_y, 2p_z$</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>$3s$</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>-1, 0, 1</td>
<td>3</td>
<td>6</td>
<td>$3p_x, 3p_y, 3p_z$</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>-2, -1, 0, 1</td>
<td>5</td>
<td>10</td>
<td>$3d_{xy}, 3d_{yz}, 3d_{xz}$, $3d_{x^2-y^2}, 3d_{z^2}$</td>
</tr>
</tbody>
</table>

From the chemistry point of view, the most important thing that we are trying to find out by studying atomic orbitals is to get as much information as possible about their shapes, so that we can describe the formation of chemical bonds between atoms. Since the wave function characterizes the orbital from nucleus to infinity, the shape of an orbital cannot be determined precisely. However, it can be determined with a high probability. There is about a 90% probability of finding the electron within a sphere of radius 100 pm surrounding nucleus. Thus orbitals can be represented by boundary surface diagrams that enclose about 90% of the total electron density in an orbital. These diagrams, for the first few orbitals, are given in Fig. 38.

The eigenfunctions that are used to draw boundary surface diagrams are related to the spherical harmonics of the eigenvalue problem for angular momentum. The spherical
harmonics of this problem are obtained by solving the problem in spherical coordinates [68]:

$$Y_{m}^{\ell}(\theta, \varphi) = \sqrt{\frac{2\ell + 1}{4\pi}} \frac{(\ell - m)!}{(\ell + m)!} (-1)^m \cdot e^{im\varphi} P_{\ell}^{m}(\cos \theta) ,$$

where $P_{\ell}^{m}(\cos \theta)$ are Legendre functions and given by

$$P_{\ell}^{m}(\cos \theta) = \frac{(-1)^m (\ell + m)!}{2\ell \cdot \ell!} \frac{(\ell - m)!}{(\ell - m)!} (1 - \cos^2 \theta)^{-m/2} \frac{d^{\ell-m}}{d(\cos \theta)^{\ell-m}} (\cos^2 \theta - 1)^\ell .$$

Details of the solution can be found in [68]. The probability distributions of the first few spherical harmonics, $|Y_{m}^{\ell}|^2$, are given in Fig. 39. As we mentioned before, the boundary surface diagrams are drawn based on these spherical harmonics. The relation describing orbital eigenfunctions using spherical harmonics is given as follows

$$S_{\ell}^{m}(\theta, \varphi) = Y_{\ell}^{m}(\theta, \varphi) \quad , \quad m_{\ell} = 0$$

$$S_{\ell}^{m}(\theta, \varphi) = \frac{1}{\sqrt{2}} \left[ \text{sign}(m_{\ell}) Y_{\ell}^{m}(\theta, \varphi) + Y_{\ell}^{-m}(\theta, \varphi) \right] \quad , \quad m_{\ell} \neq 0$$
Fig. 39. The probability distributions of first few spherical harmonics, $|Y_{\ell m}|^2$.

### 4.2.6.1. s orbitals.

The spherical harmonic and the eigenfunction corresponding to s orbitals is given by

$$S_0^0 = Y_0^0 = \frac{1}{\sqrt{4\pi}}.$$  

As seen in Fig. 38a, boundary surface diagrams of the s orbitals are merely in sphere shapes. The size of the sphere increases with the principle quantum number. Although the exact location of the electron in an atom is not known, the information given by these diagrams are sufficient to determine their chemical bonding.

### 4.2.6.2. p orbitals.

As mentioned earlier, p orbitals correspond to the orbital quantum number $\ell = 1$. The values that orbital magnetic quantum number, $m_\ell$, can have for $\ell = 1$ are -1, 0, and 1. Therefore we have three p orbitals, namely, $p_x$, $p_z$, and $p_y$. The boundary surface diagrams for these three p orbitals for $n = 2$ can be seen in Fig. 38b.
The spherical harmonics and eigenfunctions corresponding to \( p \) orbitals are given by

\[
S_0^0 = Y_0^0 = \sqrt{\frac{3}{4\pi}} \cos \theta = \sqrt{\frac{3}{4\pi}} \frac{\hat{z}}{r},
\]
\[
Y_1^\pm = \mp \sqrt{\frac{3}{8\pi}} e^{\pm i\phi} \sin \theta = \mp \sqrt{\frac{3}{8\pi}} \frac{x \pm iy}{r},
\]
\[
S_1^1 = \frac{1}{2} \sqrt{\frac{3}{\pi}} \sin \theta \sin \varphi \quad S_1^{-1} = \frac{1}{2} \sqrt{\frac{3}{\pi}} \sin \theta \cos \varphi.
\]

As seen in the figure, the orbital can be thought as two lobes where nucleus is at the center. \( p \) orbitals are not spherically symmetric like \( s \) orbitals were. These three \( p \) orbitals are identical in size, shape, and energy. They differ only in their orientations. Just like \( s \) orbitals, \( p \) orbitals also increase in size with the principle quantum number, \( n \).

4.2.6.3. \( d \) orbitals and other higher-energy level orbitals. \( d \) orbitals are the ones that correspond to the orbital quantum number, \( \ell = 2 \). They consist of five orbitals, since there are five \( m_\ell \) values for \( \ell = 2 \). The boundary surface diagrams for these five orbitals, \( d_{xy}, d_{xz}, d_{yz}, \) and \( d_{x^2-y^2} \), for \( n = 3 \) can be seen in Fig. 38c.

The spherical harmonics and eigenfunctions corresponding to \( d \) orbitals are given by

\[
S_0^2 = Y_2^0 = \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1) = \sqrt{\frac{5}{16\pi}} \frac{2z^2-x^2-y^2}{r},
\]
\[
Y_2^\pm = \mp \sqrt{\frac{15}{8\pi}} e^{\pm i\phi} \cos \theta \sin \theta = \mp \sqrt{\frac{15}{8\pi}} \frac{(x \pm iy)^z}{r^2},
\]
\[
S_2^1 = \frac{1}{2} \sqrt{\frac{15}{\pi}} \cos \theta \sin \theta \sin \varphi \quad S_2^{-1} = \frac{1}{2} \sqrt{\frac{15}{\pi}} \cos \theta \sin \theta \cos \varphi,
\]
\[
Y_2^{\pm 2} = \sqrt{\frac{15}{32\pi}} e^{\pm 2i\phi} \sin^2 \theta = \sqrt{\frac{15}{32\pi}} \frac{(x \pm iy)^2}{r^2},
\]
\[
S_2^2 = \frac{1}{4} \sqrt{\frac{15}{\pi}} \sin^2 \theta \cos 2\varphi \quad S_2^{-2} = \frac{1}{4} \sqrt{\frac{15}{\pi}} \sin^2 \theta \sin 2\varphi.
\]

All of these \( 3d \) orbitals in an atom shown here have identical orbital energies. Other \( d \) orbitals with higher-energy, such as \( 4d, 5d, \ldots \) have similar shapes but different size.
The higher energy level orbitals are $f$, $g$, ... orbitals. $f$ orbitals correspond to $l = 3$ and consist of seven orbitals with the capacity of 14 electrons. These orbitals are used by the elements with atomic numbers greater than 57.

4.2.6.4. Orbital energies. According to the original version of Bohr’s atomic model for a hydrogen atom, the energies of the orbitals depend only on the principle quantum number, $n$. Therefore the energy levels for the orbitals of a same shell are equal, although they have different electron distribution. The orbital energy levels for the hydrogen atom are shown in Fig. 40a. Different orbitals with the same energies are called degenerate orbitals. 1s orbital is the most stable condition for hydrogen’s one electron and is called the ground state. The other orbitals with higher orbitals are called excited states.

We said that all of the orbital energies in a shell are equal for the hydrogen atom. For multi-electron atoms, this cannot be the case, because each electron has to be at a different energy level (because of the Pauli exclusion principle that we discuss in the next section.) According to the central field approximation model, in which the spin-spin interactions and some relativistic effects are not taken into account, orbital energies in many-electron atoms do not depend only on the principle quantum number but also the orbitals quantum number [65]. This dependency is caused by the role of electron-electron repulsion energies in the total energy. The total energy is not dependent on the orbital magnetic quantum number, since the approximation assumes the potential to be spherically symmetric. The orbital energy levels according to the central field approximation for many-electron atoms is shown in Fig. 40b. As can be seen in the figure, some orbitals in a subshell are still degenerate. This degeneracy will be removed in the next section. The orbital energies for many-electron atoms are given in ascending order in Fig. 40c. The orbitals according to their energies in ascending order can be written as: 1s, 2s, 2p, 3s, 3p, 4s, 3d, 4p, 5s, 4f, 5d, 6p, .... Some energy levels, for example 4s and 3d, are very close and the order may differ from an atom to another. These orbitals are marked in Fig. 40c.
Fig. 40. (a) Orbital energy levels of the hydrogen atom, (b) orbital energy levels for the many-electron atoms, (c) orbital energies in ascending order.
Since electrons are more stable in the lower energy level orbitals, the orbital energy levels also show how electrons fill the orbitals in many-electron atoms. In the next section, the electron configuration in many-electron atoms is described.

4.2.7. Electron Configuration

In this section, how the electrons of known elements are distributed among the various atomic orbitals is described. This configuration is done by using quantum numbers, since they enable us to label electrons in the atom. For the sake of simplicity, we shall only consider the ground state of neutral atoms.

Some rules, effects, and principles, such as Pauli exclusion principle, Hund’s rule, building-up (aufbau) principle, shielding effect, and paramagnetism-diamagnetism concepts, that lead to construction of the periodic table are discussed following the notation used in this section.

4.2.7.1. Notations. The electron configuration can be represented in the usual spectroscopic notation, that is, the value of $n$ followed by the corresponding orbital of $l$ value ($s$ for $l = 0$, $p$ for $l = 1$, and so on) and its superscript showing the number of electrons on that particular orbital. For example, $1s^2$, $3p^4$, $5d^8$, ....

It can also be represented by an orbital diagram that shows the spin of the electrons as well:

\[
\begin{array}{c}
\uparrow \downarrow \\
1s^2
\end{array} \quad \begin{array}{c}
\uparrow \downarrow \uparrow \uparrow \\
3p^4
\end{array} \quad \begin{array}{c}
\uparrow \downarrow \uparrow \uparrow \uparrow \uparrow \uparrow \uparrow \\
5d^8
\end{array}
\]

where each box represents an atomic orbital, and the arrows denote two possible spinning motions of the electron.
The third notation we will demonstrate is called terms in spectroscopic language and denotes atomic energy levels. The terms are shown in the Russell-Saunders notation:

\[ 2S + 1 L_J. \]

Here \( S (= s) \) is the spin quantum number, which is the spin angular momentum in \( \hbar \) units, as in (4.61). \( J (= j) \) is the total angular momentum number, which is the total angular momentum given by the following equation in \( \hbar \) units:

\[ \vec{J} = \vec{L} + \vec{S}, \]

where \( \vec{L} \) and \( \vec{S} \) are orbital and spin angular momentums given in (4.60) and (4.61). And the code letters \( S (= s), P (= p), D (= d), \ldots \) correspond to the values of \( L (= \ell) = 0, 1, 2, \ldots \) as mentioned before.

4.2.7.2. The “aufbau” building-up. The \( Z \) electrons of an atom of atomic number \( Z \) occupy the lowest individual energy levels, in accordance with the Pauli exclusion principle that we will discuss shortly. The ordering of individual levels is displayed in Fig. 40c. The ground state configuration of an atom is obtained by distributing its \( Z \) electrons among a certain number of subshells in a way that all subshells but the last one – corresponding to the highest energy – are filled, and the last subshell has some empty orbitals, except for particular values of \( Z \) (2, 4, 10, 12, etc.). The electrons in this last subshell are the least tightly bound ones and called valence electrons. The aufbau principle then can be described as going from one atom with atomic number \( Z \) to the next one, with atomic number \( Z + 1 \), the \( (Z+1) \) electrons occupying the lowest energy levels allowed by the exclusion principle. In this way, the subshells are progressively filled, as seen in Table 9 [65]. Note that configurations and terms in parentheses are estimated.
# TABLE 9

Electron Configuration of the Atoms in Their Ground State

<table>
<thead>
<tr>
<th>Z</th>
<th>Symbol</th>
<th>Element</th>
<th>Configuration</th>
<th>Term</th>
<th>Ionization potential (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>H</td>
<td>hydrogen</td>
<td>1s</td>
<td>$^2S_{1/2}$</td>
<td>13.60</td>
</tr>
<tr>
<td>2</td>
<td>He</td>
<td>helium</td>
<td>1s²</td>
<td>$^1S_0$</td>
<td>24.59</td>
</tr>
<tr>
<td>3</td>
<td>Li</td>
<td>lithium</td>
<td>[He] 2s</td>
<td>$^2S_{1/2}$</td>
<td>5.39</td>
</tr>
<tr>
<td>4</td>
<td>Be</td>
<td>beryllium</td>
<td>[He] 2s²</td>
<td>$^1S_0$</td>
<td>9.32</td>
</tr>
<tr>
<td>5</td>
<td>B</td>
<td>boron</td>
<td>[He] 2s²2p</td>
<td>$^2P_{1/2}$</td>
<td>8.30</td>
</tr>
<tr>
<td>6</td>
<td>C</td>
<td>carbon</td>
<td>[He] 2s²2p²</td>
<td>$^3P_0$</td>
<td>11.26</td>
</tr>
<tr>
<td>7</td>
<td>N</td>
<td>nitrogen</td>
<td>[He] 2s²²p³</td>
<td>$^4S_{3/2}$</td>
<td>14.53</td>
</tr>
<tr>
<td>8</td>
<td>O</td>
<td>oxygen</td>
<td>[He] 2s²²p⁴</td>
<td>$^3P_2$</td>
<td>13.62</td>
</tr>
<tr>
<td>9</td>
<td>Fe</td>
<td>fluorine</td>
<td>[He] 2s²²p⁵</td>
<td>$^2P_{3/2}$</td>
<td>17.42</td>
</tr>
<tr>
<td>10</td>
<td>Ne</td>
<td>neon</td>
<td>[He] 2s²²p⁶</td>
<td>$^1S_0$</td>
<td>21.56</td>
</tr>
<tr>
<td>11</td>
<td>Na</td>
<td>sodium</td>
<td>[Ne] 3s</td>
<td>$^2S_{1/2}$</td>
<td>5.14</td>
</tr>
<tr>
<td>12</td>
<td>Mg</td>
<td>magnesium</td>
<td>[Ne] 3s²</td>
<td>$^1S_0$</td>
<td>7.65</td>
</tr>
<tr>
<td>13</td>
<td>Al</td>
<td>aluminum</td>
<td>[Ne] 3s²³p</td>
<td>$^2P_{1/2}$</td>
<td>5.99</td>
</tr>
<tr>
<td>14</td>
<td>Si</td>
<td>silicon</td>
<td>[Ne] 3s²³p²</td>
<td>$^3P_0$</td>
<td>8.15</td>
</tr>
<tr>
<td>15</td>
<td>P</td>
<td>phosphorus</td>
<td>[Ne] 3s²³p³</td>
<td>$^4S_{3/2}$</td>
<td>10.49</td>
</tr>
<tr>
<td>16</td>
<td>S</td>
<td>sulphur</td>
<td>[Ne] 3s²³p⁴</td>
<td>$^3P_2$</td>
<td>10.36</td>
</tr>
<tr>
<td>17</td>
<td>Cl</td>
<td>chlorine</td>
<td>[Ne] 3s²³p⁵</td>
<td>$^2P_{3/2}$</td>
<td>12.97</td>
</tr>
<tr>
<td>18</td>
<td>Ar</td>
<td>argon</td>
<td>[Ne] 3s²³p⁶</td>
<td>$^1S_0$</td>
<td>15.76</td>
</tr>
<tr>
<td>19</td>
<td>K</td>
<td>potassium</td>
<td>[Ar] 4s</td>
<td>$^2S_{1/2}$</td>
<td>4.34</td>
</tr>
<tr>
<td>20</td>
<td>Ca</td>
<td>calcium</td>
<td>[Ar] 4s²</td>
<td>$^1S_0$</td>
<td>6.11</td>
</tr>
<tr>
<td>21</td>
<td>Sc</td>
<td>scandium</td>
<td>[Ar] 4s²³d</td>
<td>$^2D_{3/2}$</td>
<td>6.54</td>
</tr>
<tr>
<td>22</td>
<td>Ti</td>
<td>titanium</td>
<td>[Ar] 4s²³d²</td>
<td>$^3F_2$</td>
<td>6.82</td>
</tr>
<tr>
<td>23</td>
<td>V</td>
<td>vanadium</td>
<td>[Ar] 4s²³d³</td>
<td>$^4F_{3/2}$</td>
<td>6.74</td>
</tr>
<tr>
<td>24</td>
<td>Cr</td>
<td>chromium</td>
<td>[Ar] 4s³d⁵</td>
<td>$^7S_3$</td>
<td>6.77</td>
</tr>
<tr>
<td>25</td>
<td>Mn</td>
<td>manganese</td>
<td>[Ar] 4s²³d⁵</td>
<td>$^6S_{5/2}$</td>
<td>7.44</td>
</tr>
<tr>
<td>26</td>
<td>Fe</td>
<td>iron</td>
<td>[Ar] 4s²³d⁶</td>
<td>$^5D_4$</td>
<td>7.87</td>
</tr>
<tr>
<td>27</td>
<td>Co</td>
<td>cobalt</td>
<td>[Ar] 4s²³d⁷</td>
<td>$^4F_{9/2}$</td>
<td>7.86</td>
</tr>
<tr>
<td>28</td>
<td>Ni</td>
<td>nickel</td>
<td>[Ar] 4s²³d⁸</td>
<td>$^3F_4$</td>
<td>7.64</td>
</tr>
<tr>
<td>29</td>
<td>Cu</td>
<td>copper</td>
<td>[Ar] 4s³d¹⁰</td>
<td>$^2S_{1/2}$</td>
<td>7.73</td>
</tr>
<tr>
<td>30</td>
<td>Zn</td>
<td>zinc</td>
<td>[Ar] 4s²³d¹⁰</td>
<td>$^1S_0$</td>
<td>9.99</td>
</tr>
<tr>
<td>31</td>
<td>Ga</td>
<td>galm</td>
<td>[Ar] 4s²³d¹⁰⁴p</td>
<td>$^2P_{1/2}$</td>
<td>6.00</td>
</tr>
<tr>
<td>32</td>
<td>Ge</td>
<td>germanium</td>
<td>[Ar] 4s²³d¹⁰⁴p²</td>
<td>$^3P_0$</td>
<td>7.90</td>
</tr>
<tr>
<td>33</td>
<td>As</td>
<td>arsenic</td>
<td>[Ar] 4s²³d¹⁰⁴p³</td>
<td>$^4S_{3/2}$</td>
<td>9.81</td>
</tr>
<tr>
<td>34</td>
<td>Se</td>
<td>selenium</td>
<td>[Ar] 4s²³d¹⁰⁴p⁴</td>
<td>$^3P_2$</td>
<td>9.75</td>
</tr>
<tr>
<td>35</td>
<td>Br</td>
<td>bromine</td>
<td>[Ar] 4s²³d¹⁰⁴p⁵</td>
<td>$^2P_{3/2}$</td>
<td>11.81</td>
</tr>
<tr>
<td>Z</td>
<td>Symbol</td>
<td>Element</td>
<td>Configuration</td>
<td>Term</td>
<td>Ionization potential (eV)</td>
</tr>
<tr>
<td>---</td>
<td>--------</td>
<td>-------------</td>
<td>---------------</td>
<td>-----------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>36</td>
<td>Kr</td>
<td>krypton</td>
<td>[Ar] 4s²3d¹⁰4p⁶</td>
<td>¹S₀</td>
<td>14.00</td>
</tr>
<tr>
<td>37</td>
<td>Rb</td>
<td>rubidium</td>
<td>[Kr] 5s¹</td>
<td>²S₁/₂</td>
<td>4.18</td>
</tr>
<tr>
<td>38</td>
<td>Sr</td>
<td>strontium</td>
<td>[Kr] 5s²</td>
<td>¹S₀</td>
<td>5.70</td>
</tr>
<tr>
<td>39</td>
<td>Y</td>
<td>yttrium</td>
<td>[Kr] 5s²4d²</td>
<td>²D₃/₂</td>
<td>6.38</td>
</tr>
<tr>
<td>40</td>
<td>Zr</td>
<td>zirconium</td>
<td>[Kr] 5s²4d²</td>
<td>³F₂</td>
<td>6.84</td>
</tr>
<tr>
<td>41</td>
<td>Nb</td>
<td>niobium</td>
<td>[Kr] 5s4d⁴</td>
<td>⁶D₁/₂</td>
<td>6.8</td>
</tr>
<tr>
<td>42</td>
<td>Mo</td>
<td>molybdenum</td>
<td>[Kr] 5s4d⁵</td>
<td>⁷S₃</td>
<td>7.10</td>
</tr>
<tr>
<td>43</td>
<td>Tc</td>
<td>technetium</td>
<td>[Kr] 5s²4d⁵</td>
<td>⁶S₅/₂</td>
<td>7.28</td>
</tr>
<tr>
<td>44</td>
<td>Ru</td>
<td>ruthenium</td>
<td>[Kr] 5s4d⁷</td>
<td>⁵F₅</td>
<td>7.37</td>
</tr>
<tr>
<td>45</td>
<td>Rh</td>
<td>rhodium</td>
<td>[Kr] 5s4d⁸</td>
<td>⁴F₉/₂</td>
<td>7.46</td>
</tr>
<tr>
<td>46</td>
<td>Pd</td>
<td>palladium</td>
<td>[Kr] 4d¹⁰</td>
<td>¹S₀</td>
<td>8.34</td>
</tr>
<tr>
<td>47</td>
<td>Ag</td>
<td>silver</td>
<td>[Kr] 5s4d¹⁰</td>
<td>²S₁/₂</td>
<td>7.58</td>
</tr>
<tr>
<td>48</td>
<td>Cd</td>
<td>cadmium</td>
<td>[Kr] 5s²4d¹⁰</td>
<td>¹S₀</td>
<td>8.99</td>
</tr>
<tr>
<td>49</td>
<td>In</td>
<td>indium</td>
<td>[Kr] 5s²4d¹⁰⁵</td>
<td>²P₁/₂</td>
<td>5.79</td>
</tr>
<tr>
<td>50</td>
<td>Sn</td>
<td>tin</td>
<td>[Kr] 5s²4d¹⁰⁵</td>
<td>³P₀</td>
<td>7.34</td>
</tr>
<tr>
<td>51</td>
<td>Sb</td>
<td>antimony</td>
<td>[Kr] 5s²4d¹⁰⁵</td>
<td>⁴S₃/₂</td>
<td>8.64</td>
</tr>
<tr>
<td>52</td>
<td>Te</td>
<td>tellurium</td>
<td>[Kr] 5s²4d¹⁰⁵</td>
<td>³P₂</td>
<td>9.01</td>
</tr>
<tr>
<td>53</td>
<td>I</td>
<td>iodine</td>
<td>[Kr] 5s²4d¹⁰⁵</td>
<td>²P₃/₂</td>
<td>10.45</td>
</tr>
<tr>
<td>54</td>
<td>Xe</td>
<td>xenon</td>
<td>[Kr] 5s²4d¹⁰⁵</td>
<td>¹S₀</td>
<td>12.13</td>
</tr>
<tr>
<td>55</td>
<td>Cs</td>
<td>cesium</td>
<td>[Xe] 6s²</td>
<td>²S₁/₂</td>
<td>3.89</td>
</tr>
<tr>
<td>56</td>
<td>Ba</td>
<td>barium</td>
<td>[Xe] 6s²</td>
<td>¹S₀</td>
<td>5.21</td>
</tr>
<tr>
<td>57</td>
<td>La</td>
<td>lanthanum</td>
<td>[Xe] 6s²5d</td>
<td>²D₃/₂</td>
<td>5.58</td>
</tr>
<tr>
<td>58</td>
<td>Ce</td>
<td>cerium</td>
<td>[Xe] 6s²4f⁵d</td>
<td>¹G₄</td>
<td>5.47</td>
</tr>
<tr>
<td>59</td>
<td>Pr</td>
<td>praseodymium</td>
<td>[Xe] 6s²4f³</td>
<td>⁴I₀/₂</td>
<td>5.42</td>
</tr>
<tr>
<td>60</td>
<td>Nd</td>
<td>neodymium</td>
<td>[Xe] 6s²4f⁴</td>
<td>⁵I₄</td>
<td>5.49</td>
</tr>
<tr>
<td>61</td>
<td>Pm</td>
<td>promethium</td>
<td>[Xe] 6s²4f⁵</td>
<td>⁶H₅/₂</td>
<td>5.5</td>
</tr>
<tr>
<td>62</td>
<td>Sm</td>
<td>samarium</td>
<td>[Xe] 6s²4f⁶</td>
<td>⁷F₀</td>
<td>5.63</td>
</tr>
<tr>
<td>63</td>
<td>Eu</td>
<td>europium</td>
<td>[Xe] 6s²4f⁷</td>
<td>⁸S₇/₂</td>
<td>5.67</td>
</tr>
<tr>
<td>64</td>
<td>Gd</td>
<td>gadolinium</td>
<td>[Xe] 6s²4f⁷5d</td>
<td>⁹D₂</td>
<td>6.14</td>
</tr>
<tr>
<td>65</td>
<td>Tb</td>
<td>terbium</td>
<td>[Xe] 6s²4f⁹</td>
<td>⁶H₁₅/₂</td>
<td>5.85</td>
</tr>
<tr>
<td>66</td>
<td>Dy</td>
<td>dysprosium</td>
<td>[Xe] 6s²4f¹⁰</td>
<td>⁵I₈</td>
<td>5.93</td>
</tr>
<tr>
<td>67</td>
<td>Ho</td>
<td>holmium</td>
<td>[Xe] 6s²4f¹¹</td>
<td>⁴I₁₅/₂</td>
<td>6.02</td>
</tr>
<tr>
<td>68</td>
<td>Er</td>
<td>erbium</td>
<td>[Xe] 6s²4f¹²</td>
<td>³H₆</td>
<td>6.10</td>
</tr>
<tr>
<td>69</td>
<td>Tm</td>
<td>thulium</td>
<td>[Xe] 6s²4f¹³</td>
<td>²F₇/₂</td>
<td>6.18</td>
</tr>
<tr>
<td>70</td>
<td>Yb</td>
<td>ytterbium</td>
<td>[Xe] 6s²4f¹⁴</td>
<td>¹S₀</td>
<td>6.25</td>
</tr>
<tr>
<td>Z</td>
<td>Symbol</td>
<td>Element</td>
<td>Configuration</td>
<td>Term</td>
<td>Ionization potential (eV)</td>
</tr>
<tr>
<td>----</td>
<td>--------</td>
<td>-------------</td>
<td>--------------------</td>
<td>----------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>71</td>
<td>Lu</td>
<td>lutetium</td>
<td>[Xe] 6s²4f¹⁴5d</td>
<td>²D₃/₂</td>
<td>5.43</td>
</tr>
<tr>
<td>72</td>
<td>Hf</td>
<td>hafnium</td>
<td>[Xe] 6s²4f¹⁴5d²</td>
<td>³F₂</td>
<td>7.0</td>
</tr>
<tr>
<td>73</td>
<td>Ta</td>
<td>tantalum</td>
<td>[Xe] 6s²4f¹⁴5d³</td>
<td>⁴F₃/₂</td>
<td>7.89</td>
</tr>
<tr>
<td>74</td>
<td>W</td>
<td>tungsten</td>
<td>[Xe] 6s²4f¹⁴5d⁴</td>
<td>⁵D₀</td>
<td>7.98</td>
</tr>
<tr>
<td>75</td>
<td>Re</td>
<td>rhenium</td>
<td>[Xe] 6s²4f¹⁴5d⁵</td>
<td>⁶S₅/₂</td>
<td>7.88</td>
</tr>
<tr>
<td>76</td>
<td>Os</td>
<td>osmium</td>
<td>[Xe] 6s²4f¹⁴5d⁶</td>
<td>⁵D₄</td>
<td>8.7</td>
</tr>
<tr>
<td>77</td>
<td>Ir</td>
<td>iridium</td>
<td>[Xe] 6s²4f¹⁴5d⁷</td>
<td>(⁴F₉/₂)</td>
<td>9.1</td>
</tr>
<tr>
<td>78</td>
<td>Pt</td>
<td>platinum</td>
<td>[Xe] 6s⁴f¹⁴5d⁸</td>
<td>³D₃</td>
<td>9.0</td>
</tr>
<tr>
<td>79</td>
<td>Au</td>
<td>gold</td>
<td>[Xe] 6s⁴f¹⁴5d¹⁰</td>
<td>²S₁/₂</td>
<td>9.23</td>
</tr>
<tr>
<td>80</td>
<td>Hg</td>
<td>mercury</td>
<td>[Xe] 6s²4f¹⁴5d¹⁰</td>
<td>¹S₀</td>
<td>10.44</td>
</tr>
<tr>
<td>81</td>
<td>TI</td>
<td>thalium</td>
<td>[Xe] 6s²4f¹⁴5d¹⁰₆p</td>
<td>²P₁/₂</td>
<td>6.11</td>
</tr>
<tr>
<td>82</td>
<td>Pb</td>
<td>lead</td>
<td>[Xe] 6s²4f¹⁴5d¹⁰₆p²</td>
<td>³P₀</td>
<td>7.42</td>
</tr>
<tr>
<td>83</td>
<td>Bi</td>
<td>bismuth</td>
<td>[Xe] 6s²4f¹⁴5d¹⁰₆p³</td>
<td>⁴S₃/₂</td>
<td>7.29</td>
</tr>
<tr>
<td>84</td>
<td>Po</td>
<td>polonium</td>
<td>[Xe] 6s²4f¹⁴5d¹⁰₆p⁴</td>
<td>³P₂</td>
<td>8.42</td>
</tr>
<tr>
<td>85</td>
<td>At</td>
<td>astatine</td>
<td>[Xe] (6s²4f¹⁴5d¹⁰₆p⁵)</td>
<td>²P₃/₂</td>
<td>9.5</td>
</tr>
<tr>
<td>86</td>
<td>Rn</td>
<td>radon</td>
<td>[Xe] 6s²⁴f¹⁴5d¹⁰₆p⁶</td>
<td>¹S₀</td>
<td>10.75</td>
</tr>
<tr>
<td>87</td>
<td>Fr</td>
<td>francium</td>
<td>[Rn] 7s</td>
<td>²S₁/₂</td>
<td>4.0</td>
</tr>
<tr>
<td>88</td>
<td>Ra</td>
<td>radium</td>
<td>[Rn] 7s²</td>
<td>¹S₀</td>
<td>5.28</td>
</tr>
<tr>
<td>89</td>
<td>Ac</td>
<td>actinium</td>
<td>[Rn] 7s²6d</td>
<td>²D₃/₂</td>
<td>6.9</td>
</tr>
<tr>
<td>90</td>
<td>Th</td>
<td>thorium</td>
<td>[Rn] 7s²6d²</td>
<td>³F₂</td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>Pa</td>
<td>protactinium</td>
<td>[Rn] (7s²5f²6d)</td>
<td>(⁴K₁₁/₂)</td>
<td></td>
</tr>
<tr>
<td>92</td>
<td>U</td>
<td>uranium</td>
<td>[Rn] 7s²5f³6d</td>
<td>⁵L₆</td>
<td>4.0</td>
</tr>
<tr>
<td>93</td>
<td>Np</td>
<td>neptunium</td>
<td>[Rn] 7s²5f⁴6d</td>
<td>⁶L₁/₂</td>
<td>5.8</td>
</tr>
<tr>
<td>94</td>
<td>Pu</td>
<td>plutonium</td>
<td>[Rn] 7s²5f⁶</td>
<td>⁷F₀</td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>Am</td>
<td>americium</td>
<td>[Rn] 7s²5f⁷</td>
<td>⁸S₇/₂</td>
<td>6.0</td>
</tr>
<tr>
<td>96</td>
<td>Cm</td>
<td>curium</td>
<td>[Rn] 7s²5f⁷6d</td>
<td>⁹D₂</td>
<td></td>
</tr>
<tr>
<td>97</td>
<td>Bk</td>
<td>berkelium</td>
<td>[Rn] 7s²5f⁸6d</td>
<td>⁸H₁₇/₂</td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>Cf</td>
<td>californium</td>
<td>[Rn] 7s²5f¹₀</td>
<td>⁵I₈</td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>Es</td>
<td>einsteinium</td>
<td>[Rn] 7s²5f¹₁</td>
<td>⁴I₁₅/₂</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>Fm</td>
<td>fermium</td>
<td>[Rn] (7s²5f¹₂)</td>
<td>(³H₆)</td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>Md</td>
<td>mendelevium</td>
<td>[Rn] (7s²5f¹₃)</td>
<td>(²F₁/₂)</td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>No</td>
<td>nobelium</td>
<td>[Rn] (7s²5f¹₄)</td>
<td>(¹S₀)</td>
<td></td>
</tr>
<tr>
<td>103</td>
<td>Lw</td>
<td>lawrencium</td>
<td>[Rn] (7s²5f¹₄6d)</td>
<td>(²D₃/₂)</td>
<td></td>
</tr>
</tbody>
</table>
4.2.7.3. **Pauli exclusion principle.** It is found that particles having half-integral spin values, such as electrons and protons \((s = 1/2)\), are fermions. The statement that the wave function of a system of identical fermions must be totally antisymmetric in the combined space is the generalized version of the Pauli exclusion principle. We use this principle when we have to deal with the electron configurations of atoms that contain more than one electron. The principle, then, reduces to its original form that Pauli discovered in 1925; that is, no two electrons in an atom can have the same set of four quantum numbers [3, 65]. According to this principle, if two electrons are in the same orbital, which means they have the same principle, orbital, and orbital magnetic quantum numbers, they must have different spin magnetic quantum numbers.

Let us start writing electron configurations of the elements from the one with the smallest atomic number, hydrogen. Since hydrogen has only one electron and \(1s\) has the orbital that has the lowest energy level, the single electron is shown in this orbital with an arbitrary spin orientation. The orbital diagram for hydrogen then will be as follows:

\[
H \quad \begin{array}{c}
\uparrow \\
1s^1
\end{array}
\]

Since \(L = 0, S = 1/2, \) and \(J = 1/2\) the ground term value for hydrogen is \(^2S_{1/2}\). The following element with the smallest atomic number is helium. Now we have two electrons to fill into the orbitals. Since \(1s\) orbital has the lowest energy level with the capacity of two electrons, both of the electrons must be in this orbital. That is, these electrons have the same principle, orbital, and orbital magnetic quantum numbers. According to the exclusion principle their spin magnetic quantum numbers must be different, that is \(1/2\) and \(-1/2\). So, the orbital diagram for helium will be as follows:

\[
He \quad \begin{array}{c}
\uparrow \downarrow \\
1s^2
\end{array}
\]
The ground term value for helium is $^1S_0$.

4.2.7.4. Diamagnetism and paramagnetism. Paramagnetic substances are those that are attracted by a magnet, and diamagnetic substances are those that are slightly repelled by a magnet [3]. Using these definitions as a general rule, it can be said that any atom with an odd number of electrons must be paramagnetic, since there are the same number of electrons with $m_s = 1/2$ as of electrons with $m_s = -1/2$. On the other hand, atoms containing an even number of electrons may be either paramagnetic or diamagnetic. According to these definitions, hydrogen is paramagnetic and helium is diamagnetic. These are in accordance with the experimental results, which verify correctness of the exclusion principle.

Now let us find out the electron configuration of the lithium atom with three electrons: Since $1s$ orbital can only have two electrons, its third electron should be in the next lowest energy level orbital, $2s$. The orbital diagram for lithium atom then is:

![Orbital Diagram](Li_orbital_diagram.png)

As one can see from the diagram, the lithium atom is also paramagnetic. The ground term value for lithium is $^2S_{1/2}$, since it has one electron outside of the closed shell.

4.2.7.5. Shielding effect. While configuring electrons for the lithium atom, we placed the third electron on $2s$ orbital, since it has the lowest energy level after $1s$ orbital. The reason that $2s$ orbital has a lower energy level than $2p$ orbital can be explained by the shielding effect. It can be also used to explain the rest of the ordering of the orbital energy levels shown in Fig. 40c.

$2s$ and $2p$ orbitals are larger than the $1s$ orbital. $1s$ electrons partially shield the electrostatic attractive forces of nucleus on the electrons of $2s$ and $2p$ orbitals. On the other hand, the density for the $2s$ electron near the nucleus is greater than that for the $2p$ electron.
For this reason, the 2s orbital is said to be more penetrating than the 2p orbital. 2p electron, therefore, is effected by this shielding more than 2s electron.

The electron configurations for beryllium (Z = 4) and boron (Z = 5) are 1s\(^2\)2s\(^2\) and 1s\(^2\)2s\(^2\)2p\(^1\), respectively. As one can expect from their orbital diagrams below, beryllium is a diamagnetic, and boron is a paramagnetic element:

\[
\begin{align*}
Be & & \quad \begin{array}{c} \hline \hline \uparrow \downarrow \end{array} & \hline \hline \uparrow \\
1s^2 & & 2s^2
\end{align*}
\]

\[
\begin{align*}
B & & \quad \begin{array}{c} \hline \hline \uparrow \downarrow \end{array} & \hline \hline \uparrow \downarrow & \hline \hline \uparrow \hline \hline \hline \hline \\
1s^2 & & 2s^2 & 2p^1
\end{align*}
\]

4.2.7.6. *Hund’s rule.* If we start constructing an orbital diagram for carbon atom with Z = 6, we need to make a decision about how to place two electrons of 2p orbitals seen in the electron configuration of 1s\(^2\)2s\(^2\)2p\(^2\). We know that there is no difference in terms of orbital energy levels among 2px, 2py, and 2pz orbitals. We also know from the Pauli exclusion principle that we cannot have two electrons in one orbital with the same spin configuration. Considering these knowns, we have three choices to place two 2p electrons: We can place them in one of three orbitals with opposite spin directions, or we can place them in two different orbitals with same or opposite spin directions. Hund’s rule helps us to make this decision. This rule states that the most stable arrangement of electrons in subshells is the one with the greatest number of parallel spins. Considering this rule, we have only one way of placing the two 2p electrons of carbon atom:

\[
\begin{align*}
B & & \quad \begin{array}{c} \hline \hline \uparrow \downarrow \end{array} & \hline \hline \uparrow \downarrow & \hline \hline \uparrow & \hline \hline \hline \hline \\
1s^2 & & 2s^2 & 2p^2
\end{align*}
\]
Considering these principles, rules, and effects orbital diagrams of nitrogen \((Z = 7)\), oxygen \((Z = 8)\), fluorine \((Z = 9)\), and neon \((Z = 10)\) with respective electron configurations 
\[1s^22s^22p^3, 1s^22s^22p^4, 1s^22s^22p^5, \text{ and } 1s^22s^22p^6\] 
are given as follows:

\[
\begin{array}{ccc}
N & \uparrow\downarrow & \uparrow\downarrow \\
 & 1s^2 & 2s^2 \\
 & & 2p^3 \\
\end{array}
\]

\[
\begin{array}{ccc}
o & \uparrow\downarrow & \uparrow\downarrow \\
 & 1s^2 & 2s^2 \\
 & & 2p^4 \\
\end{array}
\]

\[
\begin{array}{ccc}
f & \uparrow\downarrow & \uparrow\downarrow \\
 & 1s^2 & 2s^2 \\
 & & 2p^5 \\
\end{array}
\]

\[
\begin{array}{ccc}
Ne & \uparrow\downarrow & \uparrow\downarrow \\
 & 1s^2 & 2s^2 \\
 & & 2p^6 \\
\end{array}
\]

### 4.3. Noisy Communicational Channels Representing Atomic Orbitals

Our approach to the atom, just like any other system, is to consider it as a communication channel that communicates with its environment through other communication channels. Here, we present a way to model atomic orbitals as noisy communication channels by representing them as Fourier series that plot error content graphs of noisy communication channels on complex plane.

Let us get started by remembering the Fourier series equation for noisy communication channel error content graphs from Section 3.2.4:

\[
u(t) = \sum_{i=-\infty}^{\infty} e^{i(q+nl)t} \frac{e^{i(q+nl)t}}{(q+nl)^2}.
\] (4.62)
By observing the quantum numbers table (Table 5) from Section 4.2.5.1 and the parameters of (4.62), one can see the mathematical relation between these two sets of numbers. Let us use this relation and replace $n, p,$ and $l$ parameters of (4.62) with principle quantum number $n$, orbital quantum number $\ell$, and orbital magnetic quantum number $m_{\ell}$, respectively. Then (4.62) would become

$$u_{\ell}^{\prime} (t) = \sum_{m_{\ell} = -\ell}^{\ell} \frac{e^{j(q+nm_{\ell})t}}{(q + nm_{\ell})^2}.$$ (4.63)

An obvious interpretation of this equation would suggest that it represents a subshell depending on the value of $\ell$. Note that each component of (4.63) corresponds to an $m_{\ell}$ value. If we continue with our interpretation, each of these components would represent an orbital. Let us name these components as $u_{m_{\ell}} (t)$:

$$u_{m_{\ell}} (t) = \frac{e^{j(q+nm_{\ell})t}}{(q + nm_{\ell})^2}.$$

By extending the same analogy, we can represent the $n$th shell by taking a summation over $\ell$, from 0 to $n - 1$. Then, this series could be shown as:

$$u_{n}^{\prime} (t) = \sum_{\ell=0}^{n-1} \sum_{m_{\ell}=-\ell}^{\ell} \frac{e^{j(q+nm_{\ell})t}}{(q + nm_{\ell})^2}.$$

**Example 34.** Let us find Fourier series representations for shell $N$ ($n = 4$) and one of its subshells $4d$ ($\ell = 3$), and complex function representations for subshell $4d$’s seven orbitals ($m_{\ell} \in [-3,3]$.) For this example, $q$ is taken as 1. Fourier series $u_{4}^{\prime} (t)$ of shell the $N$ is:

$$u_{4}^{\prime} (t) = \sum_{\ell=0}^{3} \sum_{m_{\ell}=-\ell}^{\ell} \frac{e^{j(1+4m_{\ell})t}}{(1 + 4m_{\ell})^2},$$

Fourier series $u_{3}^{\prime} (t)$ of the subshell $4d$ is:

$$u_{3}^{\prime} (t) = \sum_{m_{\ell}=-3}^{3} \frac{e^{j(1+4m_{\ell})t}}{(1 + 4m_{\ell})^2},$$
and the complex functions $u_{-3}(t)$, $u_{-2}(t)$, $u_{-1}(t)$, $u_0(t)$, $u_1(t)$, $u_2(t)$, and $u_3(t)$ of 4d's seven orbitals are:

$$e^{-31\mu t/121}, \quad e^{-7\mu t/49}, \quad e^{-3\mu t/9}, \quad e^{\mu t}, \quad e^{5\mu t/25}, \quad e^{9\mu t/81}, \quad e^{15\mu t/169}$$

respectively. These functions, plotted on complex plane, are shown in Fig. 41.

![Fig. 41](image_url)

4.4. Summary

In the representation of atomic model by Fourier series of uniform polygons, each atomic orbital is represented by a component of Fourier series (a complex function.) Each
atomic subshell is represented by a Fourier series of a uniform polygon, and each atomic shell is represented by another Fourier series (actually this also corresponds to a Fourier series of a uniform polygon, as can be seen in the examples shown in Appendix A) that is the sum of all Fourier series representing its subshells.

Here we present not only a way to represent atomic energy levels with exponential components and Fourier series; we are also connecting it to a framework where they have a representations in other analysis toolsets. For instance, we can also represent atoms with roots of unity, Chebyshev polynomials and matrices, or permutations and permutation matrices. We implemented the Fourier series using numerical and symbolic calculation algorithms, then combined them with the other representations of the framework into a Mathematica implementation. This implementation is given in Appendix A and provides a look at all of these representations together.
CHAPTER 5
CONCLUSION

5.1. Summary and Conclusions

This dissertation presents an information theoretical framework to model and/or analyze systems. Along with the advancement of technology, the demand for such modeling approaches to better analyze, consequently to better understand, systems that are natural or otherwise ever getting complex products of service industry has been on the rise [69]. There is an coordinated effort to describe such complex systems more clearly and encourage research for future complex products. As part of these efforts, National Science Foundation has identified Cyber-Physical Systems (CPS) as a key area of research. The CPS are defined as systems that are composed of tightly combined and coordinated physical and computational elements. The CPS are distinguished from embedded systems by their design as a network of communicating components with physical input and output, rather than as standalone devices. Hence, by definition, the CPS are much closer to natural systems in which this communication network among components with physical input and output works seamlessly.

As design of complex systems similar to CPS and natural systems are becoming alike, this dissertation suggests to approach modeling and analysis of such systems by utilizing the noisy communication channel concept of mathematically rigorous information theory. According to this approach, just as it is suggested for CPS, components of a system, as well as the entire system itself, are considered to have a set of inputs and a set of outputs, where the inputs are considered as source signals of a corresponding communication channel, and the outputs as signals interpreted by that communication channel’s receiver.
The modeling process will vary by the particular system being modeled; however, we suggest a modeling process similar to mathematical modeling process. As illustrated in Fig. 42, starting from a natural or produced (or to be produced) real-world system, this process consists of the development of a communication channel model, analysis by rigorous mathematics of the information theory and the framework of analysis tools presented, interpretation of the analysis results, validation against the real world system, and adjustment to the model and repetition of analysis and validations until the results are satisfactory.

Fig. 42. The modeling process.

The dissertation also presents a framework of relationships (Fig. 43) founded on a uniform polygon representation of the noisy communication channel model consisting of complex analysis tool-set based on complex roots, combinatorial analysis tool-set based on a subset of permutations, Chebyshev analysis tool-set based on eigenvalues of Chebyshev matrices, and Fourier analysis tool-set based on Fourier series representation of regular and irregular uniform polygons. In addition to providing a vast set of tools to analyze noisy communication channel models for systems, this framework also presents a new perspective and greatly expanded set of tools to study science and engineering problems that are currently being approached only by one of the above-mentioned analysis tool-sets. Because of the established common mathematical notation the framework is founded on and the set of relationships it comprises, the use of mathematical tools that belong to
an analysis tool-set that has not been considered for a particular scientific or engineering problem becomes considerably easier.

A chapter has been dedicated to understand the current modeling approaches of a natural system, the atom, and to offer a noisy communication channel perspective as a showcase of our approach. Applications areas of the proposed modeling and analysis framework would be diverse. The atomic model case study introduced here is just an example, where a real-world system is being modeled to further investigate its behavior and interactions with its environment. This case study presents a different perspective to model atom in which its energy levels are represented in the forms of Fourier Series and plots of those series on the complex plane. Only one modeling of atom has been included to prove the concept; however, considering other tool-sets available in the framework, it is implied that many other models can also be produced.

5.2. Future Work

As we have indicated in the introduction of this study, we approach natural and other systems from information theoretical perspective. We consider the study of natural processes and also the study of complex systems as the study of communication systems, as illustrated in Fig. 44. This dissertation only presents one aspect of information theoretical
study of natural processes; however, there is much more room left for researchers to use our approach as a starting point and explore how other information theoretical concepts can also help to provide rigorous mathematical tools to explain the natural and complex systems.

Fig. 44. Study of natural processes as communication systems.

There is an increased interest in transdisciplinary approaches to address the multidimensional nature of many of today’s research topics in sciences, engineering, and medicine. National Science Foundation’s Cyber-Physical Systems (CPS) research program mentioned in the previous section is one example for this increased interest. National Health Institute Director’s High Risk-High Reward program, where the goal is to fundamentally change the way we think about biomedical research, is another. One example for such research planning is given in Appendix B. As illustrated in Fig. 44, we feel that all of these natural systems and other complex systems, such as CPS, can benefit from collaborative research efforts utilizing information theoretical approaches.

The elements of the relationships framework and various representations of concepts used (complex roots, eigenvalues of Chebyshev matrices, Fourier series, etc.) connect
this framework to the powerful analysis tool-sets that they are part of (complex analysis, Fourier analysis, graph theory, etc.) This means that this framework is expandable by introducing other relationships among various concepts from different analysis tool-sets. One example could be establishing a relationship between entropy or channel capacity from information theory and fast Fourier transformation from Fourier analysis using an already established relationship between information theory and Fourier analysis.

This study suffices by presenting relationships that connect each of the analysis tool-sets to noisy communication channels based on uniform polygons, and it briefly touches on other direct relationships between an analysis tool-set and another (e.g., primitive permutation matrices of combinatorial analysis and $n$th roots of unity of complex analysis.) We believe that further investigation would lead to additional such relationships, eventually mapping the analysis tool-sets in all possible ways. Furthermore, combinations of relationships (similar to the combined relationships of special set of permutations, uniform polygons, and eigenvalues of Chebyshev matrices presented in Chapter 3) can be published to provide a complete set of relationships as a useful handbook.

Exploring relationships between the framework introduced and quantum information theory in general, and quantum computation, quantum communication, and quantum neural networks in particular, could potentially benefit the relatively young research topic of quantum information theory.

There are a couple of specific topics in graph theory such as perfect shuffle and stable sets that could be potentially useful to include in the framework. Another topic that could be useful to explore is the arc duration partition functions.

Providing relationships for the introduced framework is one of the main contributions of this dissertation. At the same time, it provokes further study to include other components of each of the analysis tool-sets and, more importantly, other aspects of well established information theory. For instance, depending on the system under investiga-
tion, concepts of entropy, channel capacity, minimum code distance, etc., may proven to be useful quantitative indicators [46].

In general, any topic considered where one or more of the analysis tool-sets included in the framework is commonly used may benefit from either a new information theoretical perspective or from concepts of one of the other analysis tool-sets included in the framework that has never been considered that particular topic. For example, it is known that Chebyshev polynomials are used in experimental design, approximation theory, and interpolation theory. Noisy communication channel concept or concepts from Fourier analysis may be beneficial to apply to these areas.
REFERENCES


APPENDIX A
MATHEMATICA IMPLEMENTATION OF THE NOISY COMMUNICATION CHANNEL FRAMEWORK
Many implementations, both numerical discrete implementations in Matlab and symbolic continuous implementations Mathematica, have been used throughout this study, and some of them are available in technical reports [62, 67, 63, 42]. In order to avoid repetition and to be able to represent the entire framework in a single output, all of the implementations have been merged into one Mathematica implementation. The `Manipulate` function of Mathematica is utilized to avoid having to present many pages of figures for different parameters.

First we will describe each function used in `Manipulate`, then we will include the `Manipulate` function itself. We will conclude with sample output screen-shoots.

This defines the $\chi_k$ function introduced in Definition 44:

\[
a[k, j, n, q] := \text{Mod}[j - 1 + q (k - 1), n] + 1;
\]

Now depending on $n$ being odd or even (Corollary 1) we find Chebyshev polynomials of second kind $u_p(x)$ and $v_p(x)$ using `ChebychevU` function of Mathematica:

\[
\text{ChebPol}[n, x] := \text{If}[\text{Mod}[n, 2] == 0,
\text{Expand}[\text{ChebyshevU}[n/2 - 1, x]],
\text{Expand}[\text{ChebyshevU}[(n - 1)/2, x] + \text{ChebyshevU}[((n - 1)/2 - 1), x]]];
\]

Depending on $n$ being odd or even (Corollary 1) we find Chebyshev matrices $U_p(x)$ and $V_p(x)$ based on Remarks 51 and 52:

\[
\text{UVp}[n, x] := \text{If}[\text{Mod}[n, 2] == 0,
\text{Table}[\text{Switch}[j - i, -1, 1, 0, 2 x, 1, 1, _, 0], {i, n/2 - 1}, {j, n/2 - 1}],
\text{ReplacePart}[\text{Table}[\text{Switch}[j - i, -1, 1, 0, 2 x, 1, 1, _, 0], {i, (n - 1)/2}, {j, (n - 1)/2}], (2 x + 1), {(n - 1)/2, (n - 1)/2}]];
\]

The functions below are eigenvalue equation based on Theorems 13 and 14, eigenvalues based on sorted roots of the eigenvalue equation, the eigenvalue for $q$ of $\{nq\}$, and length of a uniform polygon side for $q$ of $\{q\}$, respectively:

\[
\text{eigens}[n] := \text{Re}[\text{Eigenvalues}[\text{UVp}[n, 1]]];
\]
\[
\text{eigensSorted}[n] := \text{Sort}[\text{N}[\text{eigens}[n]], (#2 > #1) \&];
\]
\[
\text{eigenq}[n, q] := \text{eigensSorted}[n][[q]];
\]
\[
\text{lq}[n, q] := \text{N}[\text{Sqrt}[\text{eigenq}[n, q]]];
\]

This function finds the permutations that correspond to this uniform polygon according to Definition 44:
findPerForUniPol[n_, q_, init_] := Module[{per = Range[0], perr = Range[0], d = 1, c = 1},
  If[q > 1,
   For[d = init, d <= init + (GCD[n, q] - 1) (q - 1) (q + 1), d = d + (q - 1) (q + 1),
    perr = Range[0];
    For[c = 1, c <= n/GCD[n, q], c++,
      AppendTo[perr, a[c, Mod[d, n], n, q]];]
    ];
   per = Join[per, perr];
  ];
  per = Range[n];
  per]

This function finds the adjacency matrix of a permutation according to Section 3.1.4:

findAdjMatForPer[n_, per_] := Module[{matr2 = Table[i*j*0, {i, 2 n}, {j, 2 n}], pos = 0, o = 1, posNextVal = 1},
  For[o = 1, o <= n, o++,
   pos = Position[per, o][[1, 1]]; posNextVal = If[pos == n, per[[1]], per[[pos + 1]]];
   matr2[[o, o + n]] = 1;
   matr2[[o, posNextVal + n]] = 1;
  ];
  matr2]

The first function finds the \( S(q) \) set, which is used in Theorem 10, and the second function finds whether a permutation represents an \( n \)-queens solution based on Theorems 9 and 10 and Remark 49:
The first function below generates one more permutation using Definition 8 if the polygon is compound, and the second function applies Theorem 9 to see if it corresponds to a solution:

```mathematica
findPerExtr[n_, q_] := Module[{per2 = Range[0], perr2 = Range[0], f = GCD[n, q], g = 1},
If[q > 1,
For[f = GCD[n, q], f >= 1, f--,
perr2 = Range[0];
For[g = 1, g <= n/GCD[n, q], g++,
AppendTo[perr2, a[g, f, n, q]]; ];
per2 = Join[per2, perr2]; ];
, per2 = Range[0]; ];
per2 ];
permExtr[n_, q_] := If[GCD[n + 1, q] == 1 && (Mod[n + 1, 6] == 5 ||
Mod[n + 1, 6] == 1) , True , "N/A" ];
```

This is used to draw a chessboard and place queens based on a given permutation:

```mathematica
PermutationToBoard[p_List] /; p =!= {} := Module[{n = Length[p]},
Normal[ SparseArray[ Thread[Thread[{Range[n], #}] -> 1] &[p], {n, n} ] ]
PermutationToBoard[p___] := {}
```
These draw orbital, subshell, and shell Fourier series representations of the atomic energy levels, respectively. The function `subshell` is also used to draw Fourier series representation of uniform polygons:

```math
\text{orbital}[n_, q_, l_, m_] := \text{Exp}[I (q + n*m) t] / (q + n*m)^2;

\text{subshell}[n_, q_, l_] := \sum_{m=-l}^{n-l} \text{Exp}[I (q + n*m) t] / (q + n*m)^2;

\text{shell}[n_, q_] := \sum_{l=0}^{n-1} \sum_{m=-l}^{l} \text{Exp}[I (q + n*m) t] / (q + n*m)^2;
```

The following is the function and it is documented with the code:

```math
\text{Manipulate}\[
\text{Grid}[\{\text{Grid}[\{\text{Labeled}[\text{System'CirculantGraph}[n, \{q\}, \text{VertexLabels -> "Name"},
\text{ImagePadding -> 10}],
\{\text{Text}[\{n, q\}], \text{Text}[\text{Style["Uniform Polygon", Bold]}]\}, \{\text{Bottom, Top}\}],
\text{Grid}[\{\text{Grid}[\{\text{Grid}[\{\text{Labeled}[\text{findPerForUniPol}[n, q, init],
\text{Text}[\text{Style["Permutation\ starting from " <> \text{ToString[init]}, \text{Bold},
\text{TextAlignment -> Center]], Top}],
\{\text{Labeled}[\text{permis}[n, q, init],
\text{Text}[\text{Style["Is the permutation\ a solution?", Bold, \text{TextAlignment -> Center]}], Top}\}
\}]\],
\text{Labeled}[\text{Graphics}[\{\text{ArrayPlot}[#, \text{Mesh -> True}, \text{ColorRules -> \{1 -> Black\}}, \text{ImageSize -> \{400, 400\}}][[1]]\]} &
\{\text{Apply}[\text{PermutationToBoard}, \text{Block}[\{\text{pp = Permutations[Range[n]}],
\{\text{Reverse[ findPerForUniPol[n, q, init]] \} \} ] ],
\text{Text}[\text{Style["Chessboard Representation\ of the permutation", Bold, \text{TextAlignment -> Center]}], Top},
\text{Labeled}[\text{Show}[\text{AdjacencyGraph}[\text{findAdjMatForPer[n, findPerForUniPol[n, q, init]]},
\text{VertexLabels -> Join[Table[i -> i, \{i, n\}], Table[i + n -> i, \{i, n\}]},
\text{ImagePadding -> 10}, \text{DirectedEdges -> True,}
\text{EdgeShapeFunction -> GraphElementData["FilledArrow", "ArrowSize" -> .1]],
\text{VertexCoordinates -> Join[Reverse[Table[\{0, i/2\}, \{i, n\}]],
\text{Reverse[Table[\{1, i/2\}, \{i, n\}]\}]]}],
\text{Text}[\text{Style["Shannon Representation of\ Noisy Communication Channel",
\text{Bold, \textAlignment -> Center]}}], \text{Top}\}]
\}\}
\}
\]
```

153
If $\gcd(n, q) = q \land q \neq 1$,
Grid[
{Grid[{Grid[
{Labeled[findPerExtr[n, q],
Text[Style["One extra \npermutation for \nthis compound \n\ polygon starting \nfrom q = " <> ToString[q], Bold, TextAlignment -> Center]],
Top}], {Labeled[permisExtr[n, q],
Text[Style["\nIs the permutation \n\ a solution?", Bold, TextAlignment -> Center]],
Top]}
{Labeled[Graphics[
{ArrayPlot[#, Mesh -> True, ColorRules -> {1 -> Black}, ImageSize ->
{400, 400}]][1]
]} &
(Apply[PermutationToBoard, Block[{pp = Permutations[Range[n]]},
{Reverse[findPerExtr[n, q]]}]]]
Text[Style["", Bold, TextAlignment -> Center]], Top],
Labeled[Show[
AdjacencyGraph[findAdjMatForPer[n, findPerExtr[n, q]],
VertexLabels -> Join[Table[i -> i, {i, n}], Table[i + n -> i, {i, n}]],
ImagePadding -> 10, DirectedEdges -> True,
EdgeShapeFunction -> GraphElementData["FilledArrow", "ArrowSize" ->
.1]],
VertexCoordinates -> Join[Reverse[Table[{0, i/2}, {i, n}]],
Reverse[Table[{1, i/2}, {i, n}]]]],
Text[""]
}]
}}, Frame -> True}
,
Grid[
{Text["No extra permutation, since this isn’t a compound \n\ polygon"]}
], Frame -> True
]

}]

Labeled[Grid[
{Labeled[ChebPol[n, x],
Text[Style[ToString[If[Mod[n, 2] == 0, "u_" <> ToString[n/2 - 1], "v_"
<> ToString[(n - 1)/2]], OutputForm] <> " (x) =", Bold, TextAlignment ->
Right]], Left]},
{Labeled[ChebPol[n, 1],
Text[Style[ToString[If[Mod[n, 2] == 0, "u_" <> ToString[n/2 - 1], "v_"
<> ToString[(n - 1)/2]], OutputForm] <> " (1) =", Bold, TextAlignment ->
Right]], Left]},
{Labeled[MatrixForm[UVp[n, x]],
Text[Style[ToString[If[Mod[n, 2] == 0, "U_" <> ToString[n/2 - 1], "V_"
<> ToString[(n - 1)/2]], OutputForm] <> " (x) =", Bold, TextAlignment ->
Right]], Left]}

}]

154
Chebyshev Properties

ChebPol[n, (1 - \[Lambda]/2)]

\[U_\frac{n}{2} = \text{ChebPol}[n, (1 - \frac{\Lambda}{2})] \quad (1-\frac{\Lambda}{2}) =
\]

\[V_{\frac{n-1}{2}} = \text{ChebPol}[n, (1 - \frac{\Lambda}{2})] \quad (1-\frac{\Lambda}{2}) =
\]

\[u_{\frac{n}{2} - 1} = \text{ChebPol}[n, (1 - \frac{\Lambda}{2})] \quad (1-\frac{\Lambda}{2}) =
\]

\[v_{\frac{n-1}{2}} = \text{ChebPol}[n, (1 - \frac{\Lambda}{2})] \quad (1-\frac{\Lambda}{2}) =
\]

Eigenvalues =

\[\lambda_q = \text{eigenq}[n, q] \quad (1) =
\]

\[l_q = \text{lq}[n, q] \quad (1) =
\]
Finally, three sample output screen-shots of this Mathematica implementation are shown in Fig. 45, 46, and 47.
Fig. 45. Mathematica output screen-shot for the 5 by 2 error content graph.
Fig. 46. Mathematica output screen-shot for the 6 by 2 error content graph.
Fig. 47. Mathematica output screen-shot for the 9 by 4 error content graph.
APPENDIX B

EARLY DIAGNOSIS OF GLAUCOMA
In this appendix, we will describe an ongoing research for early diagnosis of an eye disease called glaucoma (IRB: X080130004.) Supported by the National Institutes of Health and National Eye Institute, the Glaucoma Clinic at the Eye Foundation Hospital of the University of Alabama at Birmingham conducts one of the leading projects on this research. Some of the related publications produced include [70, 71, 72, 73].

This research involves data collection from various sources and analysis of the data at hand to achieve a method that helps researchers better understand the disease and its early diagnosis. This analysis is mainly involved with pattern recognition, classification, and progression. Methods used for the analysis include but are not limited to statistical analysis methods (i.e., regression, variance, probability distribution) and machine learning methods (i.e., neural networks, expert systems) [74, 75, 76, 77]. Some of the difficulties to mention that arise regarding the analysis are the variety of data sources, large quantity of data parameters, and inherent uncertainty in the classification problem.

This analysis has not yet been completed due to some funding issues. Nevertheless, we wanted to include the preliminary and planning parts of it in this appendix as an example of the noisy communication channel analysis approach presented in this dissertation. We briefly describe the collected data, then talk about the analysis methods applied so far and the challenges faced, and finally propose to analyze glaucoma using the noisy communication channel framework.

B.1. Collected Data

Data collected for the glaucoma research consist of the structural data (regarding changes in the structure of optic disc and nerve fiber layer,) functional data, histology and clinical data, and also progression data that tracks change in the other three data types with regards to time [78].

Structural data includes Scanning Laser Polarimetry (GDx), Scanning Laser Tomography (HRT), Optical Coherence Tomography (OCT), and Optic Disc Photography.
GDx consists of two 256 by 128 images (which makes more than 32,000 parameters for raw image analysis) and more than 200 additional parameters, HRT consists of an image and more than 100 parameters, OCT consists of more than 1,000 parameters, and Optic Disc Photograph is a high-resolution image.

Functional data includes two visual field analyses, namely Standard Automated Perimetry (SAP) and Short Wavelength Automated Perimetry (SWAP), and Frequency Doubling Technology Perimetry (FDT). Each of these sources provides data of about 150 parameters.

B.2. Goals, Challenges, and Methods to Overcome Challenges

Analysis is expected to yield to an expert system that, given the collected data described above as input, outputs a glaucoma diagnosis. To do this, the system needs to be trained by the experts of the field and also be provided with data classes and patterns. The classes of data involved in the current research are glaucomatous data set versus normal data set (note that there is uncertainty involved because of the overlapping of these two sets) and also African America data set versus Caucasian data set. The patterns sought in the research are differences between the data classes and a progression pattern in individual patient data that would eventually lead to a progression pattern for the patients with progressing glaucoma.

Challenges for such analysis include a) large quantity of parameters, which requires lengthy pre-processing to eliminate ineffective parameters for the specific classification or pattern recognition problem, b) variety of the data acquired, involving images, thickness maps, functional parameters, etc., and c) inherent uncertainty among classification classes. Because of the challenges listed here, statistical analysis techniques alone would not be sufficient.

Since the analysis is expected to find a structure that is unknown to us, a neural network approach seems useful [74, 75]. Although this approach overcomes some of the
difficulties, some challenges remain, such as the large quantity of data parameters requiring a larger data sample; more memory; much longer processor time, causing slower learning; and also not being able to detect uncertainties, because of its sharp decision boundaries and catastrophic forgetting due to the pattern interference. A relatively recent method called quantum neural network (QNN) addresses some of these difficulties [79, 80].

B.3. Roadmap

Currently, neural networks are developed and applied to only some parts of the data, because of the challenges of applying the analysis to the data as a whole. There exist multiple neural networks to be applied to the same data set in order to determine the most successful one. Receiver Operating Characteristic (ROC) curves are very commonly used tools in medical literature. They display the sensitivity and specificity of a diagnostic test [81] and are used to determine success of a neural network over another. Through the use of ROC curves, results will be compared to the results of published research. Afterwords, it is planned to apply QNN analysis to the same and larger parts of the data and compare results with the results of neural network analysis. Finally, we plan to apply the information theoretical framework of this dissertation to this data and compare the results.
APPENDIX C

IRB APPROVAL FORMS
Form 4: IRB Approval Form
Identification and Certification of Research
Projects Involving Human Subjects

UAB's Institutional Review Boards for Human Use (IRBs) have an approved Federalwide Assurance with the Office for Human Research Protections (OHRP). The Assurance number is FWA00005960 and it expires on September 29, 2013. The UAB IRBs are also in compliance with 21 CFR Parts 50 and 56.

Principal Investigator: OZAYDIN, BUNYAMIN
Co-Investigator(s):
Protocol Number: X080130004
Protocol Title: An Information Theoretical Modeling and Analysis of Systems

The IRB reviewed and approved the above named project on 4/1/11. The review was conducted in accordance with UAB's Assurance of Compliance approved by the Department of Health and Human Services. This Project will be subject to Annual continuing review as provided in that Assurance.

This project received EXPEDITED review.
IRB Approval Date: 4/1/11

Date IRB Approval Issued: 4/1/11

HIPAA Waiver Approved?: Yes

Marilyn Doss, M.A.
Vice Chair of the Institutional Review Board for Human Use (IRB)

Investigators please note:

The IRB approved consent form used in the study must contain the IRB approval date and expiration date.

IRB approval is given for one year unless otherwise noted. For projects subject to annual review research activities may not continue past the one year anniversary of the IRB approval date.

Any modifications in the study methodology, protocol and/or consent form must be submitted for review and approval to the IRB prior to implementation.

Adverse Events and/or unanticipated risks to subjects or others at UAB or other participating institutions must be reported promptly to the IRB.
Waiver of Informed Consent and/or Waiver of Patient Authorization

Approval of Waiver of Informed Consent to Participate in Research. The IRB reviewed the proposed research and granted the request for waiver of informed consent to participate in research, based on the following findings:
1. The research involves no more than minimal risk to the subjects.
2. The research cannot practically be carried out without the waiver.
3. The waiver will not adversely affect the rights and welfare of the subjects.
4. When appropriate, the subjects will be provided with additional pertinent information after participation.

Check one:
- [ ] and Waiver of Authorization (below)
- [ ] or Waiver of Authorization (below)
- [ ] Waiver of Authorization not applicable

Approval of Waiver of Patient Authorization to Use PHI in Research. The IRB reviewed the proposed research and granted the request for waiver of patient authorization to use PHI in research, based on the following findings:
1. The use disclosure of PHI involves no more than minimal risk to the privacy of individuals
   i. There is an adequate plan to protect the identifiers from improper use and disclosure.
   ii. There is an adequate plan to destroy the identifiers at the earliest opportunity consistent with the conduct of the research, unless there is a health or research justification for retaining the identifiers or such retention that is otherwise required by law.
   iii. There is an assurance that the PHI will not be reused or disclosed to any other person or entity except as required by law, for authorized oversight of the research study, or for other research for which the use or disclosure of PHI would be permitted.
2. The research cannot practically be conducted without the waiver or alteration.
3. The research cannot practically be conducted without access to and use of the PHI.

—OR—

Full Review
The IRB reviewed the proposed research at a convened meeting at which a majority of the IRB was present, including one member who is not affiliated with any entity conducting or sponsoring the research, and not related to any person who is affiliated with any such entities. The waiver of authorization was approved by the majority of the IRB members present at the meeting.

Date of Meeting

Signature of Chair, Vice-Chair or Designee

Date

 Expedited Review
The IRB used an expedited review procedure because the research involves no more than minimal risk to the privacy of the individuals who are the subject of the PHI for which use or disclosure is being sought. The review and approval of the waiver of authorization were carried out by the Chair of the IRB, or by one of the Vice-Chairs of the IRB as designated by the Chair of the IRB.

Date

Signature of Expedited Review

Date

The University of Alabama at Birmingham
701 20th Street South
205.934.3780
Fax 205.934.1301
info@uab.edu

Page 1 of 1